

الجمهورية الجزائرية الديمقراطية الشعبية

République Algérienne Démocratique et Populaire

Ministère de l'enseignement supérieur et de la recherche scientifique

Université de 8 Mai 1945 – Guelma -

Faculté des Mathématiques, d'Informatique et des Sciences de la matière

Département d'Informatique



Mémoire de Fin d'études Master

Filière : Informatique

Option : Ingénierie des Médias

14/867

Thème :

Interface Gestuelle : Détection, suivi Bidimensionnel et reconnaissance de gestes pour les interfaces Homme-Machine augmentées

Encadré Par :

Mr. HALLACI Samir

Présenté par :

Mohamed Zine Labidine Sahraoui

Zakarya Seridi

Juin 2014

Dédicace

Au premier lieu Ahmadou Allah qui ma donner la puissance, le courage pour terminer mes étude et réaliser ce mémoire de fin d'étude, puis je dédié ce modeste travaille qui est l'effort de plusieurs année d'étude à tous les personnes qui sont proche à mon cœur.

A ma plus cher mère grâce à leur amour, soutien, sacrifices que dieu la protéger.

A mon père que dieu le donnée les grand paradis.

A Ma grand-mère Touha que dieu la protéger, et mon grand-père Alhadje dieu a son mène.

A Mon frère Amine qui mon encourage beaucoup dans tous les années d'étude, et ma sœur sara pour leur aide et les conseils dans tous le parcours d'étude.

A mes oncles : Kfali Hocine, Zineb, Mila, Bahia, Biba, Nouha, Tounse, Hamza, Abd alnoure, Kfali abde alrahman.

A mon encadreur monsieur Hallaci samir Je le remercie pour sa disponibilité, ses encouragements, son soutien moral, sa sympathie et son amitié sincère dont il a fait preuve le long de ce travail.

A mon binôme hamada et ça famille

A tous mes amies qui je ne pe zakarya

ux citer mes surtout : Amir, Titou, Abd alali, Hocine « gandoula », Haroun, Ali, bibik, dieu a son mène, Koukou, Ridha, Radwane.

A toute la département de l'informatique.

Zakarya Seridi

Dédicaces

*Je dédie ce modeste travail à très cher Mère
pour leur encouragement, leur soutien, leur
compréhension et leur tendresse et leurs
sacrifices qui m'ont servi de stimulateur
essentiel.*

*Je dédie également à mes chers frères et sœurs
qui m'ont toujours assisté.*

Sans oublier mes amis de promotions « IM »

*A tous mes amis et mes collègues et à tous les
gens que je compte pour eux*

A tous ceux qui m'aiment

Mohamed

Résumé :

L'interaction homme-machine gestuelle constitue un axe de recherche particulièrement actif avec un potentiel applicatif important. Le canal gestuel est un moyen de communication particulièrement riche et la main un objet articulé complexe. Cependant, L'interprétation automatique de gestes basée sur la vision par ordinateur offre de nouvelles possibilités d'interaction avec l'ordinateur, plus naturelles et intuitives qu'avec les périphériques classiques.

Dans ce contexte, notre travail consiste à remplacer le fonctionnement d'un écran tactile ou d'une souris par un système de vision avec une caméra. Ainsi, le problème fondamental est de suivre, en temps réel, le mouvement de la main et des doigts à partir de leurs projections dans les images, avant d'en reconnaître leurs postures.

Les contraintes industrielles qui ont guidé nos travaux nous ont orienté vers une approche par apparence, avec des hypothèses réduites afin que le système soit peu contraignant pour l'utilisateur. Les différentes étapes abordées concernent la détection de la main basée sur la couleur de peau, l'extraction de caractéristiques invariantes, le suivi du mouvement et la comparaison de descripteurs de forme pour la reconnaissance de postures 2D.

Enfin, La robustesse du système sera confrontée aux conditions réelles d'une démonstration d'un monde réelle.

Mots clés : Reconnaissance de geste, Interaction Homme-Machine (IHM), Segmentation, Extraction de caractéristiques, Classification.

Sommaire

Sommaire :	1
Liste des figures :	4
Liste des tableaux :	5
Liste des schémas :	6
Liste des abréviations :	7
Introduction générale :	8
Chapitre I reconnaissance de gestes:	9
1. Introduction :	9
2. Le geste :	9
2.1. Les gestes de la main :	9
2.1.1. La fonction ergotique :	10
2.1.2. La fonction épistémique :	10
2.1.3. La fonction sémiotique :	10
2.2. Taxonomie des gestes de Quek :	10
3. Utilisation des gestes :	10
3.1. Reconnaissance de langage des signes :	13
3.1.1. Composition d'un signe :	14
3.1.2. Problèmes de gestes :	14
3.1.2.1. Complexité du vocabulaire :	15
3.1.2.2. Relation entre les deux mains :	15
3.2. Réalité virtuelle :	15
3.3. Réalité augmentée :	15
4. Les outils de reconnaissance d'un geste:	16
4.1. Les gants de données :	17
4.2. Écrans tactiles :	18
4.3. La Camera :	18
4.4. Kinect :	19
5. Système de reconnaissance de geste :	21
6. Type de système de reconnaissance de geste :	21
6.1. Système de reconnaissance de geste avec lien physique :	21
6.2. Système de reconnaissance de geste sans lien physique (basé vision) :	21
7. Reconnaissance de la main :	22
7.1. Approches basées apparence :	22
7.2. Approches basées modèle 3D :	23
8. Détection de visage :	24
9. Domaine d'application :	25

9.1. Reconnaissance de la langue des signes :	26
9.2. Reconnaissance des nombres :	26
9.3. Le contrôle des robots :	26
9.4. Contrôle de la télévision :	26
9.5. Contrôle d'éditeur graphique :	26
10. Les travaux existants :	27
11. Conclusion :	27
Chapitre II système de reconnaissance de gestes basé vision:	28
1. Introduction :	28
2. Segmentation :	29
2.1. Types de segmentation :	29
2.1.1. Seuillage d'Otsu :	29
2.1.2. Segmentation par différence image :	30
2.1.2.1. Différence d'images successives :	30
2.1.2.2. Soustraction du fond :	30
2.1.3. Segmentation par région :	31
2.1.4. Segmentation par contour :	31
2.1.5. Segmentation par couleur :	32
3. Détection de la couleur de peau :	33
3.1. Méthodes paramétriques :	33
3.2. Méthodes non paramétriques :	33
4. Les espaces colorimétriques :	34
4.1 Espace de couleur HSV :	34
4.2 Espace de couleur RGB :	35
4.3 Espace de couleur YCbCr :	36
4.4. HSCbCr :	37
4.5. Les travaux existants :	38
5. Post des traitements :	39
5.1 Morphologie mathématique :	39
5.1.1. Dilatation :	39
5.1.2. Erosion :	40
5.1.3. Ouverture :	40
5.1.4. Fermeture :	41
5.2. Les résultats des travaux existants (post de traitement) :	41
6. Extraction des caractéristiques :	42
6.1. Carte de distance :	42
6.2. L'enveloppe Convexe :	43

6.2.1. Quickhull :	44
6.3. Les travaux voisins pour les extractions des caractéristiques :	45
7. Classification :	46
7.1. Classification basé Apprentissage :	46
7.1.1. Machines à Vecteurs Supports(SVM) :	46
7.1.2. Les réseaux de neurones artificiels :	48
7.1.3. Méthode de k plus proche voisin :	49
7.1.4. Model de Markov cache :	50
7.2. Classification basé règle :	51
7.3. Les travaux existantes pour la classification :	51
8. Conclusion :	51
Chapitre III conception et implémentation	53
1. Introduction :	53
Partie I : Conception :	54
I.1. Architecture générale de notre système :	55
I.2. Segmentation :	56
I.2.1. Détection de la couleur de peau :	56
I.2.2. Filtrage :	58
I.2.3. Étiquetage des objets :	59
I.2.4. Élimination des objets non-concerne :	60
I.3. Extraction des caractéristiques:	61
I.3.1. Détection de visage :	61
I.3.2. Extraction des caractéristiques morphologiques :	62
I.3.2.1. Centre de la main :	63
I.3.2.2. Détection de la paume :	65
I.3.2.3. Détection de poignet :	65
I.3.2.4. Détection des bouts des doigts :	66
I.3.2.5. Détection des creux :	68
I.4. Classification:	68
I.5. taux de reconnaissance :	72
Partie II: Implémentation	75
II.1. Environnement de travail :	75
II 2. Description de l'application :	76
II.3. Testes et résultats :	78
II. Conclusion :	79
Conclusion générale:	81
Bibliographie:	82

Liste des figures

Figure I.1 : Taxonomie des gestes.....	10
Figure I.2: exemple de geste ayant un sens différent selon leur culture	12
Figure I.3 : Exemples de gestes statiques et dynamiques	12
Figure I.4 : Exemple de configuration, d'orientation et mouvement de la main	14
Figure I.5 : La phrase, " MOI GARÇON ", constituée des signes [MOI] et [GARÇON].....	14
Figure I.6: système qui présente la réalité augmentée.....	16
Figure I.7: Exemples de périphériques d'entrée.....	17
Figure I.8: représentation d'un gant numérique	18
Figure I. 9: gant coloré	19
Figure I. 10: Kinect	20
Figure I. 11: Utilisation de kinect pour des jeux de vidéo	20
Figure II. 1 : Segmentation par région	31
Figure II. 2: Segmentation par couleur.....	33
Figure II. 3: Différentes présentation de la couleur de peau	34
Figure II. 4: Présentation des couleurs dans l'espace HSV	35
Figure II. 5: Présentation des couleurs dans l'espace RGB.	36
Figure II. 6: Les résultats du model HSCbCr.....	37
Figure II. 7: Exemple de dilatation. A gauche : l'image originale. A droite : sa dilatée.	40
Figure II. 8: Exemple d'érosion. A gauche : l'image originale. A droite : son érodée.	40
Figure II. 9: A gauche : image original A milieu : ouverture. A droite : fermeture.....	41
Figure II. 10 : Détection du centre avec la carte de distance.....	43
Figure II. 11 :L'enveloppe convexe.	43
Figure II.12:L'enveloppe convexe de la main.....	43
Figure II.13 :P1, P2 deux points avec Xmin et Xmax.....	44
Figure II.14:p3 plus loin point.....	44
Figure II. 15: Obtention de l'enveloppe convexe.....	45
Figure II. 16: exemple de classification des gestes avec SVM.	47
Figure II. 17: Représentation d'un réseau de neurone avec 2 entrees 2 neurones 1 sortie	48
Figure II.18 : Exemple de classification avec KNN (K=3).....	50
Figure III. 1 : résultat de la phase d'étiquetage	60
Figure III. 2 : largeur et la hauteur d'un objet.....	60
Figure III. 3 : résultat de la détection de visage par la transformé de Hough.	62
Figure III. 4: résultat de détection de centre avec la boite englobante.....	63
Figure III. 5 : résultat de détection de centre de la main (point rouge).	64

Figure III. 6: résultat de détection de la paume	65
Figure III. 7 : exemple illustratif de détection de poignet	65
Figure III. 8: résultat de détection de poignet	66
Figure III. 9: détection des bouts des doigts par la courbure du contour	66
Figure III. 10: les pixels détectés comme des bouts des doigts.....	67
Figure III. 11: résultat final de détection des bouts des doigts.....	67
Figure III. 12: détection des creux entre les doigts	68
Figure III. 13: résultat de détection des creux (point noir).....	68
Figure III. 14: notre approche de reconnaissance a base des règles.....	69
Figure III. 15: le changement de l'angle lors de la rotation de la main.....	70
Figure III. 16: l'image à gauche (angle juste), image droite (angle erronée).....	70
Figure III. 17: calcule de l'angle de rotation par le poignet.....	70
Figure III. 18: le résultat de la rotation inverse de la main	71
Figure III. 19: résultat de reconnaissance de geste.....	71
Figure III. 20 : taux de reconnaissance cas 1	73
Figure III. 21: taux de reconnaissance cas 2	73
Figure III. 22: taux de reconnaissance cas 3	74
Figure III. 23 : Interface de QT.....	76
Figure III. 24:l'interface de notre application.	76
Figure III. 25 : fenêtre principale quand l'utilisateur clique sur démarrer	78
Figure III. 26: carte météo.....	79
Figure III. 27: Paint	79

Liste des tableaux

Tableau I.1 : les applications des travaux existants au plusieurs domaines	27
Tableau II. 1: les espaces couleurs utilisent pour la détection de peaux.	38
Tableau II.2: Les résultats obtenus avec des déférents espaces de couleur.	38
Tableau II. 3: les résultats de près traitement des travaux voisins	41
Tableau II. 4: les résultats obtcnus pour l'extraction des caractéristiques	45
Tableau II. 5: les résultats obtiennent par les travaux voisins.....	51
Tableau III. 1 : comparaison entre les espaces colorimétrique.	57
Tableau III. 2 : les différents résultats de filtrage.....	59
Tableau III. 3 : les objets qui restent après l'élimination.....	61
Tableau III. 4: symboles des gestes.....	72
Tableau III. 5: taux de reconnaissance cas 1.	72

Tableau III. 6: taux de reconnaissance cas 2	73
Tableau III. 7: taux de reconnaissance cas 3	74
Tableau III. 8: les gestes pour remplacer la souris et clavier.	77

Liste des schémas

Schéma I. 1:Les deux modèles de geste	22
Schéma I. 2: Étapes de reconnaissance du geste basé apparence.....	23
Schéma I. 3: Reconnaissance du geste basé modèle 3D	24
Schéma II. 1 : Etapes de reconnaissance d'un geste	28
Schéma III.1 : Architecture générale de notre système.....	55
Schéma III.2 : les phases pour l'extraction des caractéristiques morphologiques	63

Liste des abréviations :

IHM : Interface Homme- Machine.

LS : Langue des Signes.

SRG : Systèmes de Reconnaissance de Gestes.

ASL : Américaine signe langage.

JSL : japonaise signe langage.

ArSL : Arabe signe langage.

RGB : Red Gren Bleu.

HSV : Hue Saturation Value.

SVM : Support Vector Machine.

ANN : Artificial Neural Networks.

K-PPV(KNN) : K plus porches voisin.

MMC : Model de Markov Cache.

GD : Grand Distance.

GUI : Graphical User Interface (Interface utilisateur graphique).

Open cv : open-source computer vision.

WIMP : Window, Icon, Menu, Pointing device.

PUI : Perceptual User Interface.

Introduction générale :

Lorsque la parole ne suffit plus ou n'est pas adaptée à la situation, l'homme utilise le geste comme moyen de communication alternatif ou complémentaire. Il existe différentes types de communications gestuelles suivant les besoins d'expressivité (gestes de main, visage, corps).

Les gestes de la main sont un canal de communication naturel et intuitif chez l'homme pour interagir avec son environnement. Ils servent à désigner ou à manipuler des objets, à renforcer la parole, ou à communiquer basiquement dans un environnement bruité.

Les gestes de la main et de visage peuvent avoir une signification différente suivant la langue ou la culture par exemple dans la langue des signes chaque pays à sa propre langue de signes, ainsi le geste est l'un des canaux de communications les plus riches dans le domaine des Interfaces Homme-Machine.

Différents périphériques permettent d'interagir avec un ordinateur par l'intermédiaire de la main (souris, gants de données, écrans tactiles).

Notre travail s'inscrit dans une démarche pour remplacer la souris ou clavier de l'ordinateur par la main de l'utilisateur, parce qu'elle a une grande variété de gestes, par ailleurs, les progrès scientifiques et techniques offrent de nouvelles possibilités d'interaction, plus naturelles et intuitives, basées sur le canal gestuel, mais qui n'est pas évident non plus et surtout pas facile à réaliser.

Notre mémoire s'articule autour de trois chapitres principaux, Le premier chapitre a pour but de donner un aperçu sur la reconnaissance de gestes et la détection de visage, puis de présenter les différentes techniques et approches que soit théoriques ou matériels et les différents domaines utilisant le geste comme moyen de communication ou de manipulation.

Dans le deuxième chapitre, nous présentons une étude approfondie sur les différentes méthodes et techniques suivi par un processus de reconnaissance gestuel. L'objectif est de présenter les méthodes adaptées au contexte de notre application.

Enfin, pour le dernier chapitre on présente la conception de notre système en détail, ainsi qu'il sera consacré aux différentes options d'implémentation retenues pour la réalisation de notre système.

Chapitre I : Reconnaissance de geste

Chapitre I : Reconnaissance de gestes

1. Introduction :

Les recherches effectuées dans le domaine de l'Interface homme-machine visent à améliorer les performances des utilisateurs d'un système informatique, non pas en s'intéressant au système lui-même, mais plutôt à son utilisation. La parole et les gestes apparaissent comme un moyen spontané pour une personne de communiquer avec son environnement [1].

Parmi tous les canaux de communications dont dispose l'être humain, le canal gestuel est sans doute le plus riche parce qu'il permet d'agir sur le monde physique d'une part et d'autre part il sert à utiliser le canal d'information. De plus, ce canal fonctionne dans les deux sens, comme un moyen d'émission et de réception d'information.

Les gestes de la main sont un canal de communication naturel et intuitif chez l'homme pour interagir avec son environnement. Ils servent à désigner ou à manipuler des objets aussi 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 [1].

Le but essentiel des systèmes de reconnaissance des gestes et de visage est de créer une interaction naturelle entre l'homme et l'ordinateur où les gestes reconnus peuvent être utilisés pour commander ou donner une information significative sans avoir besoin d'utiliser les dispositifs classiques (Clavier, Souris).

Les possibilités de communication avec les ordinateurs peuvent être aujourd'hui réduites à l'utilisation manuelle ou corporelle des mains, des gestes et des postures. [2].

2. Le geste :

Le geste est souvent défini comme étant un signe manuel ou corporel qui permet d'illustrer les mots du langage, de les compléter ou de les appuyer. Le geste peut aussi être utilisé pour remplacer la parole, c'est le cas pour la langue des signes où les gestes jouent un rôle important dans la communication non-verbale.

D'une manière générale, le geste est assimilé à tous les mouvements d'une partie du corps. Le geste de la main est à la fois un moyen d'action, de perception et de communication. [1]

2.1. Les gestes de la main :

Le geste est composé de trois fonctionnalités : le geste ergotique, le geste épistémique, le geste sémiotique.

Chapitre I : Reconnaissance de gestes

2.1.1. 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.

2.1.2. 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.

2.1.3. La fonction sémiotique :

La main joue le rôle d'organe d'expression pour l'émission d'informations visuelles. Cela comprend la langue des signes, le geste co-verbal, qui accompagne la parole, ou les gestes permettant une communication basique lorsqu'on ne peut pas utiliser la parole, comme dans un environnement bruyé ou en plongée sous-marine. [1]

Dans le cadre de la reconnaissance de gestes pour les IHM, nous nous intéressons plus particulièrement à la fonction sémiotique. Cette fonction est la plus riche et la plus complexe. Elle peut être décomposée en plusieurs catégories, et différentes classifications ont été proposées.

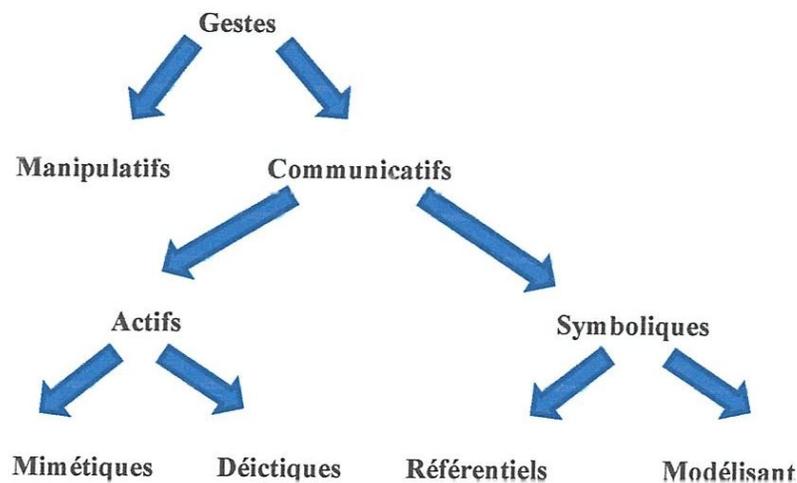


Figure I.1 : Taxonomie des gestes [3].

2.2. Taxonomie des gestes de Quek :

Une classification bien adaptée au domaine de l'interaction homme-machine est la taxonomie de Quek [3], qui décompose les gestes en gestes manipulatifs, correspondant aux fonctions ergotique et épistémique, et en gestes communicatifs, correspondant à la fonction

Chapitre I : Reconnaissance de gestes

sémiotique. Les gestes communicatifs sont décomposés en gestes actifs et en gestes symboliques (*Figure I.1*).

✚ Les gestes symboliques :

Ne sont pas compréhensibles directement, il faut être initié pour comprendre leur signification. Il s'agit par exemple des gestes des langues des signes. Ils sont décomposés en deux :

- **Les gestes référentiels :**

Qui faisant directement référence à un objet ou à un concept (p. ex. le frottement du pouce et de l'index pour évoquer l'argent).

- ✓ Il est possible de faire référence à un objet en montrant sa forme ou son volume qu'il occupe dans l'espace [3].

- **Les gestes modélisant :**

Les gestes de modélisation sont souvent conjoints à d'autres moyens de communication, exemple la parole.

- ✓ Qui modélisent un état ou une opinion et s'emploient souvent en addition à d'autres moyens de communication (p. ex. pour donner une idée de la taille d'un objet).

- **Ambiguïté culturelle :**

Le sens d'un geste peut avoir un sens différent selon la culture et l'origine du communicant, ainsi, le geste repris par la (*Figure I.2*) dans lequel la configuration des doigts représente un O, à un sens différent selon son origine.

Pour un Américain le O signifie good.

✚ Les gestes actifs :

Sont directement liés à leur interprétation et sont utilisés en complément de la parole.

Ils sont décomposés en deux :

- ✓ **Les gestes mimétiques :**

Ils sont souvent exprimés par le mime de l'utilisateur référent, sa position spatiale ou temporelle par rapport à d'autres. Ils sont caractérisés par leur iconicité. Consistant à mimer une action.

- ✓ **Les gestes déictiques (gestes de pointage) :**

Sont très utilisés pour l'interaction homme-machine, car le doigt représente un dispositif de pointage naturel et très pratique. Sont classifiés en trois groupes selon le contexte.

Chapitre I : Reconnaissance de gestes

- Les gestes de pointage spécifiques : sont effectués lorsque le sujet fait allusion à un objet ou un lieu en particulier.
- Les gestes de pointage générique : permettent l'identification d'une classe d'objets en indiquant un de ses représentants.
- Les gestes de pointage métonymiques : l'auteur du geste référence une entité liée à l'objet qu'il pointe [1].

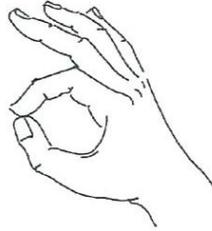


Figure I.2: exemple de geste ayant un sens différent selon leur culture [1]

Il existe deux autres catégories de gestes : les gestes statiques (postures), et les gestes dynamiques.

Il faut distinguer la position et la configuration de la main, en combinant ces deux aspects on obtient la classification proposée par Harling et Edwards [4] :

- ☞ Position statique, configuration statique (les postures).
- ☞ Position statique, configuration dynamique (les gestes dynamiques ou de manipulation).
- ☞ Position dynamique, configuration statique (p. ex. les gestes de pointage).
- ☞ Position dynamique, configuration dynamique (p. ex. la langue des signes).

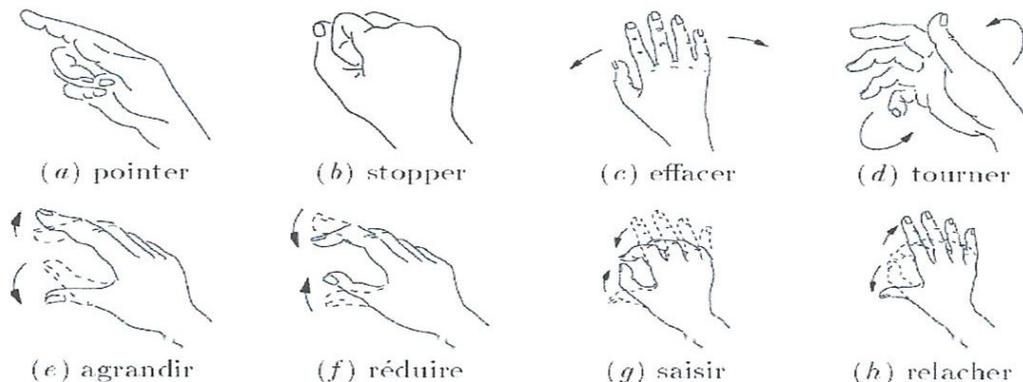


Figure I.3 : Exemples de gestes statiques et dynamiques [1]

3. Utilisation des gestes :

Les trois principales catégories d'applications profitent des débuts de la reconnaissance des gestes. La reconnaissance de langues des signes est un domaine important pour appliquer la reconnaissance des gestes. Deux autres courants de l'interaction homme-machine utilisent des langues de commande gestuelles :

- La réalité virtuelle.
- La réalité augmentée.

Les systèmes de réalité virtuelle plongent les utilisateurs dans un environnement 3D virtuel, on peut l'utiliser dans un domaine éducatif par exemple : la description des phénomènes physiques ou la description de molécules chimiques.

Ces systèmes utilisent les gants numériques et souvent un casque de vision 3D.

Les systèmes de réalité augmentée fusionnent le monde physique réel et le monde électronique et informatisé.

3.1. Reconnaissance de langage des signes :

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

L'expression de phrases en LS ne se réduit pas aux gestes produits par les deux mains et le visage, c'est le corps tout entier qui peut être mis à contribuer pour exprimer une phrase. On peut distinguer trois principales parties qui interviennent : les mains, la tête et la parole, chaque phrase est constituée d'une suite de gestes de la main ou de visage que l'on appelle signes et qui sont agencés suivant une syntaxe régie par une logique spatiale et temporelle.

La reconnaissance automatique de langue des signes est un rapport important pour la communauté des sourds-muets, les systèmes de conversion signes-paroles permettent de discuter avec des entre eux ou bien de proposer des téléphones de la langue des signes [5].

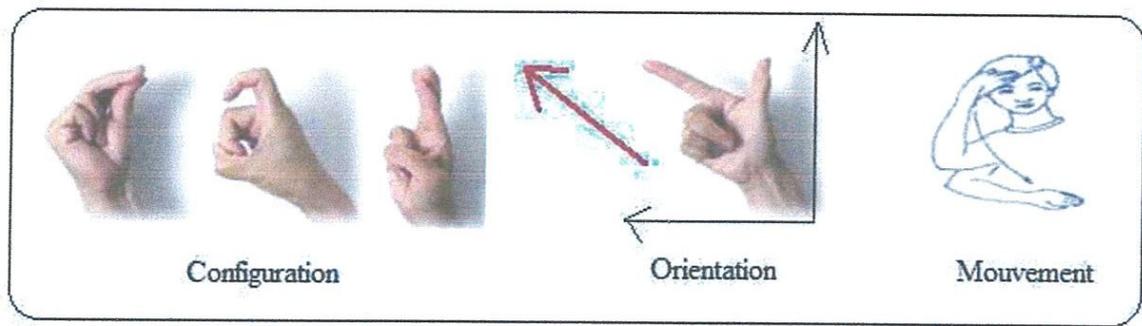


Figure I.4 : Exemple de configuration, d'orientation et mouvement de la main

3.1.1. Composition d'un signe :

En général chaque geste d'une main peut être décomposé en quatre paramètres qui sont indépendants et peuvent être aussi bien dynamiques qu'invariants durant l'émission du signe.

- **La configuration** : Elle correspond à la forme de la main définie par les doigts et la paume, voir l'exemple dans la *Figure I.4*.
- **L'orientation** : Elle est définie par deux axes de la main indiqués dans la *Figure I.4*.
- **Le mouvement** : correspond à la trajectoire décrite par la main (ligne, arc de cercle...), voir *Figure I.4*.
- **L'emplacement** : c'est la position de la main par rapport aux autres organes du corps.

Dans la *Figure I.5*, le premier signe à gauche veut dire [GARÇON] dans la langue des signes française. Le second signe indique que c'est [MOI] aussi dans la langue des signes française.



Figure I.5 : La phrase, " MOI GARÇON ", constituée des signes [MOI] et [GARÇON]

3.1.2. Problèmes de gestes :

Dans un système de reconnaissance de forme classique ou l'en trouve la diversité des signes et l'interaction entre les mains portent beaucoup de problèmes au niveau de la reconnaissance de ces différents gestes, on trouve parmi ces problèmes celui du début et fin d'un geste, tels que le système ou l'application qui ne reconnaît pas si un geste a commencé ou bien a terminé, mais le problème majeur dans tous les systèmes de reconnaissance et la complexité de son vocabulaire, dans le cas d'utilisation des deux mains dans le signe exprimé le problème peut aller plus loin, ce dernier va donc être concentré sur les manières de

Chapitre I : Reconnaissance de gestes

distinguer entre ces différents cas pour ne pas reconnaître un signe à deux mains alors qu'il s'agissait de deux signes distincts effectués au même moment.

3.1.2.1. Complexité du vocabulaire :

Selon [6] la complexité du vocabulaire est l'un des principaux problèmes que l'on trouve en reconnaissance de forme. En effet, si un signe est composé de plusieurs paramètres, le vocabulaire peut être composé d'un très grand nombre de signes.

Dans le cas où une main possède N configurations possibles, M types de mouvements, P emplacements et Q orientations, on a $N*M*P*Q$ signes possibles. Avec la possibilité d'avoir des signes qui combinent les deux mains, on se voit la création d'un système de reconnaissance qui peut devenir très complexe et aussi il nécessite un long apprentissage. En réalité, toutes les combinaisons possibles de paramètres ne sont pas utilisées par les signeurs. Car ils ne sont pas combinés au hasard, mais en fonction de critères syntaxiques et sémantiques. Toutefois le nombre de signes reste important pour remédier à ce problème, il existe une solution qui consiste à traiter séparément les divers composants d'un signe au lieu de traiter le signe dans sa globalité [7].

3.1.2.2. Relation entre les deux mains :

L'interaction entre les deux mains est de trois types, soit les deux mains produisent ensemble un signe, soit une main effectue un signe et l'autre est au repos, soit les deux mains effectuent séparément deux signes de manière plus ou moins indépendante.

Le problème rencontré va être donc de savoir distinguer entre ces différents cas pour ne pas tomber dans un signe à deux mains alors qu'il s'agissait de deux signes distincts effectués au même moment (ou vice-versa).

3.2. Réalité virtuelle :

La réalité virtuelle consiste à plonger l'utilisateur dans un environnement d'images de synthèse. Cette immersion dans un environnement virtuel peut se faire grâce à un casque, ou dans une pièce dédiée munie de plusieurs écrans ou d'un vidéoprojecteur. L'utilisateur est donc en immersion complète dans un environnement dans lequel il peut interagir, notamment par le biais de gestes. [8]

3.3. Réalité augmentée :

La réalité augmentée mélange le monde physique et des informations virtuelles, en surimposant ces informations dans le champ de vision de l'utilisateur.

Chapitre I : Reconnaissance de gestes

Ce type de système est fondé sur une analyse par l'ordinateur du monde environnant l'utilisateur, au moyen d'un système de vision, de type caméra vidéo par exemple, ou de capteurs spécifiques.

Des informations virtuelles sont alors projetées sur des éléments réels, par exemple par l'intermédiaire d'un vidéoprojecteur.

Une des applications principales de la reconnaissance de gestes en réalité augmentée est de rendre une surface interactive : une table, un tableau ou un bureau, sur lesquels des images sont projetées. L'utilisateur peut alors interagir avec des objets réels ou virtuels. Ces systèmes permettent aussi une interaction multi-utilisateurs. [8]



Figure I.6: système qui présente la réalité augmentée [8].

4. Les outils de reconnaissance d'un geste:

La majorité des systèmes d'exploitation repose sur un dispositif de pointage, généralement la souris, qui permet d'interagir avec des éléments graphiques tels que des fenêtres, des icônes et des menus.

On présente : les différents dispositifs permettant une interaction avec l'ordinateur : les périphériques d'entrée (souris et autres dispositifs de pointage), les gants de données, les caméras vidéo et les écrans tactiles.

Pour interagir avec un ordinateur, la souris s'est imposée comme le périphérique d'entrée indispensable. Il existe aussi la boule de commande, ou « trackball » (*Figure I.7*). Le pavé tactile ou « touchpad » est un dispositif de pointage spécifique aux ordinateurs portables, permettant de remplacer la souris. Il s'agit d'une surface sensible de faible dimension, utilisant la capacité électrique.

Pour les jeux vidéo, le périphérique usuel est la manette. Récemment, les constructeurs ont cherché à développer d'autres types d'interactions pareilles aux boutons, par exemple des vibrations pour le retour d'effort, ou des capteurs de mouvements. Ainsi, la Wii mote (*Figure I.7*) de Nintendo a marqué un tournant dans les périphériques de jeux vidéo. Cette manette est

Chapitre I : Reconnaissance de gestes

équipée des capteurs qui lui permettent de se repérer dans l'espace et de retranscrire les mouvements de l'utilisateur à l'écran [8].



Figure I.7: Exemples de périphériques d'entrée [8].

4.1. Les gants de données :

Par rapport à la richesse de l'information véhiculée par des gestes de la main, les possibilités de communication avec les ordinateurs sont réduites avec la souris et le clavier. Des dispositifs spécialisés pour une application sont apparus. Ainsi, pour l'acquisition de données en trois dimensions, des périphériques d'entrée 3d fournissent à l'ordinateur des informations sur la position de la main, voire sur sa configuration pour les plus évolués.

C'est le cas des gants de données (ou gant électronique, gant numérique), qui sont munis de capteurs fournissant la position de la main et les angles des articulations des doigts.

Les gants de données sont utilisés de longue date pour la reconnaissance de la langue des signes et de système de la réalité virtuelle et augmentée, car ils fournissent les positions précises et fiables des articulations de la main. Malheureusement, ces gants sont couteux et sont encombrants, leur utilisation est contraignante pour l'utilisateur.

On présente ci-dessous un type de gant de données sans fil avec ces fonctionnalités et opérations offertes :

- **Le Gant DG5VHand 2.0 :** C'est un gant de données sans fil basé sur la dernière technologie Bluetooth pour une bande passante élevée, le gant V-Hand (V-Main) fonctionne avec une seule pile à la charge de 3.5V - 5V et dispose d'une connectivité jusqu'à la portée de 10 mètres, dans ce gant de données 5 propriétaires capteurs de flexion sont utilisés pour une sensibilité élevée. Le Capteur de Courbure Bi-Flex est un capteur qui modifie la résistance quand il coude, la flexion peut avoir lieu dans les deux sens. V-Hand est une Platform totalement indépendante, le Bluetooth est connecté via le port COM et il lit les commandes et les actions à travers le capteur du gant.

Chapitre I : Reconnaissance de gestes



Figure I.8: représentation d'un gant numérique. [8]

4.2.Écrans tactiles :

Les écrans tactiles combinent à la fois entrée et sortie, avec la visualisation sur un écran et un dispositif de pointage pour interagir directement avec l'information affichée à l'écran. Cette technologie est utilisée pour des écrans de la taille d'un moniteur d'ordinateur, par exemple pour les guichets de billetterie automatique ou pour les assistants personnels digitaux avec éventuellement l'utilisation complémentaire d'un stylet.

Il existe différentes technologies pour les écrans tactiles : capacitive, résistive, infrarouge, à ondes de surface. De manière générale, les écrans tactiles souffrent de différents inconvénients : ils sont onéreux, limités en surface, sujet à l'usure et sensibles aux rayures.

Les écrans tactiles permettent la reconnaissance de gestes simples, appelés gestes de dessins. Par exemple utilisent les gestes de commande. Il existe aussi des alphabets simplifiés pour les permettant la réalisation de lettres d'un seul trait [8].

4.3. La Camera :

L'approche caméra permet de détecter le geste (mains, visage et corps) en analysant le flux vidéo acquit image par image afin de distinguer les régions de la peau du reste.

La détection de la peau dans une image sans hypothèse préalable, car c'est un problème très difficile à résoudre en raison de la très grande variabilité de la forme et la couleur à détecter (Orientation et tailles quelconques, éclairage variable, couleur de peau différente)

Pour simplifier le problème de détection des configurations de la main, mais avec un taux de réalisme plus élevé par rapport à celui offert par les mains et le visage directement (peau). Des marqueurs sont positionnés sur les extrémités des doigts ou bien des gants couleur (voir *Figure I.9*) sont portés par l'utilisateur.

Ces marqueurs sont facilement détectables dans les images par des algorithmes faciles et classiques en vision par ordinateur, la classification sera faite par compter le nombre de

Chapitre I : Reconnaissance de gestes

zones qui contiennent la même couleur et donné le sens approprié au geste fait par la main ou visage[9].

Avec l'augmentation de la puissance des ordinateurs de bureau et l'apparition de caméras bon marché, il est désormais possible de développer des systèmes de reconnaissance des gestes fonctionnant en temps réel.

Les caractéristiques telles que le taux de rafraîchissement ou la résolution varient d'une caméra à une autre. Des valeurs élevées de ces caractéristiques sont avantageuses pour avoir une mise à jour fréquente des images et un niveau de détail important. Toutefois, un compromis est nécessaire, car des valeurs trop importantes augmentent la complexité et le temps de traitement des données.

Par ailleurs, les caractéristiques du capteur (CCD ou CMOS) et de l'optique ont une influence primordiale sur la qualité de l'image obtenue et la sensibilité à l'éclairage [8].

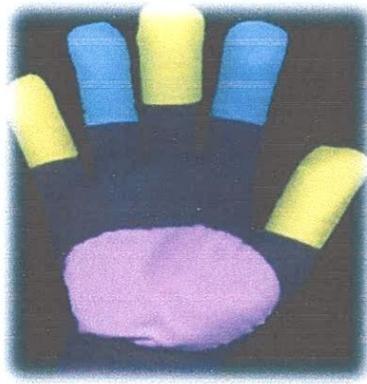


Figure I. 9: gant coloré [8]

4.4. Kinect :

Kinect c'est la dernière technologie de la reconnaissance de geste, son lancement a eu lieu le 10 novembre 2010 par Microsoft, c'est un petit appareil constitué de trois composants : un moteur, un LED, et un microphone, le moteur permet d'incliner kinect le long de l'axe horizontal afin de régler son champ de vision, le LED peut prendre 3 couleurs (rouge, vert et bleu) [10], caractérisé de 4 microphones qui permettent de savoir d'où le son se provient, la caméra permet la capture vidéo dans des différents formats, une caméra infrarouge, c'est elle qui capture la grille et permet de renvoyer des informations de profondeur.

Chapitre I : Reconnaissance de gestes

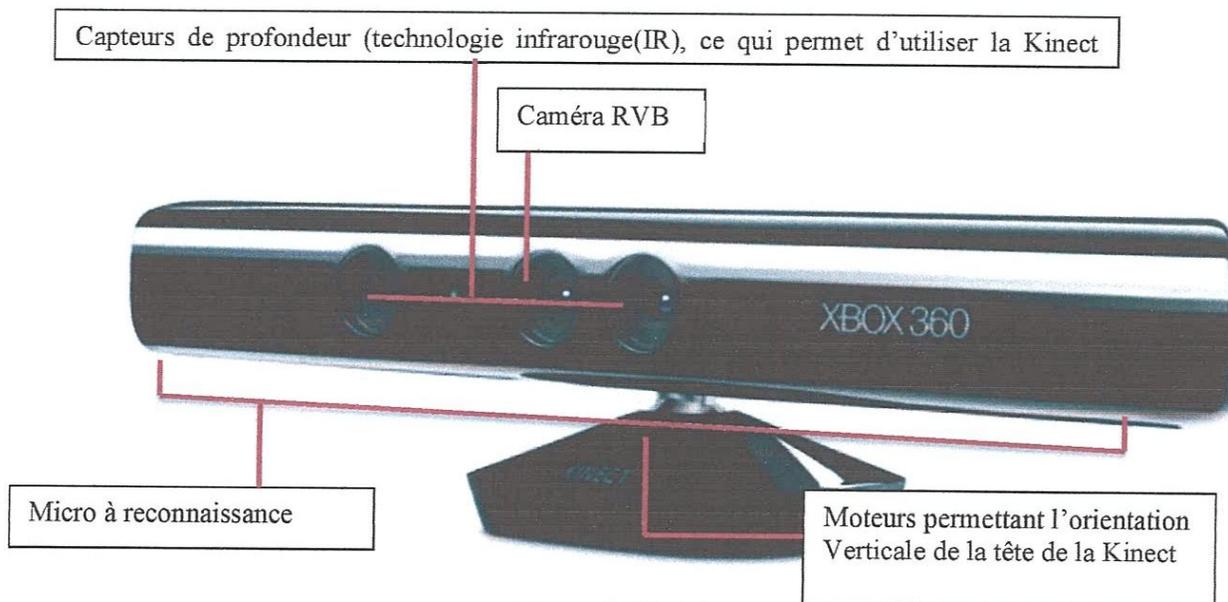


Figure I. 10: Kinect

La combinaison de tout ça permet à Kinect de visualiser les personnes présentes devant l'écran ainsi que leur position physique, même en profondeur. Un logiciel s'occupe ensuite d'interpréter ces données pour les transposer à l'écran, en arrivant ainsi à reproduire le squelette d'un joueur via une vingtaine de points (articulations et extrémités du corps) [10].

Avec kinect y'aurai plus besoin d'utilise une manette ou une souris pour donner une commande, il suffit de tendre notre main devant nous et celle-ci apparaît immédiatement à l'écran sous forme de curseur .les ventes de kinect ont décollé immédiatement avec plus de 10 millions de capteurs Kinect vendus le 11 mars 2011, en décrochant au passage le Guinness

World Record de l'objet d'électronique grand public le plus rapidement vendu de tous les temps.

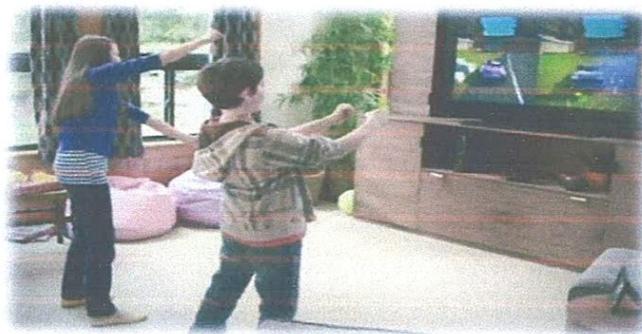


Figure I. 11: Utilisation de kinect pour des jeux de vidéo

5. Système de reconnaissance de geste :

La reconnaissance de geste 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 à proprement parler. L'objectif de base de la reconnaissance des gestes est d'avoir un système automatisé qui permet d'identifier les gestes spécifiques de l'homme et aussi utiliser ces gestes pour contrôler les périphériques ou environnement virtuel [11].

La reconnaissance de geste nous offre un moyen de communication très riche et fort entre les êtres humains et les machines et fournit une interface compatible pour l'homme et nous permet d'interagir avec les machines de façon plus aisée et sympathique.

6. Type de système de reconnaissance de geste :

Tous les systèmes de reconnaissance de gestes (SRG) doivent avoir en entrée des données pour accomplir leur tâche. Différentes technologies sont utilisées pour la collecte de ces données. Selon la technologie utilisée, les systèmes de reconnaissance de gestes peuvent être classifiés en deux catégories : SRG avec lien physique et SRG sans lien physique.

6.1. Système de reconnaissance de geste avec lien physique :

Ce type de système suit une approche instrumentale et utilise un périphérique pour capturer le geste. Cette approche peut détecter exactement la position de la main, les régions des doigts et de la paume, ainsi que leurs orientations et configurations, mais utilise des périphériques externes qui sont très chers et oblige l'utilisateur d'être connecté avec l'ordinateur physiquement, comme les gants quant parle donne : (4.1. Les gants de données). [8]

6.2. Système de reconnaissance de geste sans lien physique (basé vision) :

La vision par ordinateur est la solution la plus naturelle. Elle nous libère d'employer des périphériques, tels que le gant numérique, et se base sur la technique par laquelle nous percevons les gestes.

Elles s'appuient pour cela sur des techniques de vision par ordinateur. Mais cette solution est la plus difficile à mettre en œuvre tout cela revient aux limitations actuelles en vision par ordinateur. [12]

7. Reconnaissance de la main :

Les technologies basées vision utilise une main nue pour extraire les informations nécessaires pour la reconnaissance, ces méthodes sont naturelles et faciles, et l'utilisateur interagit directement avec le système. Il existe plusieurs techniques pour détecter les informations de la main après certaines opérations de prétraitements, ces méthodes peuvent être divisés on deux parties :

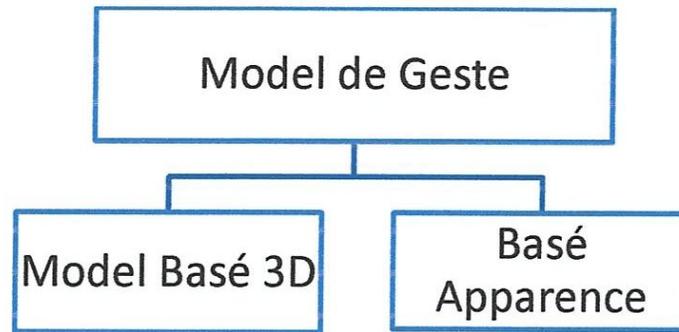


Schéma I. 1: Les deux modèles de geste

7.1. Approches basées apparence :

La majorité des modèles basés sur l'apparence utilisent les paramètres dérivés d'images dans des modèles enregistrés dans une base de données.

Ces paramètres modèles seront comparés avec les paramètres extraits de la vidéo pour définir le sens du geste, ces paramètres peuvent être: contours et des arêtes, des moments d'image, parfois d'autre ces paramètres sont enregistrés dans des modèles sous forme des numéros qui caractérise la main. Beaucoup de ces paramètres sont également utilisés comme éléments dans l'analyse des gestes.

Un autre groupe de modèles utilise positions du bout des doigts comme paramètres. la méthode commune dans cette approche est la détection de la région de la peau dans l'image, mais cette approche aussi à des inconvénients qui sont le changement dans les conditions de la luminance et les objets qui ont la même couleur de la peau dans l'arrière-plan [12].

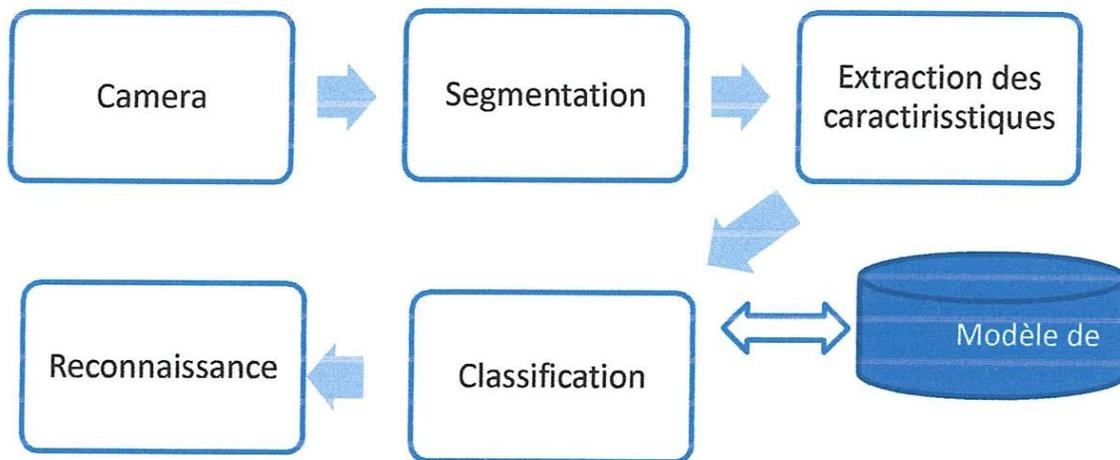


Schéma I. 2: Étapes de reconnaissance du geste basé apparence

7.2. Approches basées modèle 3D :

Les approches basées modèle 3D, une détection de la main avec deux caméras (la vision stéréoscopique) puis une reconstruction 3D (triangulation 3D) en construisant un modèle de main 3D, ensuite ce modèle est projeté sur une image qui sera présentée en une image contour, puis une erreur est calculée selon une fonction erreur [12].

Mais l'inconvénient est que plusieurs caractéristiques de la main peuvent être perdues durant l'opération de projection.

Dans les plus part des systèmes de reconnaissance de geste basé 3D ils utilisent des modèles de main construits avant puis en entre dans la reconnaissance.

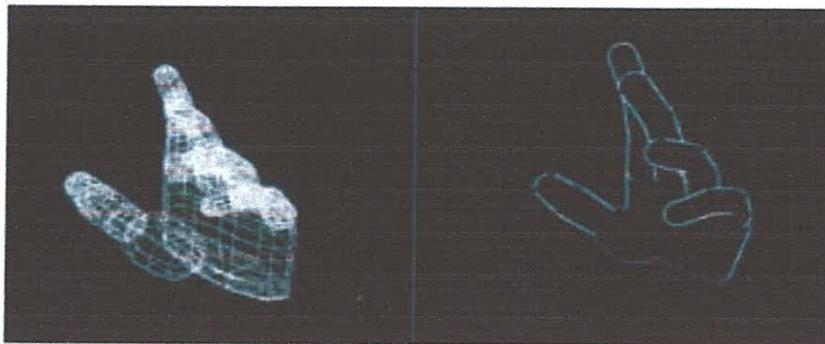


Figure I.12 : Modèle 3D d'une main [8]

Chapitre I : Reconnaissance de gestes

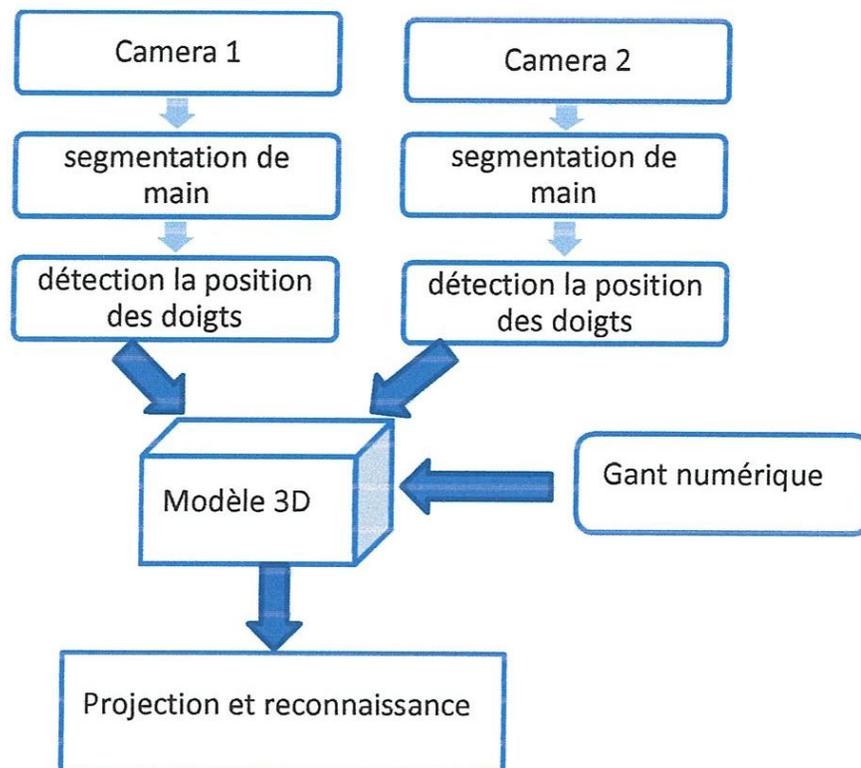


Schéma I. 3: Reconnaissance du geste basé modèle 3D

8. Détection de visage :

Au cours des 40 dernières années, les chercheurs ont commencé à développer des systèmes pour détecter la présence ou l'absence d'un visage dans une image et de la reconnaissance de ce visage, ses chercheurs, dans des disciplines différentes telles que le traitement d'image, la reconnaissance de formes, la vision par ordinateur et la psychologie, ont effectué des recherches sur le traitement du visage [13].

La détection du visage est une tâche préliminaire nécessaire à la plupart des techniques d'analyse du visage, elle est donc un champ d'investigations primordial dans le domaine de la vision par ordinateur et du traitement d'images.

Ses applications sont diverses, citons par exemple, la biométrie, l'Interface Homme machine (IHM), la reconnaissance de gestes comme notre application, la détection de sourires dans les appareils photo récents, les outils de suivi de visages par des webcams motorisées, la recherche de visages dans des grandes bases de données [13].

Les techniques utilisées dans ce domaine sont généralement issues du domaine de la reconnaissance des formes. Exemple de ses techniques :

La transformé de Hough, HaarCascade, Template matching-methods.

Chapitre I : Reconnaissance de gestes

La détection du visage est l'étape préliminaire dans le cadre d'un système de contrôle d'accès à un bâtiment, à une porte d'embarquement ou à un ordinateur.

Plus précisément ses applications peuvent être citées comme suite :

- **En vidéosurveillance :**

Son intérêt est de déterminer la présence d'êtres humains pour surveiller automatiquement des zones sensibles et puis remplacer des systèmes de sécurité actuels comme ceux de clefs, de badges ou de mots de passe.

- **En biométrie faciale :**

Parmi des développements récents dans ce domaine, on trouve des applications de la détection du visage dans la reconnaissance du visage en trois dimensions ou dans l'utilisation d'infrarouges pour analyser la cartographie vasculaire du visage.

- **En études comportementales :**

En couplant à une reconnaissance de geste le cas de notre application, de posture, ou à un suivi de regard, la détection du visage et la reconnaissance de geste permettront des études automatiques pour comprendre les sourds-muets et communiquent avec des personnes normales [13].

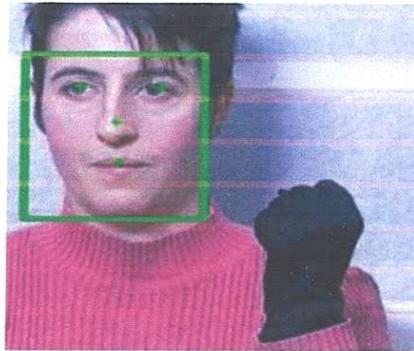


Figure I. 13: détection de visage [14]

9. Domaine d'application :

Avec l'évolution des technologies d'acquisition et des techniques de reconnaissance des gestes, alors la reconnaissance de geste est donc un champ d'investigations primordial dans le domaine de la vision par ordinateur et du traitement d'images. Ses applications sont diverses, citons par exemple, l'Interface Homme machine (IHM), Reconnaissance de la langue des signes, la biométrie, pour la reconnaissance de personnes avec la forme de leur main.

Plus précisément ses applications peuvent être citées comme suite :

Chapitre I : Reconnaissance de gestes

9.1. Reconnaissance de la langue des signes :

La langue des signes a reçu une attention particulière depuis son utilisation dans certains sujets lors de la conversation sur les réseaux sociaux. Beaucoup de systèmes ont été proposés pour la reconnaissance des gestes des différentes langues. Pour la reconnaissance des signes de la langue américaine.

(ASL : Américaine signe langage), la reconnaissance des signes de la langue japonaise (JS L) pour connaître 42 alphabets et 10 mots. Aussi la reconnaissance de la langue arabe(ArSL) [15].

9.2. Reconnaissance des nombres :

Une autre application récente de geste de la main est la reconnaissance des chiffres. Ils ont proposé un système automatique qui pourrait isoler et de reconnaître la signification du geste de mouvement de la main des chiffres arabes de 0 à 9 en temps réel [15].

9.3. Le contrôle des robots :

Contrôler le robot en utilisant des gestes est considéré comme l'une des applications les plus intéressantes dans ce domaine. Ils ont proposé un système qui utilise les signes de numérotation, ce système compte les cinq doigts de la main pour contrôler le robot. Les ordres sont donnés au robot pour effectuer une tâche particulière, où chaque signe a une signification particulière et représente une fonction particulière, par exemple, le chiffre "un" signifie "avancer", "cinq" signifie "stop", et ainsi de suite[15].

9.4. Contrôle de la télévision :

Les Postures et les gestes de la main sont utilisés pour commander sur la télévision.

Donc un ensemble de gestes de la main est utilisé pour contrôler les activités de télévision, telles qu'allumer et éteindre la télévision, augmenter et diminuer le volume, couper le son, et changer le canal en utilisant la main ouverte et fermer [15].

En peut utiliser pour commander aux d'autres machines électroniques telles que les postes radio des voitures.

9.5. Contrôle d'éditeur graphique :

Le contrôle des éditeurs graphiques nécessite que la main doit être localisé et suivi comme une opération de prétraitement, donc ils ont utilisé des gestes dynamiques pour dessiner et éditer des graphes, aussi l'utilisation des formes particulières telle que : des rectangles, triangles, cercle, des lignes horizontales et verticales, et les commandes

Chapitre I : Reconnaissance de gestes

nécessaires pour l'édition graphique sont : copier, supprimer, déplacer, échanger, annuler et fermer [15].

10. Les travaux existants :

Plusieurs chercheurs développent dans le domaine de la reconnaissance des gestes et du visage pour applique à des domaines différents ce tableau exprime les travaux et le domaine qui sera appliqué sur lui :

Travail	Domaine d'application
[8]	Suivi tridimensionnel de la main et reconnaissance de gestes pour les Interfaces Homme-Machine
[14]	Reconnaissance automatique des gestes de la langue française parlée complétée
[24]	Système pour la reconnaissance des signes indiens
[40]	Reconnaissance des signes
[1]	Reconnaissance de gestes en vision par ordinateur
[7]	Reconnaissance de langage des signes américains
[34]	Suivi et de la main et reconnaissance des signes

Tableau I.1 : les applications des travaux existants au plusieurs domaines

11. Conclusion :

Depuis toujours, l'homme rêve de machines intelligentes, capables de dialoguer avec lui, de le comprendre, et de répondre aux questions qu'il s'est toujours posées. Il rêve de pouvoir interagir avec ses machines intelligentes comme avec ses semblables. Dans ce chapitre nous avons présenté diverses facettes de la reconnaissance de gestes. Dans un premier temps on a défini c'est quoi le geste puis on a vu les outils et les approches pour la reconnaissance du geste et finalement quelques domaines d'application sont présentés. Le chapitre suivant présente en détail les différentes méthodes utilisées dans chaque étape du processus de reconnaissance de geste spécifique aux systèmes basés vision.

Chapitre II : Système de Reconnaissance des gestes basé vision

Chapitre II : Système de reconnaissance de gestes basé vision

1. Introduction :

La reconnaissance des gestes est une tâche difficile, sujette à de nombreux travaux de recherche.

Dans une approche basée vision la main est représentée par un ensemble de caractéristiques morphologiques de forme ou de position. Celles-ci peuvent être spatiales et temporelles, ainsi être utilisées et capable de suivre de la reconnaissance des gestes.

Ces informations sont extraites à partir des images en couleurs ou aux niveaux de gris, des images binaires ou du contour, obtenus par une étape de segmentation. La plupart des chercheurs [15] classent les systèmes de reconnaissance de gestes en trois étapes principales qui suivent l'acquisition de l'image d'entrée à partir de la caméra.

Ces étapes sont les suivantes: Prétraitement et segmentation l'extraction des caractéristiques et la classification ou la reconnaissance voir Schéma II.1.



Schéma II. 1 : Etapes de reconnaissance d'un geste

Dans un premier temps, nous traitons le flux vidéo acquis directement de la caméra : les images prises par la camera sont préparées pour les traitements ultérieurs. Cette préparation peut consister en un changement de codage du format des images, l'adaptation de notre système aux conditions de l'éclairage, et en une séparation de la région image contenant les mains. La deuxième étape extrait des caractéristiques de la région segmentée permettant la différenciation des gestes considérés. La troisième et la dernière étape : la classification, sera faite selon des approches différentes.

Dans ce chapitre une étude approfondie sur les différentes méthodes et techniques suivies dans un processus de reconnaissance d'un geste est effectuée. L'objectif est d'imposer le moins de contraintes possibles à l'utilisateur, et détaillé des méthodes adaptées au contexte de notre application. En effet, de telles méthodes reposent généralement sur un certain nombre d'hypothèses sur l'environnement extérieur : l'éclairage, la scène, objets ou sur la couleur du fond. Dans notre travail, il s'agit d'un environnement extérieur, où il y'a le problème de fortes variations de luminosité.

2. Segmentation :

La segmentation est la division de l'image dans des parties et zones selon des méthodes et critères. Dans le cadre de la reconnaissance du geste la segmentation est une approche qui permet de différencier les pixels de peau des autres pixels. Ainsi, on divise l'image d'entrée en régions séparées par des contours: régions blanches (représentent peau) et régions noirs (qui présentent le reste de l'image). Généralement un cadre de sélection est utilisé pour cette séparation. Prenons l'exemple d'une séparation basé sur la couleur de la peau pour l'extraction de la peau, il faut exprimer l'image en d'autres espaces de couleurs extraits à partir de l'espace de couleur de base qui est le RGB, on cite à titre d'exemple HSV ou YCrCb [16], ensuite on définit un intervalle ou un seuil pour chaque composante de HSV ou YCrCb qui caractérise la couleur de peau.

Après l'extraction de la peau on obtient une image binaire bruitée, on procède alors à l'élimination de ce bruit et ce par l'application des morphologiques mathématique sur les parties blanches (qui représente dans notre cas la « peau »).

L'élimination de certains bruits existants dans l'image facilite l'étape suivante qui est l'opération d'extraction des caractéristiques.

Cette section présente les différentes techniques de segmentation, les différents espaces colorimétriques utilisés pour la segmentation basée sur le teint de la peau et les méthodes utilisées pour éliminer le bruit.

2.1. Types de segmentation :

Il existe plusieurs méthodes de segmentation, Nous présentons dans ce qui suit cinq méthodes de segmentation: la première est l'application d'un seuillage la seconde utilise le mouvement pour localiser la main, la troisième utilise la notion de région, la quatrième utilise la notion de frontière et la dernier s'appuie sur la couleur particulière de la peau.

2.1.1. Seuillage d'Otsu :

Le seuillage d'une image en niveaux de gris I consiste à associer la valeur 0 (noir) à tous les pixels dont la valeur est inférieure au seuil α et la valeur 1 (blanc) à tous les autres. On obtient une image seuillée S , aussi appelée masque binaire ou silhouette :

$$\forall(x, y), S(x, y) = \begin{cases} 1 & \text{si } I(x, y) \geq \alpha \\ 0 & \text{sinon} \end{cases} \quad (\text{II.1})$$

L'inconvénient de cette méthode est qu'il faut déterminer le seuil manuellement, de façon empirique, et qu'un seuil donné est adapté à des conditions d'illumination bien précises. or, il est très difficile de trouver un seuil qui donne de bons résultats dans toutes les situations.

Chapitre II : Système de reconnaissance de gestes basé vision

2.1.2. Segmentation par différence image :

La technique de différence d'image est bien connue en vision par ordinateur, elle consiste à soustraire une image par une autre, pixel à pixel, ce qui suppose que la caméra soit fixe afin qu'un pixel de l'image représente toujours le même lieu de l'espace au cours du temps. Cette approche se situe au niveau pixel, ce qui signifie qu'elle ne prend pas en compte les relations qui existent entre des pixels voisins [8].

La valeur des pixels est supposée stable dans le temps. Cependant, les variations de luminosité de la scène peuvent faire varier cette valeur. Ces méthodes sont donc très sensibles aux variations de luminosité, aux ombres, et aux changements du fond. Il faut également prendre en considération le bruit d'acquisition de la caméra.

Il existe deux types de méthodes pour la différence d'images :

- La différence d'images successives, qui détecte les objets en mouvement.
- La soustraction du fond, qui détecte les objets n'appartenant pas au fond.

2.1.2.1. Différence d'images successives :

Avec I_t l'image courante et I_{t-1} l'image précédente, l'image de différence D s'obtient par :

$$D(x, y, t) = |I(x, y, t) - I(x, y, t - 1)| \quad (\text{II.2})$$

2.1.2.2. Soustraction du fond :

La soustraction du fond permet de détecter les objets au premier plan. Le fond correspond à la partie statique de la scène, qui ne change pas au cours du temps. La méthode consiste à soustraire une image de référence I_{ref} , correspondant à la scène sans objet, à l'image courante I_t :

$$D(x, y, t) = |I(x, y, t) - I_{ref}(x, y)| \quad (\text{II.3})$$

L'image de référence est prise à l'initialisation du système, en supposant qu'il n'y a pas d'objet dans la scène. Elle peut aussi être calculée en moyennant les N premières images de la séquence. La soustraction peut être effectuée avec l'image en niveaux de gris, ou en couleur sur chacune des trois composantes RGB. L'image différence est seuillée (par exemple avec la méthode d'Otsu) pour obtenir un masque binaire.

Le problème de cette méthode est sa sensibilité aux variations de luminosité. En effet, si la valeur des pixels correspondant au fond varie de façon trop importante, ces pixels sont détectés comme des objets lors du seuillage. Cette méthode permet de détecter un changement global de luminosité, lent ou rapide, faisant varier le fond de façon importante (lumière allumée ou éteinte).



Figure II. 2: Segmentation par couleur [14]

La détection de la main et de visage dans les images d'un flux vidéo est une problématique importante, commune à la plupart des systèmes de reconnaissance des gestes et de visage. Cette première étape est primordiale, car elle conditionne les résultats de la suite des traitements.

3. Détection de la couleur de peau :

La segmentation de la couleur de peau est une étape préliminaire dans plusieurs applications tel que la détection de mouvement humain, l'interaction homme machine et la vidéosurveillance et dans notre cas la reconnaissance des gestes.

Dans la littérature scientifique il existe plusieurs types de méthodes pour la discrimination entre les pixels peau et non peau. Ces méthodes peuvent être réparties en trois catégories : méthodes paramétriques, non paramétriques, et les méthodes définissant la peau comme cluster explicite [18].

3.1. Méthodes paramétriques :

Le modèle paramétrique gaussien est basé sur l'utilisation d'un modèle gaussien définissant la distribution de la couleur de peau dans un espace de couleur donné. Une étape d'apprentissage permet de calculer les paramètres des gaussiennes, et ainsi de calculer la probabilité qu'un pixel donné soit un pixel de peau.

Cette méthode suppose que la distribution de couleur de peau puisse être modélisée par une gaussienne, ce qui n'est pas forcément évident. [8]

3.2. Méthodes non paramétriques :

Les méthodes non paramétriques estiment la couleur de peau par l'histogramme de la base d'apprentissage sans la création préalable d'un modèle de peau, la distribution de la couleur de peau est modélisée par un histogramme. Ceci a l'avantage de ne pas faire de supposition sur le type de distribution. On peut ensuite calculer la probabilité qu'un pixel donné soit un pixel de peau. Le temps de calcul nécessaire est plus faible que pour les

Chapitre II : Système de reconnaissance de gestes basé vision

méthodes paramétriques, car il suffit de prendre la valeur dans l'histogramme, mais l'espace de stockage nécessaire est plus important, puisqu'il faut stocker l'histogramme en mémoire au lieu de quelques paramètres pour les gaussiennes [8].

Depuis [8] Phung et al. ont réalisé une comparaison des différentes méthodes, et ils montrent que la méthode avec les histogrammes et la classification bayésienne est parmi les plus performantes. Ils montrent également que les performances sont moins bonnes lorsque la chrominance seule est utilisée pour la classification. La prise en compte de la composante de luminance dans l'histogramme permet d'être plus robuste aux variations de luminosité. Les méthodes basées sur un mélange de gaussiennes sont plus gourmandes en temps de calcul, alors que leurs performances ne sont pas meilleures.

La méthode la plus simple est « Explicit Skin Cluster » qui consiste essentiellement à définir les pixels peau par seuillage donnant les coordonnées de l'intervalle de la couleur peau dans un espace de couleur approprié. Cette méthode est très populaire grâce à la simplicité de sa mise en œuvre et aussi parce qu'elle ne nécessite aucun apprentissage préalable [18].

4. Les espaces colorimétriques :

Nous présentons plusieurs espaces colorimétriques pour la détection de la couleur de peau par seuillages: RGB, HSV, YCbCr et HSCrCb, le choix de ces espaces de couleur parmi tant d'autres est justifié par leur popularité dans la littérature scientifique d'une part et par le fait que nous n'avons trouvé les seuils de la couleur de peau que pour ces espaces d'autre part.

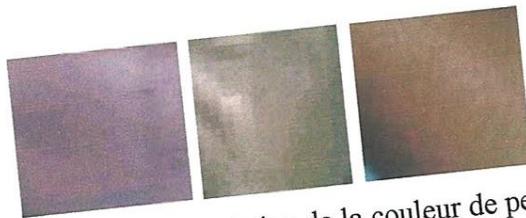


Figure II. 3: Différentes présentations de la couleur de peau [18]

4.1 Espace de couleur HSV :

C'est un modèle de représentation dit "naturel", c'est-à-dire proche de la perception physiologique de la couleur par l'œil humain. Il consiste à décomposer la couleur selon des critères physiologiques (Figure II.4) :

- **la teinte** (en anglais *Hue*), correspondant à la perception de la couleur,
- **la saturation**, décrivant la pureté de la couleur, c'est-à-dire son caractère vif ou terné,
- **la luminance** (en anglais *value*), indiquant la quantité de lumière de la couleur, c'est-à-dire son aspect clair ou sombre. Dans l'espace HSV, l'information d'intensité est représentée par

Chapitre II : Système de reconnaissance de gestes basé vision

le canal V, pour cette raison, ce canal doit être négligé dans le processus de détection de la peau, on ne considère que les canaux H et S qui représentent l'information chromatique.

Avec : $H \in [0 - 360^\circ]$ et $S \in [0 - 1]$ et $V \in [0 - 1]$ (II.4).

La segmentation de la couleur de peau selon l'espace HSV nécessite une conversion de l'image originale de l'espace RGB (souvent les images acquises se présentent dans cet espace) à l'espace HSV dans lequel la couleur de peau se présente dans les intervalles H et S suivants : $0 < H < 50^\circ$ et $0.23 < S < 0.68$. [18] (II.5).

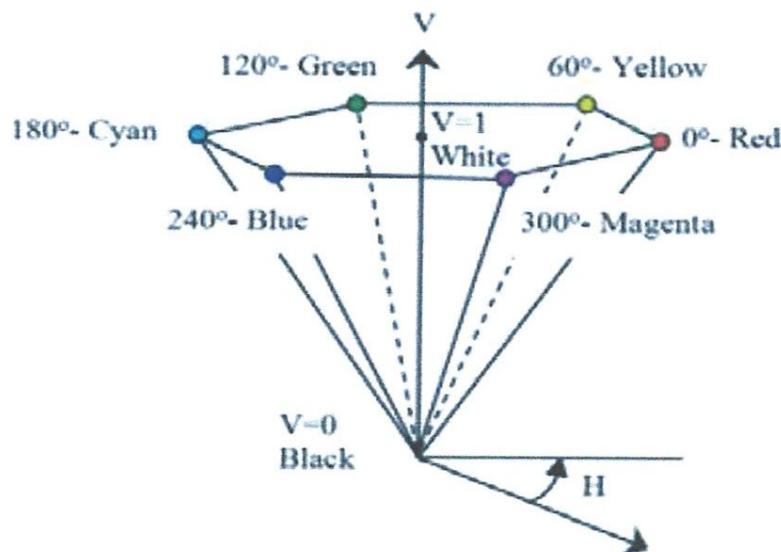


Figure II. 4: Présentation des couleurs dans l'espace HSV [19]

4.2 Espace de couleur RGB :

Cet espace de couleur correspond à la façon dont les couleurs sont généralement codées informatiquement, ou plus exactement à la manière dont les tubes cathodiques des écrans d'ordinateurs représentent les couleurs.

Etant donné que le codage RGB repose sur trois composantes proposant la même gamme de valeurs, on le représente généralement graphiquement par un cube dont chacun des axes correspond à une couleur primaire (Figure II.5).

Avec : $R, G, B \in [0 - 255]$.

Le seuillage appliqué sur l'espace RGB, s'applique aux trois canaux (Red, Green & Blue), plus une troisième contrainte, celle de l'éclairage de l'image [18].

➤ Cas éclairage fort :

Chapitre II : Système de reconnaissance de gestes basé vision

$$\left. \begin{aligned} (R > 95) \text{ et } (G > 40) \text{ et } (B > 20) \text{ et } ((\max(R, G, B)) - \min([R, G, B])) > 15) \\ \text{Et } (ABS(R - G) > 15) \text{ et } (R > G) \text{ et } (R > B). \end{aligned} \right\} \quad (\text{II.6}).$$

➤ Cas éclairage faible :

$$(R > b) \text{ et } (G > B) \text{ ou } ((R > 220) (G > 210) \text{ et } (B > 170) \text{ et } (ABS(R - G) \leq 15)). \quad [\text{18}] \quad (\text{II.7}).$$

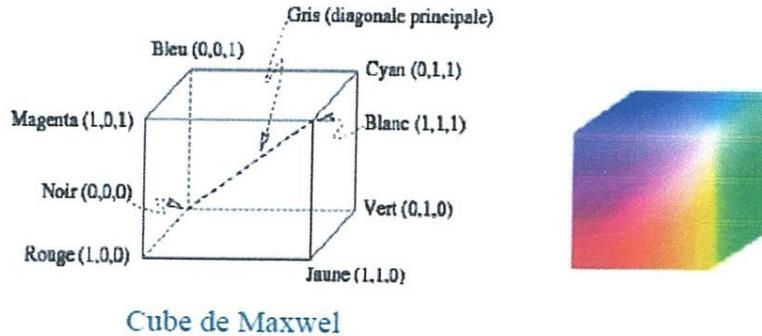


Figure II. 5: Présentation des couleurs dans l'espace RGB [20].

4.3 Espace de couleur YCbCr :

Utilisé souvent dans la compression des images, la luminance (appelée aussi Luma) est séparée de la couleur, présentée par la valeur Y et calculée par une somme pondérée des composantes de R, G et B :

$$Y = 0,299 * R + 0,587 * G + 0,114 * B \quad (\text{II.8})$$

Les deux autres composantes de cet espace représentent l'information chromatique et elles sont calculées :

$$Cr = R - Y.$$

$$Cb = B - Y.$$

Comme pour l'espace HSV, seul les deux canaux Cb et Cr sont retenus à l'application du seuillage.

Deux seuillages ont été retenus :

➤ Seuillage YCbCr(1) :

$$((Cb \geq 85) \text{ et } (Cb \leq 135)) \text{ et } ((Cr \geq 135) \text{ et } (Cr \leq 180)). \quad (\text{II.9}).$$

➤ Seuillage YCbCr(2) :

$$((Cb \geq 77) \text{ et } (Cb \leq 127)) \text{ et } ((Cr \geq 133) \text{ et } (Cr \leq 173)). \quad [\text{18}]. \quad (\text{II.10}).$$

Chapitre II : Système de reconnaissance de gestes basé vision

4.4. HSCbCr :

Le HSCbCr c'est un espace de couleurs obtenu par l'hybridation des deux espace HSV et YCbCr, le H et le S sont pris de l'espace HSV et le Cb et Cr sont pris de l'espace YCbCr, selon [21] les performances de HSCbCb sont bonne que celle du YCrCb. si l'image contient beaucoup du rouge alors le résultat de la peau sera faible pour YCbCr, tandis que la même image avec l'espace HSCbCr donne des résultats exacte.

Dans la Figure II. 6, (a) représente les images originales, (b) représente les images HSCbCr, et finalement(c) représente les images YCbCr.

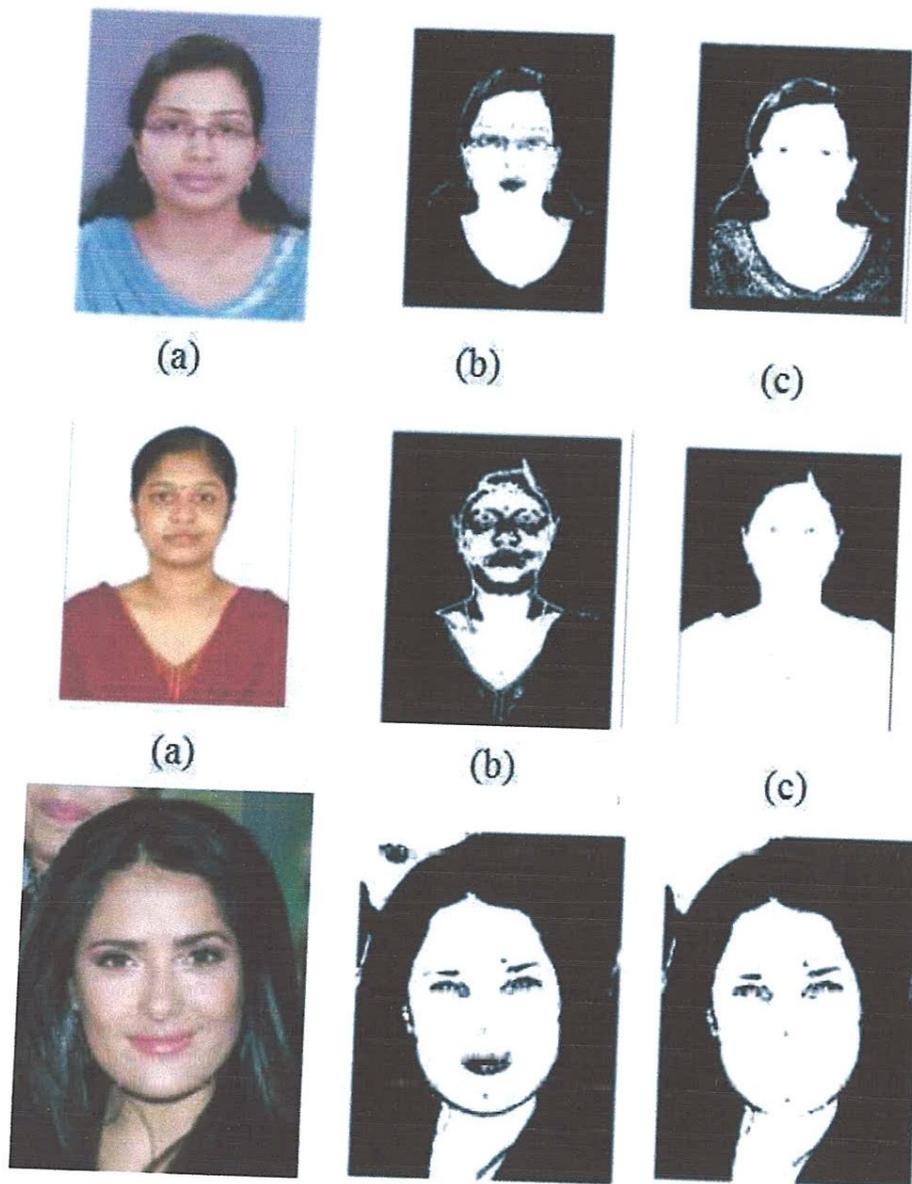


Figure II. 6: Les résultats du model HSCbCr. [21]

Chapitre II : Système de reconnaissance de gestes basé vision

4.5. Les travaux existantes :

Les espaces de couleurs	Les travaux
RGB	[10], [18], [23], [32], [34]
Normalized-RGB	[23]
HSV-YUV	[18], [15], [30], [31], [34]
YCbCr	[18], [8], [31], [32], [14], [33], [21], [34]
HSCbCr	[21]

Tableau II. 1: les espaces couleurs utilisent pour la détection de peaux.

Travail	Image d'entrée	Détection de peaux	Méthode utilise	Espace de couleur utilise
[34]			Non spécifique	RGB+ YCbCr + HSV
[10]			carte de probabilité	YCbCr
[14]			Loi gaussienne	YCbCr
[32]			Non spécifique	RGB
[19]			Non spécifique	HSV
[18]			Non spécifique	RGB-HSV
[37]			Classification des régions de peaux	HSV

Tableau II.2: Les résultats obtenus avec des différents espaces de couleur.

Chapitre II : Système de reconnaissance de gestes basé vision

5. Post des traitements :

Un prétraitement de l'image est nécessaire pour éviter le bruit causé par les conditions d'acquisition de l'image et de sa numérisation.

Les méthodes présentées dans ce chapitre agissent exclusivement au niveau pixel. Par conséquent, il n'est pas garanti que les pixels détectés comme appartenant à un objet soient connectés entre eux.

De plus, il y a généralement des fausses détections, qui se manifestent par des pixels blancs isolés, ainsi que des pixels de l'objet non détectés, qui se manifestent par des « trous » dans le masque. Il est donc nécessaire de filtrer l'image binaire pour ne garder que la région correspondant à la main [8].

Pour supprimer les pixels isolés en applique un filtre robuste pour le bruit, le filtre médian de convolution (3* 3) ou (5* 5), L'avantage principal de ce filtre est qu'il n'introduit pas l'effet flou sur les contours tel que les autres filtres, puis en appliquent des morphologies mathématiques.

5.1 Morphologie mathématique :

Les opérateurs morpo mathématiques se sont initialement appliqués sur des images en noir et blanc (Matheron et Serra, 1965). Ils ont ensuite été étendus à des images en niveaux de gris par Dougherty en 1978.

Pour les applique à des images couleurs, il suffit alors de les applique séparément à chaque composent couleur [22].

Les opérateurs morpo mathématiques sont définis à partir des petits ensembles (B) discrets appelle « éléments structurants ».

Les éléments structurants permettent de définir le type de voisinage que l'on souhaite considérer.

$$K_c = \begin{bmatrix} 010 \\ 111 \\ 010 \end{bmatrix} \quad K_h = \begin{bmatrix} 000 \\ 111 \\ 000 \end{bmatrix} \quad K_s = \begin{bmatrix} 111 \\ 111 \\ 111 \end{bmatrix} \quad K_v = \begin{bmatrix} 010 \\ 010 \\ 010 \end{bmatrix} \quad (\text{II.11}).$$

5.1.1. Dilatation :

C'est l'ensemble des points tels que lorsque (B) est centré sur un de ces points il y a une intersection non vide entre A et B.

On obtient l'image dilatée de la forme (A) en remplaçant tout pixel par le maximum des intensités des voisins selon la forme (B).

$$A \oplus B = \bigcup (A+b : b \in B) = \bigcup (B+a : a \in A) = \{x : (-B + x) \cap A \neq \emptyset\} \quad (\text{II.12}).$$

Chapitre II : Système de reconnaissance de gestes basé vision

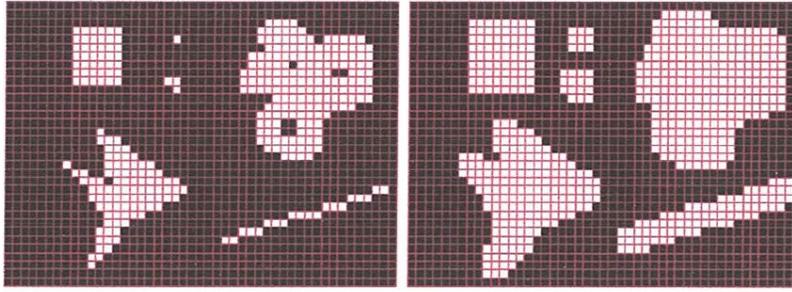


Figure II. 7: Exemple de dilatation. A gauche : l'image originale. A droite : sa dilatée [22].

- Tous les objets vont grossir d'une partie correspondant à la taille de l'élément structurant.
- Si des objets sont situés à une distance moins grande que la taille de l'élément structurant, ils vont fusionner.
- L'existence d'un trou dans l'objet, il sera comblés.

5.1.2. Erosion :

L'érodé de (A) par (B) correspond à l'ensemble des points tels que si B est centré sur ces points, B est entièrement inclus dans A.

$$C=A\ominus B=\bigcap (A-b:b\in B)=\{x:(B+x)\subset A\} \quad (\text{II.13}).$$

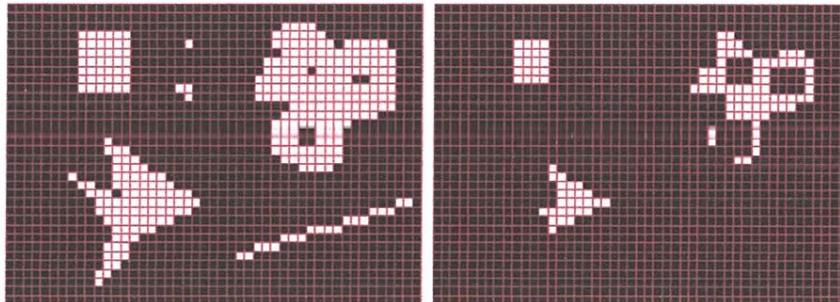


Figure II. 8: Exemple d'érosion. A gauche : l'image originale. A droite : son érodée. [22]

On remarque généralement qu'après avoir effectué ces deux opérations morphologiques

- Les objets de taille inférieure à celle de l'élément structurant vont disparaître.
- Des objets reliés entre eux vont être séparé.
- L'existence d'un trou dans l'objet, il sera accentué. [22]

5.1.3. Ouverture :

L'ouverture est une opération d'érosion suivie d'une opération de dilatation.

- Disparition des structures et détails fin.
- Elle permet d'éliminer les morceaux de contours bruités. [22]

Chapitre II : Système de reconnaissance de gestes basé vision

5.1.4. Fermeture :

La fermeture est une opération dilatation suivi d'une opération de d'érosion.

- Fusionnement des structures les plus proches.
- Elimination des trous [22].

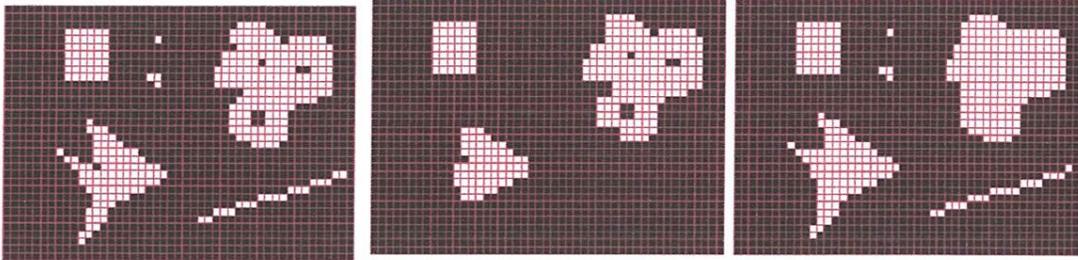


Figure II. 9: A gauche : image original A milieu : ouverture. A droite : fermeture [22].

5.2. Les résultats des travaux existants (post de traitement) :

Travail	Image d'entrée	Image résultat	Méthode utilise
[14]			Holes filling et Filtre médian
[8]			Filtre médian et ouverture
[36]			Image après érosion
[35]			Filtration
[38]			Morphologique opération et filtration

Tableau II. 3: les résultats de près traitement des travaux voisins

Chapitre II : Système de reconnaissance de gestes basé vision

6. Extraction des caractéristiques :

Dans cette étape, les caractéristiques doivent être extraites pour spécifier chaque geste, plusieurs approches sont utilisées dans cette étape. Cette section va parler sur les plus célèbre approches d'extraction de caractéristiques, parmi ces caractéristiques le center de la main, la pomme de la main et les bouts des doigts.

- **Le centre de la main :**

Est un point important car il est souvent le premier à être localisé et est très utilisé pour le suivi de la main. Il permet de représenter la position globale de la main, ce qui est parfois suffisant pour reconnaître certains gestes, tels que des gestes de la langue des signes.

Il est aussi très utile de connaître la position du centre pour détecter les bouts des doigts, avec par exemple la distance au centre.

Un des problèmes pour la détection du centre est de savoir si l'utilisateur a le bras nu ou non. En effet, si celui-ci est nu, la détection de la couleur de peau fournit une région contenant la main et l'avant-bras. Cela pose aussi des problèmes pour la détection des doigts et le calcul des caractéristiques de forme [8].

Ce qui nous amène à la nécessité de détecter le poignet pour séparer la main de l'avant-bras, afin d'éviter d'imposer des contraintes aux utilisateurs telles que le port d'un vêtement à manches longues ou des gants colorés.

- **Les bouts des doigts :**

Sont des points très importants pour le suivi et la reconnaissance de gestes simples en comptant le nombre de doigts, les bouts des doigts sont les points du contour qui se trouvent aux extrémités de la région de la main. Ainsi, il est possible de détecter ces points en utilisant des connaissances a priori sur la morphologie de la main, par exemple en calculant la distance par rapport au centre de la main [8].

6.1. Carte de distance :

Cette méthode se base sur l'idée que le centre de la main est le point situé à la plus grande distance des bords de la main. La transformée en distance d'une image binaire associée à chaque pixel de l'objet la distance au pixel du contour le plus proche.

Une carte de distance est associée à une métrique (distances euclidienne, de Manhattan), cette métrique modifie le poids donné aux pixels verticaux, horizontaux, et diagonaux.

➤ Le centre de la main est déterminé comme étant le maximum de la carte de distance.

Chapitre II : Système de reconnaissance de gestes basé vision

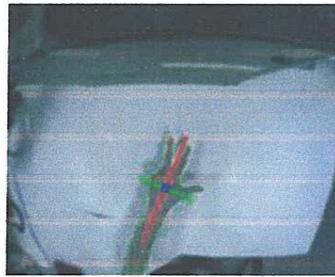


Figure II. 10 : Détection du centre avec la carte de distance [8].

6.2. L'enveloppe Convexe :

L'enveloppe convexe d'un objet ou d'un regroupement d'objets géométriques est l'ensemble convexe le plus petit parmi ceux qui le contiennent. Par exemple dans la (Figure II.11) il existe 10 points.

L'hexagone rouge dans la figure est l'enveloppe convexe de ce groupe de point. Les six point qui compose le hexagone sont appelés « les points de l'enveloppe ».

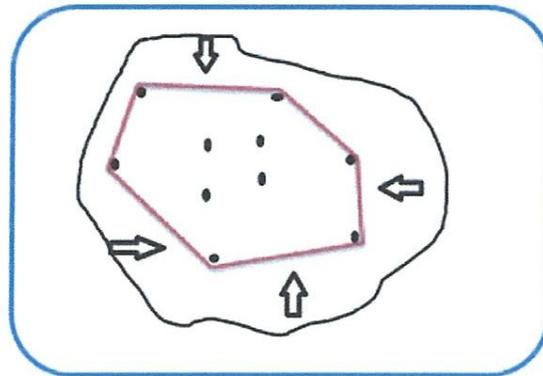


Figure II. 11 :L'enveloppe convexe.

La partie de la main contient des contours convexes et concaves et ces contours portent généralement les informations que nous voulons pour extraire les informations voulues dans la main.

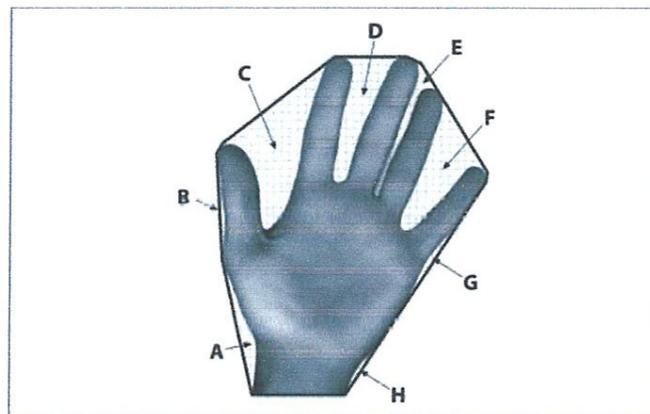


Figure II.12:L'enveloppe convexe de la main [29].

Chapitre II : Système de reconnaissance de gestes basé vision

Il existe plusieurs algorithmes pour le calcul de l'enveloppe convexe parmi ces algorithmes on cite : l'algorithme de Brute Force, GiftWrapping, Quickhull, Divide et Conquer ,etc. On va choisir un de ces algorithmes et donner son algorithme de fonctionnement.

6.2.1. Quickhull :

Quickhull est un algorithme pour le calcul de l'enveloppe convexe, cet algorithme consiste aux étapes suivantes:

- Choisir deux points (p_1, p_2), le premier contient la valeur maximale de la coordonnée x et le deuxième aura la valeur minimale de x , donc ces deux points vont faire partie de l'enveloppe convexe.
- On trace avec ces deux points une ligne qui coupe l'ensemble des points en deux régions, (voir Figure II.13).

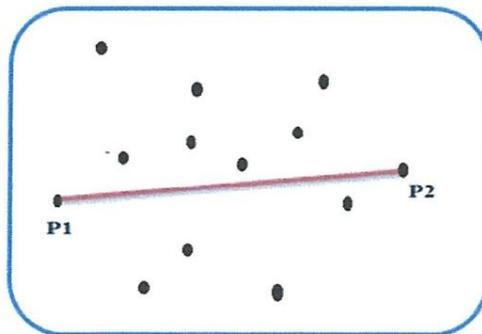


Figure II.13 : P_1, P_2 deux points avec X_{min} et X_{max} .

- Choisir dans une seule région le point (p_3) le plus loin de la ligne et former un triangle avec ce dernier, (voir Figure II.14).

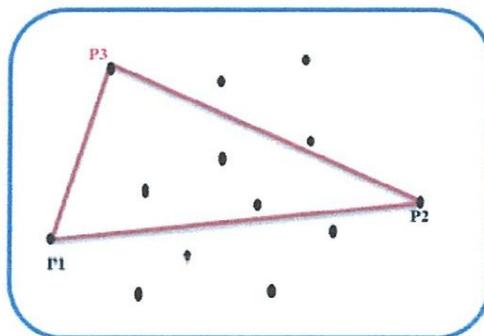


Figure II.14: p_3 plus loin point.

- Les points situés à l'intérieur de ce triangle ne peuvent pas faire partie de l'enveloppe convexe et peuvent être donc ignorés dans les prochaines étapes.

Chapitre II : Système de reconnaissance de gestes basé vision

- Répétez les deux étapes précédentes jusqu'à ce qu'il ne reste plus de points à l'intérieur du triangle, les points sélectionnés vont former l'enveloppe convexe, (voir Figure II.15).

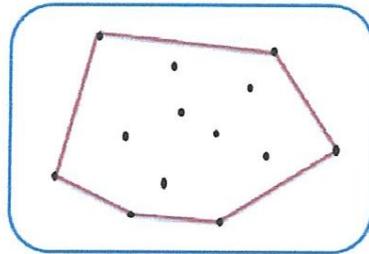


Figure II. 15: Obtention de l'enveloppe convexe.

6.3. Les travaux voisins pour les extractions des caractéristiques :

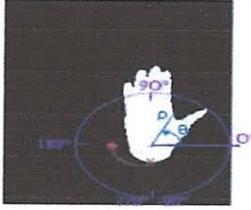
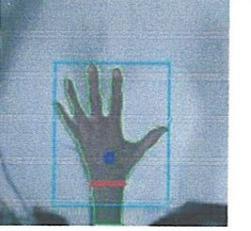
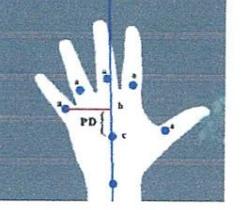
Travaux	Méthode utilisée	Type d'extraction	Exemple
[30]	l'enveloppe convexe	Enveloppe la main pour extraire les doigts	
[14]	La distance et l'angle appliqué sur le contour de la main	Le centre de la main	
[31]	Algorithme pour la détection des doigts appliqué sur contour de la main	Extraction des doigts	
[8]	la carte de distance et réduire la taille de la boîte englobant	Le centre de la main et détection du poignet pour séparer la main de l'avant-bras	
[23]	Caractéristiques géométriques	Le centre de la main et les doigts	

Tableau II. 4: les résultats obtenus pour l'extraction des caractéristiques

Chapitre II : Système de reconnaissance de gestes basé vision

7. Classification :

La classification est la dernière étape dans le processus de reconnaissance, l'essentiel de cette étape et d'identifier à quel classe appartient chaque geste on faisant une comparaison entre les caractéristiques extraites du geste actuelle et de celui enregistré dans la base de données, il existe deux types de classification :

- ✓ Classification basé Apprentissage.
- ✓ Classification basé Règle.

7.1. Classification basé Apprentissage :

Il existe 2 principaux types d'apprentissage:

- ❖ Apprentissage supervisé: les différentes classes des formes sont connus a priori et on dispose d'exemples déjà étiquetés.
- ❖ Apprentissage non supervisé: aussi appelé classification automatique. Dans ce cas les classes d'appartenance des formes est inconnues a priori. L'apprentissage non supervisé consiste à construire automatiquement ces classes.

7.1.1. Machines à Vecteurs Supports(SVM) :

Le classificateur SVM (en anglais *Support Vector Machine*, SVM) a été conçu pour une séparation de deux ensembles de données. Il est considéré donc comme un classificateur binaire.

C'est à dire qu'ils permettent de créer une surface de décision entre deux classes définies dans un même espace. Pour cela, ils construisent une frontière de décision par projection des caractéristiques provenant d'un espace d'origine dans un espace de caractéristiques de dimension supérieure (voir infini) dans le but de rendre les classes linéairement séparables.

Il existe l'hyperplan choisi est celui qui maximise la marge de séparabilité entre les deux ensembles de données. La sélection de l'hyperplan dans un espace de caractéristiques nécessite d'évaluer un produit scalaire dans cet espace. Ce qui peut être très coûteux en temps et en complexité si l'espace est de très grande dimension. Heureusement, ce calcul n'est pas obligatoire grâce à une opération mathématique appelé kernel.

Le kernel calcule le produit scalaire de deux points dans l'espace de dimension supérieur sans avoir à les projeter.

Chapitre II : Système de reconnaissance de gestes basé vision

- ❖ la validation des données.
- ❖ la gestion des risques.

Avec c'est application les ANN utilisent donne notre cas le domaine de la reconnaissance des gestes pour classaient les gestes de la main.

Les avantages :

- Fonctionnement en temps réel: calculs ANN peut être effectués en parallèle, et les périphériques matériels spéciaux sont conçus et fabriqués qui profitent de cette capacité.
- Adaptive apprentissage: Une capacité d'apprendre à effectuer des tâches sur la base des données fournies pour la formation ou l'expérience initiale.

Les inconvénients :

- ❖ Les réseaux de neurones produisent des modèles souvent incompréhensibles et demande un longtemps d'apprentissage.
- ❖ Le réseau de neurones a besoin de formation pour fonctionner. L'architecture d'un réseau de neurones est différente de l'architecture des microprocesseurs qui a donc besoin d'être imité. Nécessite du temps de traitement élevé pour les grands réseaux de neurones.

7.1.3. Méthode de k plus proche voisins :

La méthode de k plus proches voisins (K plus porches voisin K-PPV, ou K Nerest Neighbors en anglais KNN) est un algorithme de raisonnement à partir de cas qui est dédié à la classification qui peut être étendu à des tâches d'estimation. Le but de cet algorithme est de prendre des décisions en se basant sur un ou plusieurs cas similaires déjà résolus en mémoire. Dans ce cadre, et Contrairement aux autres méthodes de classification (réseaux de neurones, SVM...etc.) l'algorithme de KNN ne construit pas de modèle à partir d'un échantillon d'apprentissage, mais c'est l'échantillon d'apprentissage, la fonction de distance et la fonction de choix de la classe qui fonctionnent comme des classes des voisins les plus proches, qui constituent le modèle.[27]

On utilise plusieurs distance pour trouver le plus proche voisin exemple ;

$$\text{La distance Euclidienne : } \Delta (P(i,j),P(k,l)) = \sqrt{(k-i)^2 + (l-j)^2}. \quad (\text{II.13}).$$

$$\text{La distance par Blocs (Manhattan) : } \Delta (P(i,j),P(k,l)) = (|k-i|+|l-j|) \quad (\text{II.14}).$$

$$\text{La distance tour d'échiquier : } \Delta (P(i,j),P(k,l)) = \max(|k-i|+|l-j|) \quad (\text{II.15}).$$

$$\text{La distance de Minkowski : } d(Xi, yi) = \left(\sum_{i=1}^k |Xi - yi|^q \right)^{1/q}. \quad [22] \quad (\text{II.16}).$$

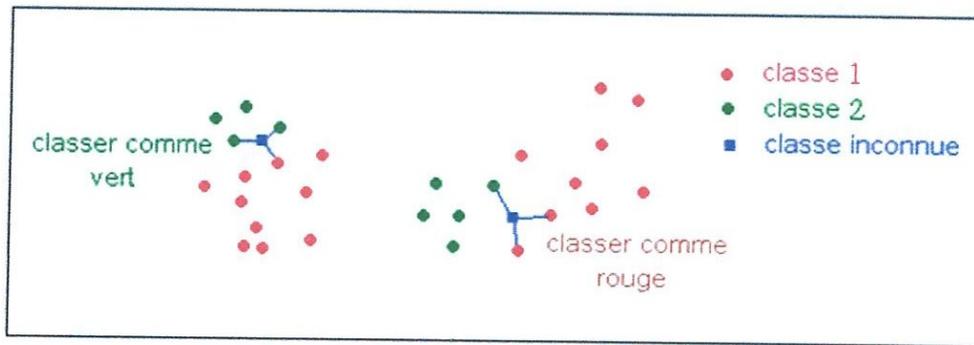


Figure II.18 : Exemple de classification avec KNN (K=3) [25].

Les avantages :

- ❖ La qualité de la méthode s'améliore en introduisant de nouvelles données sans nécessiter la reconstruction d'un modèle. Ce qui représente une différence majeure avec des méthodes telles que les réseaux de neurones.
- ❖ Propriétés théoriques de généralisation.

Les inconvénients :

- ❖ Temps de classification : la méthode ne nécessite pas d'apprentissage ce qui implique que tous les calculs sont effectués lors de la classification. Contrairement aux autres méthodes qui nécessitent un apprentissage (éventuellement long) mais qui sont rapides en classification.
- ❖ Méthode donnera de mauvais résultats Si le nombre d'attributs pertinents est faible relativement au nombre total d'attributs, car la proximité sur les attributs pertinents sera noyée par les distances sur les attributs non pertinents.

7.1.4. Model de Markov cache :

Le classificateur de Markov caché (MMC) appartient à la classe des classificateurs d'entraînement, il représente un modèle statistique, dans lequel la classe de geste de correspondance la plus probable est déterminé par un vecteur de caractéristiques donné, sur la base des données d'apprentissage.

Cette méthode a été largement exploitée pour la reconnaissance des gestes temporelle. Un MMC est constitué d'un ensemble d'états et des transitions d'état avec des probabilités d'observation. Pour regarder un geste, un distinct MMC est formé et la reconnaissance du geste est basée sur la génération de probabilité maximale par un MMC particulier. Cette méthode souffre également de temps de formation impliqués et la nature de travail complexe que les résultats sont imprévisibles à cause de la nature cachée [28].

Chapitre II : Système de reconnaissance de gestes basé vision

7.2. Classification basé règle :

Dans les approches à base de règles, la règle est créée sur la base des vecteurs de caractéristiques, et ceux qui correspondent avec la règle peuvent être considérés comme le résultat final, cette méthode a des avantages comme elle a ces inconvénients aussi.

Les avantages :

- ❖ L'avantage majeur est la vitesse d'exécution en temps réel. Au contraire des méthodes basé sur l'apprentissage qui prennent beaucoup de temps pour la reconnaissance, cette méthode donne des résultats avec une grande rapidité.

Les inconvénients :

L'inconvénient de cette approche est qu'elle dépend de la capacité de l'être humain à générer plusieurs règles qui assurent l'obtention d'une reconnaissance parfaite [28].

7.3. Les travaux existantes pour la classification :

Ce tableau représente le nombre de gestes testés et reconnus avec ses méthodes et les bases utilisées pour la classification :

Travail	Nombre des gestes reconnues	Nombre des gestes testés	Pourcentage de reconnaissance	Méthode de classification	Base de données utilisée
[23]	20	200	100% pour 14 gestes et 90% pour 15-20 Gestes	Distance euclidienne	Base privée
[39]	31	130	90.45%	distribution Gaussien	Base privée
[40]	6	60	95%	Distance euclidienne	Base privée
[41]	26	1040	98.7	Les réseaux neurones	Langage des signes américains (ASL)

Tableau II. 5: les résultats obtenus par les travaux voisins

8. Conclusion :

Dans ce chapitre, nous avons présenté les étapes du processus de reconnaissance d'un geste dans un système basé vision. On a vu quelques techniques utilisées pour extraire les

Chapitre II : Système de reconnaissance de gestes basé vision

caractéristiques contenus dans une image, ainsi que ce n'est pas facile de trouver une méthode pour la segmentation qui soit performante, forte aux variations de luminosité ou à un fond complexe, et qui respecte la contrainte du temps réel. Mais chaque approche a des avantages comme elle a des inconvénients.

Les méthodes basées sur la couleur de peau sont beaucoup plus adaptées à notre problématique. Aussi nous avons présenté une méthode basée sur un seuillage des composantes de chrominance, Cb et Cr.

Elle permet une détection très rapide concernant les mains, mais la robuste méthode repose sur l'utilisation d'un gant coloré qui facilite la détection de la main sans contrainte de variations de luminosité ou à un fond complexe ainsi que l'utilisation d'un capteur infrarouge (kinect), qui a la capacité de capter même en nuit, puis l'utilisation des méthodes pour extraire des caractéristiques de la main dans le but d'obtenir une meilleure classification pour la reconnaissance des gestes de la main.

Toutes ses méthodes nous intéressent pour exploiter une conception d'un système de reconnaissance des gestes de la main en temps réel, dans le chapitre suivant on présente une conception détaillée pour notre système.

Chapitre III : Conception et Implémentation

Chapitre III : Conception et implémentation

1. Introduction :

Ce chapitre se concerne à la conception et le développement d'un système de vision par ordinateur pour la reconnaissance de gestes de la main. Nous cherchons à répondre aux besoins de conception d'une Interface Homme-Machine dont l'objectif est de transformer un ordinateur usuel en surface gestuelle augmentée et de permettre à l'utilisateur de se servir de son doigt comme dispositif de pointage et des gestes pour le contrôle.

Les gestes de la main sont un canal de communication naturel et intuitif chez l'homme pour interagir avec son environnement. Ils servent à désigner ou à manipuler des objets, à renforcer la parole, ou à communiquer basiquement dans un environnement bruité. Ils peuvent aussi représenter un langage à part entière avec la langue des signes.

Ainsi, dans le domaine des Interfaces Homme-Machine (IHM), la main peut servir à pointer (pour remplacer la souris), à manipuler des objets (pour la réalité augmentée ou virtuelle), ou à communiquer par gestes avec un ordinateur. Par rapport à la richesse de l'information véhiculée par les gestes de la main, les possibilités de communication avec les ordinateurs sont aujourd'hui réduites avec la souris et le clavier. L'interaction homme-machine est basée actuellement sur le paradigme WIMP (Window, Icon, Menu, Pointing device) qui présente les bases fonctionnelles d'une interface graphique informatique (GUI «Graphical User Interface»). La majorité des systèmes d'exploitation repose sur ce concept, avec un dispositif de pointage, généralement la souris, qui permet d'interagir avec des éléments graphiques tels que des fenêtres, des icônes et des menus, de façon plus intuitive qu'avec une interface textuelle (en ligne de commande). En utilisant les gestes de la main, l'interface devient perceptuelle (PUI « Perceptual User Interface»).

Les systèmes de reconnaissance de gestes ont d'abord utilisé des gants électroniques munis de capteurs fournissant la position de la main et les angles des articulations des doigts. Mais ces gants sont onéreux et encombrants, d'où l'intérêt croissant pour les méthodes de vision par ordinateur. En effet, avec les progrès techniques et l'apparition de caméras bon marché, il est désormais possible de développer des systèmes de reconnaissance de gestes basés sur la vision par ordinateur, fonctionnant en temps réel.

Toutefois, la main étant un organe complexe, déformable, comportant de nombreux degrés de liberté au niveau des articulations, il est difficile de reconnaître sa forme à partir d'images sans un certain nombre de contraintes, d'où. Pour les systèmes de reconnaissance de geste on peut citer plusieurs difficultés :

- ✓ La complexité de l'environnement,

Chapitre III : Conception et implémentation

- ✓ La variance de luminosité,
- ✓ La complexité de l'arrière-plan (conflits),
- ✓ Présence du visage dans la scène ou non ? (champ de vision de la caméra)
- ✓ Utilisation de deux mains ou une seule
- ✓ Le bras nu ou non,
- ✓ La complexité des gestes si dynamiques ou statiques.

Dans ce qui suit, le chapitre est décomposé en parties, partie I : conception qui illustre les méthodes que nous utilisons pour réaliser notre système ; partie II : implémentation des concepts développés dans la partie I.

Partie I : Conception :

Pour réaliser notre système nous suivrons 3 étapes : (segmentation, extraction des caractéristiques et classification) ses dernier sont les principaux étapes pour développer la plupart des systèmes de détection et reconnaissance de geste [8], [14], [30], [31], [37]. Ses étapes sont bien détaillé dans le schéma III.1 qui représente l'architecture générale de notre système.

I.1. Architecture générale de notre système :

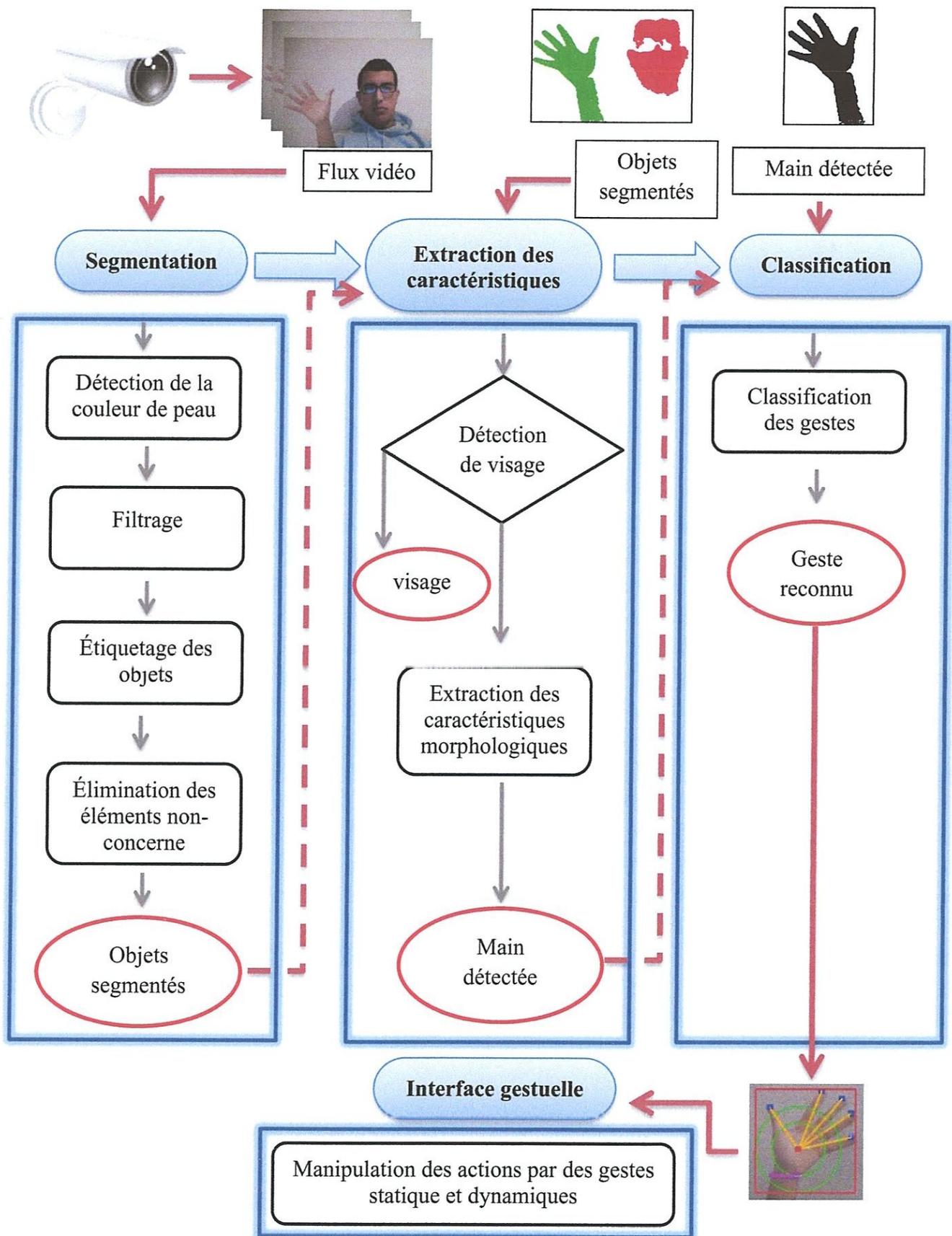


Schéma III.1 : Architecture générale de notre système.

Chapitre III : Conception et implémentation

En préambule : Nous utilisons un flux vidéo acquis directement par une caméra puis nous traitons ce vidéo, image par image, pour chaque image on extraie les régions de la peau ensuite on applique un post de traitement pour améliorer le résultat de détection, après on segmente les objets de l'image et élimine tous les éléments subtiles. Pour chaque objet étiqueté l'objectif est de vérifier si c'est une main ou un visage, on commence par éliminer les objets visages, puis il ne reste que les objets supposés main, alors on l'extraie leurs caractéristiques morphologiques pour connaitre est-ce-que il s'agit réellement d'une main ou non ? Si cet objet est une main on passe à l'étape de classification pour reconnaître le geste.

I.2. Segmentation :

L'étape de segmentation est une étape très importante pour la réalisation d'un système de reconnaissance de geste, elle influence les autres étapes qui se suivent, pour obtenir un bon résultat de reconnaissance il faut faire un algorithme de segmentation robuste, par ailleurs, il faut prendre en compte la contrainte du temps réel, car la segmentation n'est que la première étape d'un système de reconnaissance de gestes.

Il existe de nombreuses méthodes (chapitre II, types de segmentation), plus ou moins performantes suivant les suppositions faites sur l'environnement (simple, complexe, condition d'illumination...), une approche basée sur la soustraction d'image de fond utilisée par [31] pour résoudre le problème de complexité de l'environnement, mais avec cette approche on ne peut pas changer la scène ou la position de caméra (environnement statique), d'autres travaux utilisant un gant coloré comme [14] pour obtenir un bon résultat, mais cette solution oblige l'utilisateur à porter toujours ces gants, et ça rend le système inflexible, une autre solution est d'acquérir les images à l'aide d'un capteur infrarouge (kinect) cette solution est performante mais elle reste toujours coûteuse.

La segmentation par détection de la couleur de peau c'est la méthode la plus optimale et qui a été appliquée par plusieurs travaux voir : (Tableau II. 1, Tableau II. 2), dans notre travail nous appliquons cette méthode parce que le temps de calcul est faible ainsi que les résultats sont acceptables pour la démarche de notre système en temps réel, mais elle souffre du problème de variation de luminosité et des objets de la scène qui ont une couleur proche de la couleur de la peau, pour bien détailler (chapitre II, types de segmentation).

I.2.1. Détection de la couleur de peau :

L'analyse de la couleur de la peau est très utilisée pour la détection du visage et des mains. La couleur de la peau présente une distribution caractéristique dans certains espaces

Chapitre III : Conception et implémentation

colorimétriques, cette propriété peut être utilisée pour segmenter les régions de couleur de peau [8]. Il existe plusieurs espaces colorimétriques utilisés pour détecter les régions de la peau : RGB, HSV, YCbCr, ces espaces ne donnent pas des résultats parfaits alors les recherches sont courantes dans ce domaine telle que [21] qui fait une hybridation entre les deux espaces HSV et YCbCr et ça a donné comme résultat un nouvel espace colorimétrique HSCbCr.

Le choix de l'espace est très important pour cela nous comparons les performances de ces derniers donc nous trouvons des différents résultats, des études comparatives comme [18] ont été élaborées et chacun trouve qu'un espace est mieux que l'autre par rapport à l'environnement de travail, de notre part on a trouvé que l'espace YCbCr donne un meilleur résultat voici le tableau au-dessous.

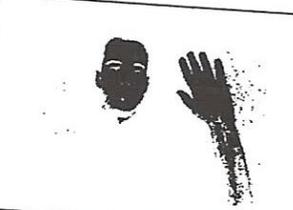
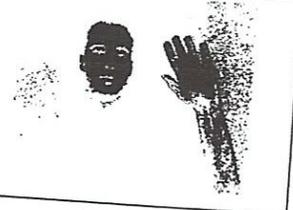
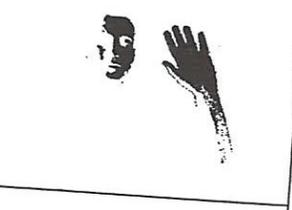
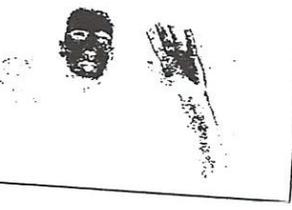
Image original	Espace colorimétriques	Image résultat
	YCbCr(1)	
	YCbCr(2)	
	HSV	
	RGB	
	HSCbCr	

Tableau III. 1 : comparaison entre les espaces colorimétriques.

Chapitre III : Conception et implémentation

Il existe différentes méthodes pour classer les pixels de couleur de la peau. Deux grands types de modèles pour la classification : paramétrique et non-paramétrique. Le temps de calcul nécessaire pour le modèle non-paramétrique est plus faible que le temps du modèle paramétrique [8], nous testons une méthode non-paramétrique qui consiste à modéliser la couleur de peau par un histogramme dans l'espace YCbCr, et à utiliser une règle de décision bayésienne afin de classer les pixels peau et les pixels non-peau ça donne un résultat très similaire avec l'espace YCbCr simple, parce que cette méthode repose sur l'histogramme de CbCr si ce dernier n'est pas bon la méthode bayésienne n'améliore pas le résultat, par conséquent nous améliorons notre résultat par d'autre traitement.

I.2.2. Filtrage :

Les méthodes présentées dans la section précédente exclusivement au niveau pixel. Par conséquent, il n'est pas garanti que les pixels détectés comme appartenant à un objet soient connectés entre eux. De plus, il y a généralement des fausses détections, qui se manifestent par des pixels noir isolés, ainsi que des pixels de l'objet non détectés, qui se manifestent par des « trous » dans le masque. Il est donc nécessaire de filtrer l'image binaire pour ne garder que la région correspondant à la main.

Pour filtrer l'image binaire résultante de la phase de détection de la couleur de peau plusieurs méthodes ont été appliqué voir (Tableau II.3). Au premier lieu nous essayons les opérateurs morphologiques : érosion suivie d'une dilatation (ouverture), l'érosion pour supprimer les éléments négligeables et la dilatation pour remplir les trous, l'image résultante n'est pas satisfaisante car l'érosion n'élimine pas tout l'élément subtil puis si en applique une suite d'érosion ces éléments sont éliminé mais a conséquence les trous qui se trouve à l'intérieur de l'objet sont agrandir et nous ne pouvons pas boucher ces trous avec la dilatation alors nous tentons avec le filtre médian 5*5, le résultat mieux que les opérateurs morphologiques mais le temps de calcul est un peu élevé mais n'influence pas sur le temps de réponse de notre système, et on à nous combinais les deux phases (opérateurs morphologiques, filtre médian 5*5) le résultat est un peu similaire avec ce que donne le filtre médian 5*5 et relativement coûteux en temps de calcul, et n'est pas très fiable pour notre système qui fonction en temps réel alors nous utilisons juste le filtre médian 5*5, le **Tableau III.2** montre les résultats que nous avons obtenu dans cette phase.

Chapitre III : Conception et implémentation

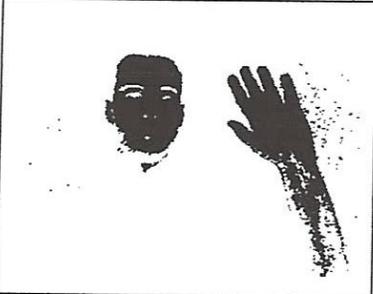
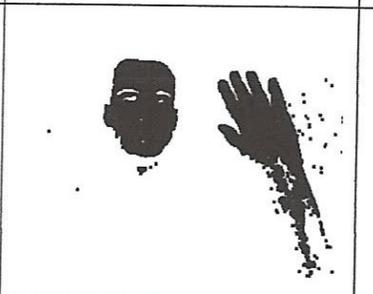
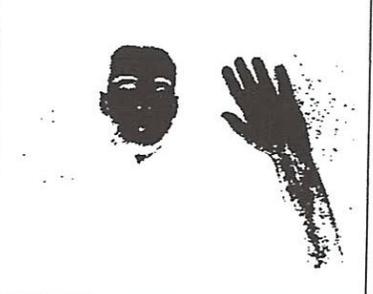
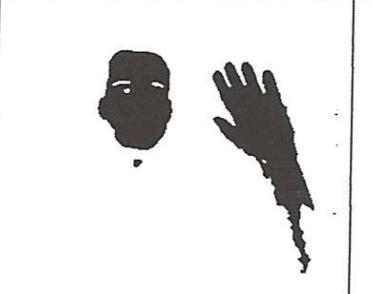
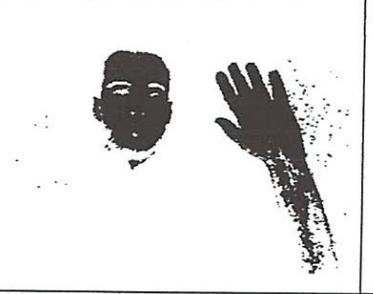
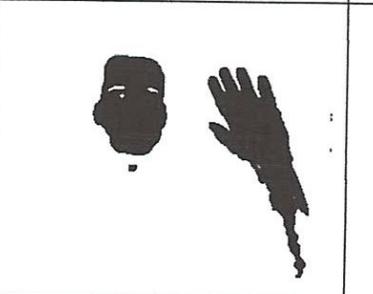
Image binaire (YCbCr)	Opération	Résultat	Temps de calcul
	Ouverture		47 ms
	Filtre médian 5*5		53 ms
	Ouverture + Filtre médian 5*5		100 ms

Tableau III. 2 : les différents résultats de filtrage.

1.2.3. Étiquetage des objets :

La plupart des systèmes de reconnaissance de geste ont minimisé la complexité de l'environnement par l'utilisation d'une seule main dans la scène comme : [1] [8] [30] [31] [36] [38] [40], à cause de la facilité des traitements et le taux de reconnaissance généralement élevé parce que il n'y pas des autres objets qui peuvent être détectés comme une main, une autre raison c'est la faiblesse de temps de calcul, mais avec cette condition le système devient extrémiste et n'est pas flexible.

Dans notre système nous ne prendrons pas cette condition en compte ce qui le rend robuste et fonction dans un environnement qui contient plusieurs objets (mains, visages), pour cela nous segmentons les objets de l'image binaire résultante de la phase précédente et de connaître la position, largeur et la hauteur de chaque objet segmenté, la Figure III.1 montre le résultat de l'étiquetage (chaque objet enveloppé par une boîte englobante).

Chapitre III : Conception et implémentation

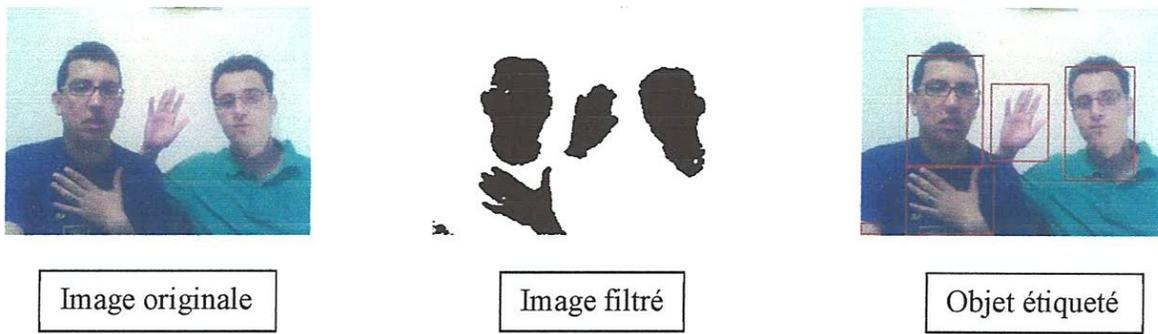


Figure III. 1 : résultat de la phase d'étiquetage

I.2.4. Élimination des objets non-concerne :

Dans la phase précédente on obtiennent une image binaire contenant des objets étiqueté, ces objets sont des mains et visages et aussi des éléments qui sont de petite taille et qui ne sont pas éliminer dans la phase de filtrage, en trouvant aussi dans cette phase des objets qui ont une grande largeur et une petite hauteur ou l'inverse qui ne présentera jamais une main, les objets avec une grande largeur et hauteur mais le nombre de pixel est très petit par rapport à leur dimension, dans ce cas-là nous éliminons tous ces objets non-concerne.

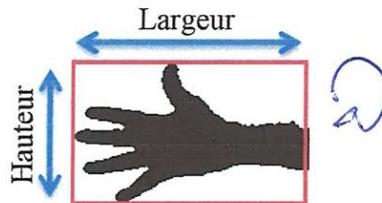


Figure III. 2 : largeur et la hauteur d'un objet

$$\text{Taux} = \frac{\text{hauteur}}{\text{largeur}} \quad \text{Si } (\text{Taux} > \text{seuil}_1 < \text{ou } \text{Taux} < \text{seuil}_2) \text{ alors éliminer l'objet}$$

$$\text{Pourcentage} = \frac{\text{nombre des pixels noir}}{\text{hauteur} * \text{largeur}} \quad \text{Si } (\text{Pourcentage} < \text{seuil}) \text{ alors éliminer l'objet}$$

Si (largeur < seuil ou hauteur < seuil) alors éliminer l'objet.

Supposons que les objets qui restent après cette phase sont des mains ou des visages, à la prochaine étape nous aurons une décision que vraiment sont des mains ou des visages et on élimine les objets qui ont une dimension ainsi que le pourcentage qui ne leurs permet pas de passé à cette étape, le **Tableau III.3** montre les résultats de cette phase.

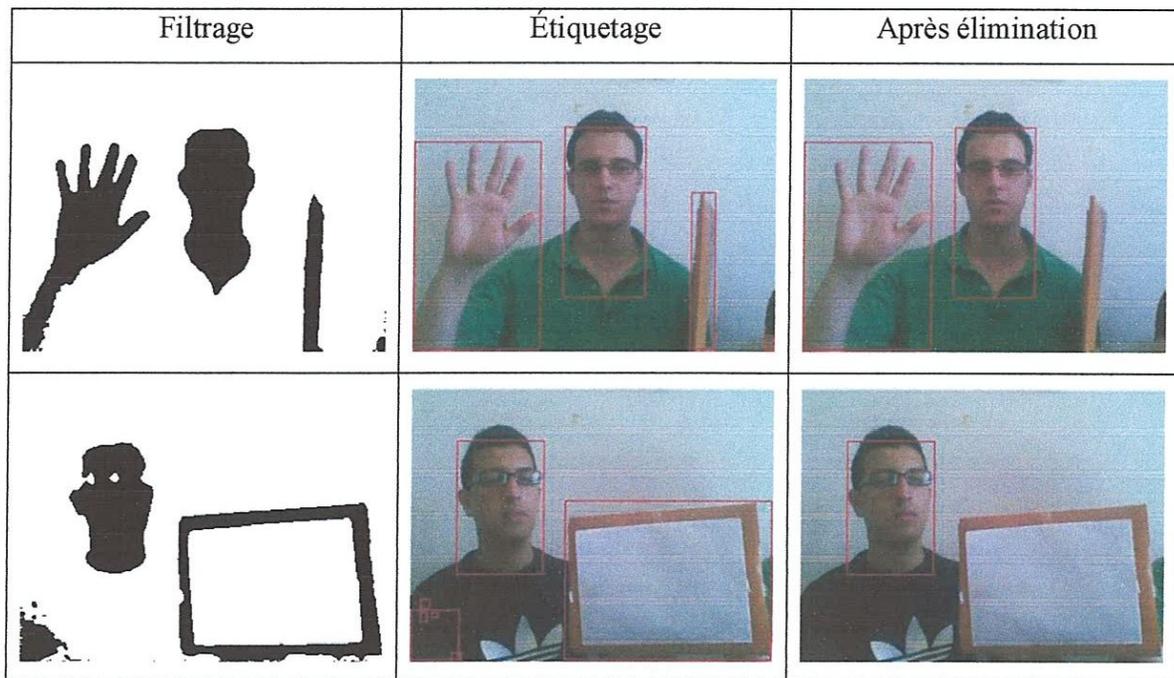


Tableau III. 3 : les objets qui restent après l'élimination

Dans le premier exemple l'objet qui est situé à gauche de visage est éliminer car sa hauteur est très grande par rapport à sa largeur, dans le deuxième exemple le cadre est éliminer parce que les pixels noir sont plus inférieure par rapport à leur taille, les autres objets à par le visage sont des fausses détections qu'ils ne sont pas effacer dans la phase de filtrage mais leurs tailles très petites alors ne seront pas introduit dans la prochaine étape.

I.3. Extraction des caractéristiques:

Dans cette étape nous nous intéressons à extraire les caractéristiques de chaque objet segmenter par l'étape précédente pour le reconnaître : main, visage ou fausse détection, ces caractéristiques sont utilisés par la suite pour l'interface gestuelle de notre système, différent méthode sont utilisé, dans notre cas le visage peut entrer dans la scène alors il faut le détecter pour le différencier avec la main.

I.3.1. Détection de visage :

La détection de visages en temps réel est un sujet de plus en plus exploré ces dernières années. Les chercheurs développent plusieurs algorithmes tels que Hough, HaarCascade, Template matching methods, La transformée de Hough elliptique autorise une vitesse de détection assez rapide pour envisager un traitement en temps réel, mais par rapport à d'autres algorithmes de segmentation de visages un des avantages de cette méthode réside dans la robustesse : elle permet de détecter des visages de petite taille, des visages atypiques (yeux

Chapitre III : Conception et implémentation

fermés, avec des lunettes) [43], dans notre traitement les yeux sont bouchés dans la phase de filtrage, alors on ne peut pas détecter le visage par l'algorithme de HaarCascade qui se base sur la détection des yeux [13], à cause de ça nous utilisons la transformée de Hough qui se base sur la forme elliptique de visage.

Dans la majorité des cas, les visages intéressants à détecter sont positionnés selon une direction verticale, donc il n'est pas indispensable de définir l'orientation précise des visages détectés. Une ellipse verticale a pour expression :

$$\frac{X^2}{lh^2} + \frac{Y^2}{lv^2} = 1 \quad (\text{III.1})$$

Avec : $X = x - x_c$ et $Y = y - y_c$

(x_c, y_c) : le centre de l'ellipse, (x, y) : un point de contour de l'ellipse, (lv, lh) : demi-hauteur, demi-largeur de l'ellipse [43], la **Figure III.3** montre le résultat de détection de visage par transformé de Hough.



Figure III. 3 : résultat de la détection de visage par la transformée de Hough.

À la fin de cette phase nous obtenons des visages, les autres objets ont une couleur de la peau et des dimensions qui leur permettent d'être des mains, mais on ne peut pas encore décider que ces objets sont des mains ou non; dans la phase suivante nous extrairons les caractéristiques morphologiques qui représentent la main pour décider.

1.3.2.Extraction des caractéristiques morphologiques :

Chaque main de l'être humain a un ensemble des caractéristiques morphologiques : (le centre, la paume, les bouts des doigts, les creux entre les doigts), dans cette phase nous nous intéressons de les extraire, pour chaque caractéristique une priorité de détection parce que on ne peut pas connaître certains spécifiquement sauf avec l'utilisation des autres.

Pour connaître les caractéristiques de la main nous suivons l'ordre qui se représente dans le schéma ci-dessous [8], [37], [14]:

Chapitre III : Conception et implémentation

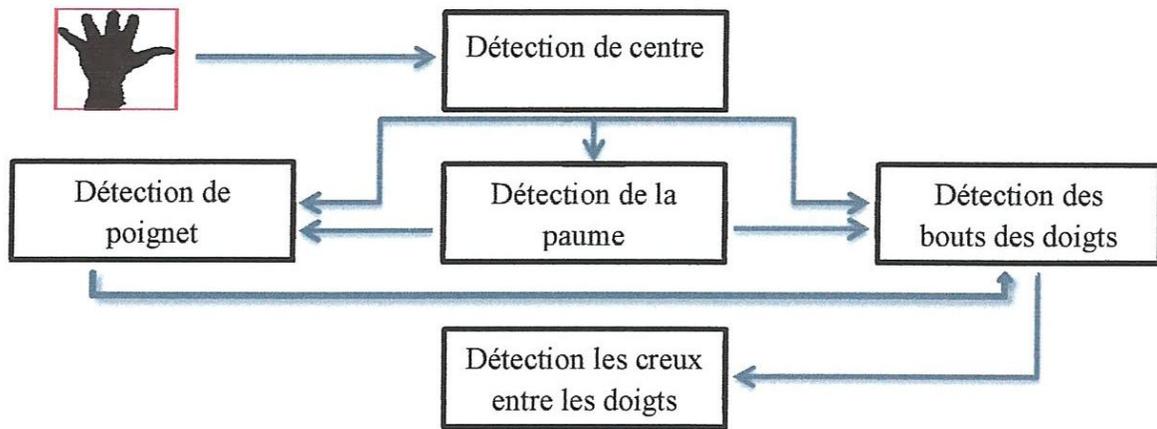


Schéma III.2 : les phases pour l'extraction des caractéristiques morphologiques

I.3.2.1. Centre de la main :

Le centre de la main est un point important car il est souvent le premier à être localisé et est très utilisé pour le suivi de la main. Il permet de représenter la position globale de la main, Il est aussi très utile de connaître la position du centre pour détecter les bouts des doigts, avec par exemple la distance au centre [8].

Un des problèmes pour la détection du centre est de savoir si l'utilisateur a le bras nu ou non. En effet, si celui-ci est nu cela pose des problèmes pour la détection des doigts et le calcul des caractéristiques de forme. Ce qui nous amène d'éviter d'imposer des contraintes aux utilisateurs telles que le port d'un vêtement à manches longues.

Différentes approches existent telles que : la détection avec la boîte englobante qui se base sur la l'intersection de l'axe verticale de demi-largeur avec l'axe horizontal de la demi-hauteur ce qui donne le centre de la boîte qui est le centre de la main, l'inconvénient de cette approche est que si l'utilisateur à le bras nu elle donne un centre qui ne correspondant pas au centre de la main, la figure suivante montre le résultat de cette méthode.

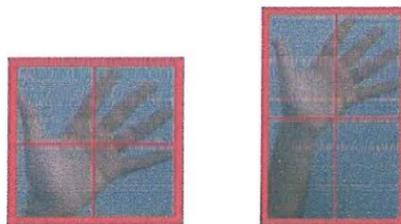


Figure III. 4: résultat de détection de centre avec la boîte englobante

Autre approche basé sur le calcul de centre de gravité avec les moments géométriques qui se basent sur la formulation classique des moments bidimensionnels, un moment d'ordre $p + q$ s'écrit :

Chapitre III : Conception et implémentation

$$m_{pq} = \sum_{(x,y)} x^p y^q I(x,y), \quad (p,q) \in \mathbb{N}^2 \quad (\text{III.2})$$

Le moment d'ordre 0, m_{00} , donne la surface de la main. Le centre de gravité (x_G, y_G) de la région de la main peut être calculé à partir des moments d'ordre 1 de cette région [8] :

$$(x_G, y_G) = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right) \quad (\text{III.3})$$

Cette méthode est influence par a la taille et la forme de la main si la forme de la main change (avant-bras nu) le centre de gravité ne correspond pas véritablement au centre de la main [14].

La carte de distance c'est l'approche qui utilise la transformée de distance d'une image binaire, cette dernier associe à chaque pixel de l'objet la distance (Euclidienne, Manhattan, d'échiquier, Minkowski, voir (chapitre II, Méthode de k plus proche voisin)) au pixel du contour le plus proche [14], le centre de la main correspond au pixel de l'objet qui a une grande distance **GD** par rapport d'autre pixel. Le temps de calcule est assez rapide par rapport les autre approche qui en a définir mais leur avantage c'est de ne pas prendre en compte la forme de la main (avant-bras), nous utilisons cette approche pour éviter d'impose des contraintes aux utilisateurs telles que le port d'un vêtement à manches longues la **Figure III.5** montre le résultat de détection de centre de la main par la carte de distance.

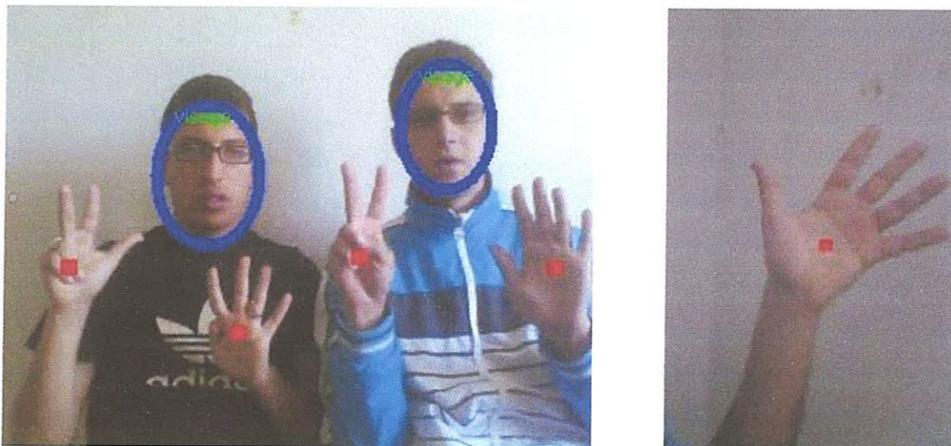


Figure III. 5 : résultat de détection de centre de la main (point rouge).

Pour diminuer les fausses détections nous éliminons tous les objets qui ont un **GD** très petit qui ne correspond jamais une main.

Chapitre III : Conception et implémentation

I.3.2.2. Détection de la paume :

La phase de détection de la paume est très importante car à l'aide de ce résultat on peut extraire d'autres caractéristiques (poignet, doigts), pour déterminer la paume de la main nous utilisons la distance **GD** obtenue ans la phase précédente comme un rayon de le cercle qui a un centre correspond au centre de la main. Le résultat de cette phase est illustré dans la **Figure III.6** :



Figure III. 6: résultat de détection de la paume

Le cercle vert représente la paume, le point rouge représente le centre de la paume et la main, la ligne bleue représente **GD** qui est le rayon de cercle.

I.3.2.3. Détection de poignet :

L'extraction de poignet est une phase qu'on ne peut pas éliminer dans notre système car leur résultat jeu un rôle important dans l'extraction des bouts des doigts et l'élimination de l'avant-bras alors pour obtenir des résultats satisfaisantes dans l'extraction de ces caractéristiques nous utilisons le cercle de la paume pour détecter le poignet, l'intersection de l'objet avec ce cercle au-dessous le centre de la main engendre les deux points extrémités de poignet.

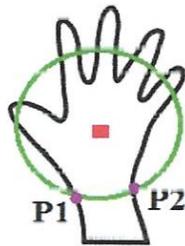


Figure III. 7 : exemple illustratif de détection de poignet

Par l'utilisation de cette méthode nous obtenons les résultats suivants :



Figure III. 8: résultat de détection de poignet

I.3.2.4. Détection des bouts des doigts :

Les bouts des doigts sont des points très importants pour les suivis et la reconnaissance des gestes. Pour les gestes de pointage, le doigt est un moyen simple et intuitif pour remplacer la souris, pour pointer sur un écran ou sur des surfaces interactives [8].

Il existe de nombreuses méthodes pour détecter les doigts basées sur le contour tel que la courbure du contour ou avec la distance au centre de la main.

La détection des doigts avec la courbure de contour basé sur le calcul de l'angle entre les vecteurs PP_1 et PP_2 qui tel que P est situé entre les deux point de contour avec la même distance si l'angle est dans un intervalle qui correspond l'angle de doigt alors en prend le point P comme un bout la figure suivante représente la détection avec la courbure du contour.

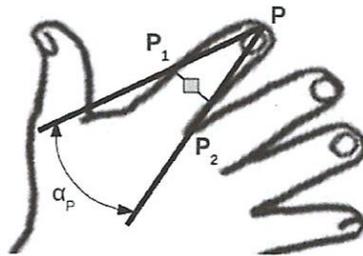


Figure III. 9: détection des bouts des doigts par la courbure du contour [37]

L'inconvénient de cette méthode c'est qu'elle donne des fausses détections, par exemple tous les points de contour qui ont une courbure comme les bouts des doigts.

Nous utilisons cette l'approche de détection avec la distance au centre de la main qui se base sur le calcul de distance (euclidienne) entre le centre de l'objet et les points de leur contour, les maximums locaux de la courbe de distance obtenue correspondent aux bouts des doigts avec cela on peut détecter plusieurs locaux qui ne correspondent pas aux bouts des doigts, pour éviter ce problème nous utilisons deux seuils :

Chapitre III : Conception et implémentation

Seuil1 = $GD \cdot Q_1$ tel que $Q_1 = 5/3$ [14]

Seuil2 = $GD \cdot Q_2$ tel que $Q_2 = 3$

Les bouts des doigts sont les points de contour qui ont des distances D avec le centre de l'objet inférieur à seuil2 et supérieur à seuil1.

Nous utilisons aussi pour la réduction des fausses détections un algorithme qui teste les pixels du contour, si les pixels qui sont situés au-dessus du pixel P sont du fond alors P est considéré comme un bout de doigt, la figure ci-dessous représente cet algorithme.

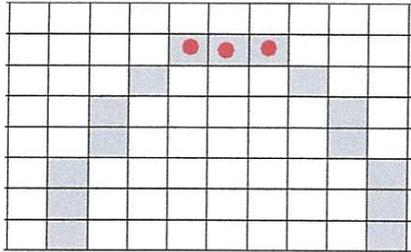


Figure III. 10: les pixels détectés comme des bouts des doigts

Les trois pixels résultant de la **Figure III. 10** dépassent les deux conditions précédentes, Pour tous les couples des pixels calculons, l'angle (pixel1, centre de la main, pixel2), S'il est inférieur à un seuil les deux pixels sont regroupés [14]. Le résultat de cet algorithme engendre un seul bout à chaque doigt, **Figure III. 11** le résultat de cette phase.



Figure III. 11: résultat final de détection des bouts des doigts

Le grand cercle est défini par le centre de la main et le rayon ($GD \cdot Q_1$), les points bleus sont les bouts des doigts détectés, les lignes jaunes ce les distances entre les bouts et le centre.

Chapitre III : Conception et implémentation

I.3.2.5. Détection des creux :

La détection des creux entre les doigts est une phase pour confirmer c'est les bouts des doigts qu'ont été détectés correspondent vraiment à des doigts, nous utilisons un algorithme qui consiste à définir le point P qui correspond le demi segment qui lier les deux bouts des doigts successif puis calcule la distance qui correspond au parcours verticale des pixels au-dessous de point P jusqu'à en arriver à un pixel Y de la main qui définit le creux. Si cette distance est inférieure à un seuil alors cet objet ne ressemble pas aux caractéristiques de la main qui a un vide entre les doigts, la figure au-dessous illustre cet algorithme.

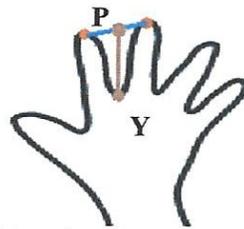


Figure III. 12: détection des creux entre les doigts

Nous utilisons cet algorithme dans notre système qui donne les résultats suivants :

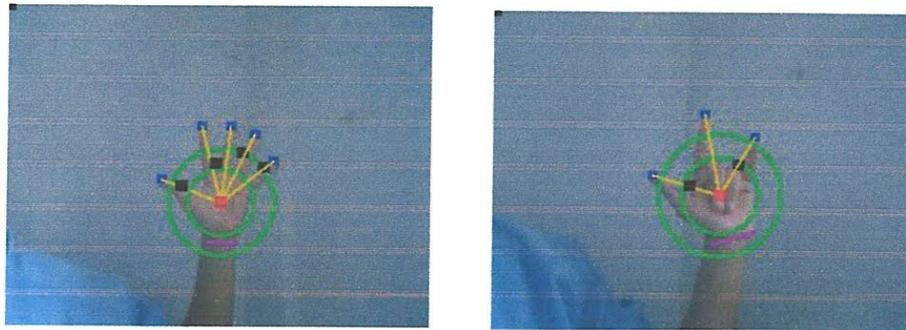


Figure III. 13: résultat de détection des creux (point noir)

Après l'étape de l'extraction des caractéristiques nous connaissons les objets main ou visage qui entre dans la scène, dans la prochaine étape notre objectif est la reconnaissance des gestes des mains pour l'utiliser dans l'interface gestuelle.

I.4. Classification:

À la fin de l'étape précédente nous obtenons le nombre des doigts, mais cette information n'est pas satisfaisante pour reconnaître le geste par exemple quand nous trouvons un seul doigt on ne peut pas différencier est ce que un pouce ou index ou autre doigt ? L'étape de classification consiste à distinguer chaque doigt qui engendre la reconnaissance de geste.

Chapitre III : Conception et implémentation

Il existe dans la littérature plusieurs méthodes pour la classification telle que : Machines à Vecteurs Supports (SVM), Les réseaux de neurones artificiels, Méthode de k plus proche voisin (KNN), Model de Markov cache (MMC), pour plus de détaille (chapitre II, Classification), ces approches sont utilisés par des travaux voisin pour la reconnaissance des gestes voir (Tableau II,4). Les méthodes qu'on a défini précédemment sont basé sur l'apprentissage supervisé et non supervisé, le premier consiste à définir un ensemble d'étiquette pour chaque geste et le classer par la comparaison avec ces étiquette [28], l'inconvénient de ces méthodes c'est qu'elles donnent un taux de reconnaissance faible quand l'utilisation de petites étiquettes et si on définit un grand nombre la comparaison prend beaucoup de temps et ça influence sur notre système qui fonctionne en temps réel.

L'apprentissage non supervisé appelée aussi (classification automatique) consiste à construire automatiquement les classes qui sont inconnues a priori l'inconvénient de cette approche c'est qu'elle prend beaucoup de temps pour la construction des classes.

La classification à base des règles consiste sur la création des règles avec l'utilisation des vecteurs des caractéristiques, et ceux qui correspondent avec la règle peuvent être considérés comme un résultat final, l'avantage majeur de cette approche est la vitesse d'exécution, au contraire des méthodes basé sur l'apprentissage qui prennent beaucoup de temps pour la reconnaissance, mais elle est difficile à réaliser car elle dépend de la capacité de l'être humain à générer plusieurs règles assurant l'obtention d'une reconnaissance parfaite.

Nous utilisons la classification à base des règles grâce à leur vitesse qui marche avec notre système en temps réel, pour cela nous proposons notre approche de reconnaissance de geste qui consiste à mesurer les angles entre les bouts des doigts et le centre de la main, cet angle fait la différence entre les doigts, la figure au-dessous montre notre approche.

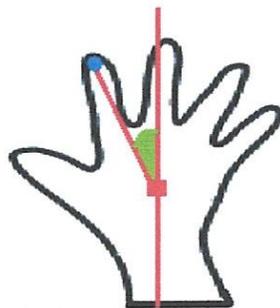


Figure III. 14: notre approche de reconnaissance a base des règles

Chapitre III : Conception et implémentation

L'angle est situé entre l'axe vertical et le bout de doigt comme montre la figure, ce qui pose un problème lors de la rotation de la main, cet angle varier et ne corresponde pas à notre règle qui nous définissons dans la figure suivante.

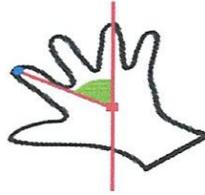


Figure III. 15: le changement de l'angle lors de la rotation de la main

Pour résoudre ce problème il faut trouver l'angle de rotation de la main pour faire une rotation inverse, [30] extraire cette angle avec une méthode qui consiste à trouver le pixel de contour plus haut de la main puis il calcul cette angle qui se situe entre le centre de la main C et ce point et l'axe verticale, car il travaille sur une main de cinq doigts mais cette approche ne donne pas le vrais angle dans quelle que contraintes. La figure suivant montre les résultats de cette approche:

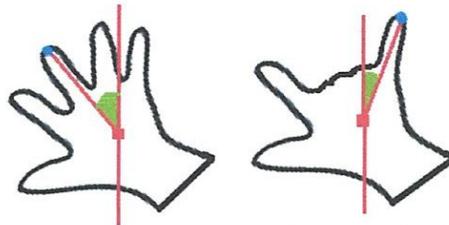


Figure III. 16: l'image à gauche (angle juste), image droite (angle erronée)

Nous utilisons le poignet pour calculer l'angle de rotation de la main, avec cette méthode en trouve cet angle entre le centre de la main et le poignet et l'axe verticale ; ça donne des résultats juste dans tous les cas, la figure au-dessous montre cet algorithme :

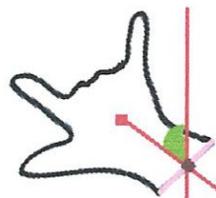


Figure III. 17: calcule de l'angle de rotation par le poignet

Après le calcul de l'angle de rotation de la main en fait une rotation inverse de la main par rapport à leur centre. la **Figure III.18** montre le résultat de la rotation inverse

Chapitre III : Conception et implémentation

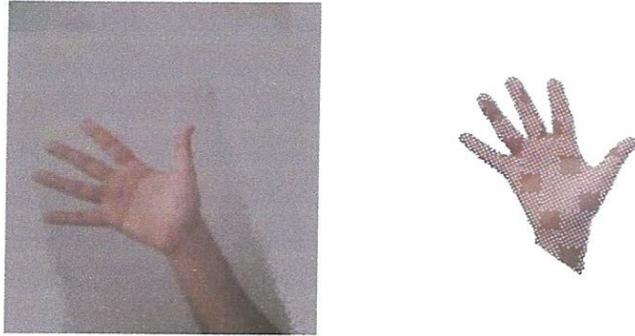


Figure III. 18: le résultat de la rotation inverse de la main

Après la rotation inverse nous pouvons appliquer notre approche pour reconnaître les gestes, les images ci-dessous représentent les résultats de reconnaissance de geste.



Figure III. 19: résultat de reconnaissance de geste

Chapitre III : Conception et implémentation

Nous utilisons le résultat de cette étape dans une interface gestuelle pour l'interaction homme-machine d'une manière plus naturelles et intuitives avec des gestes statiques et dynamiques dans la deuxième partie nous expliquons notre application.

I.5. taux de reconnaissance :

Nous obtenons le taux de reconnaissance par l'application des tests expérimental en temps réel sur dix gestes, le tableau suivant montre les résultats de la reconnaissance dans trois cas d'environnements différents:

1^{er} cas : - mauvais éclairage (non contrôler) – utilisateur non habituer – visage non présent sur la scène – bras nu.

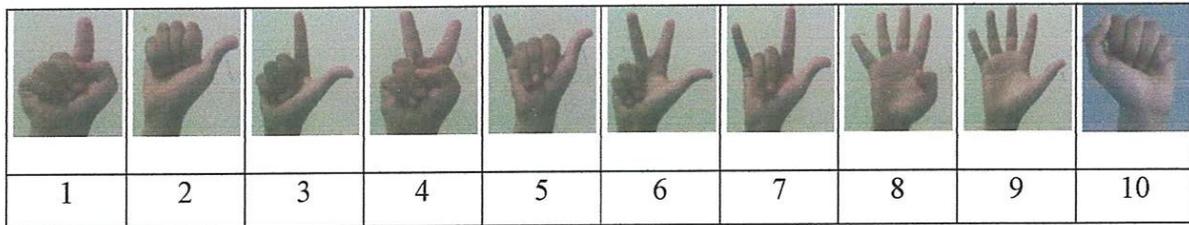


Tableau III. 4: symboles des gestes

Geste	Nombre des gestes testé	Nombre des gestes reconnues	Taux de reconnaissance	Temps de calcul moyenne
1	198	147	74,24%	0.139 s
2	177	149	84,18%	0.138 s
3	315	245	77,77%	0.145 s
4	335	305	91,04%	0.150 s
5	238	203	85,29%	0.146 s
6	338	282	83,43%	0.159 s
7	242	236	97,52%	0.150 s
8	195	175	89,74%	0.169 s
9	302	246	81,45%	0.189 s
Taux total	2340	1988	84,95%	0.153 s

Tableau III. 5: taux de reconnaissance cas 1.

Chapitre III : Conception et implémentation

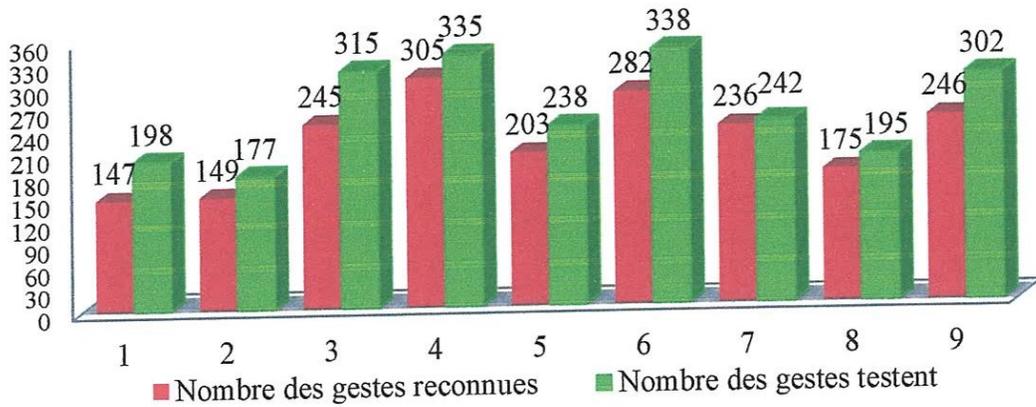


Figure III. 20 : taux de reconnaissance cas 1

2^{eme} cas : - bon éclairage – utilisateur habitué – visage non présent sur la scène – bras nu.

Geste	Nombre des gestes testé	Nombre des gestes reconnues	Taux de reconnaissance	Temps de calcul moyenne
1	220	216	98,18%	0.137 s
2	319	316	99,05%	0.137 s
3	243	241	99,17%	0.141 s
4	414	412	99,51%	0.144 s
5	211	203	96,20%	0.140 s
6	263	262	99,61%	0.155 s
7	410	409	99,75%	0.146 s
8	398	318	82,41%	0.166 s
9	309	300	97,08%	0.184 s
10	494	482	97,75%	0.128 s
Taux total	3281	3159	96,28%	0.147 s

Tableau III. 6: taux de reconnaissance cas 2

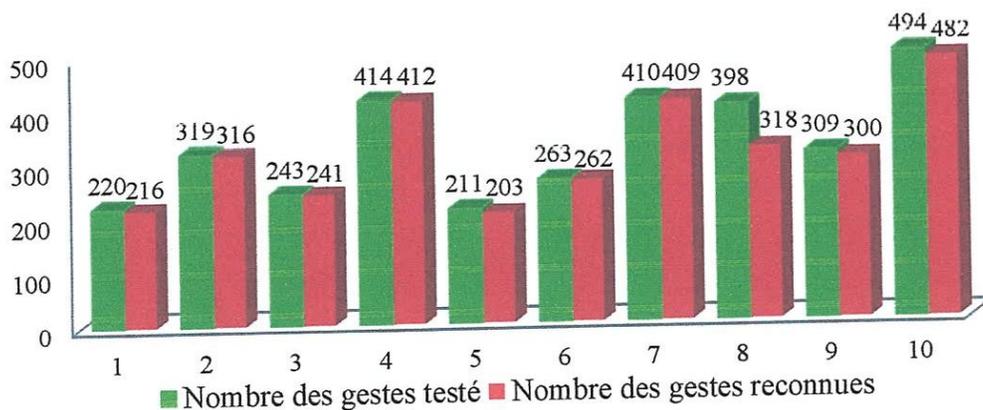


Figure III. 21: taux de reconnaissance cas 2

Chapitre III : Conception et implémentation

3^{eme} cas : - bon éclairage – utilisateur habitué – visage présent sur la scène – bras nu.

Geste	Nombre des gestes testé	Nombre des gestes reconnues	Taux de reconnaissance	Nombre de fausse détection de visage	Taux de fausse détection (visage)	Temps de calcul moyenne
1	218	210	96,33%	3	0,01%	0.177 s
2	332	327	98,49%	7	0,02%	0.163 s
3	269	268	99,62%	1	0,03%	0.176 s
4	248	243	97,98%	0	0%	0.174 s
5	251	241	96,01%	3	0,01%	0.177 s
6	301	298	99,66%	3	0,09%	0.161 s
7	264	261	97,31%	0	0%	0.176 s
8	212	188	88,67%	6	0,02%	0.226s
9	253	250	98,81%	4	0,075%	0.253 s
10	263	260	98,85%	1	0,03%	0.170 s
Taux total	2611	2546	97,51%	28	0,01%	0.185 s

Tableau III. 7: taux de reconnaissance cas 3

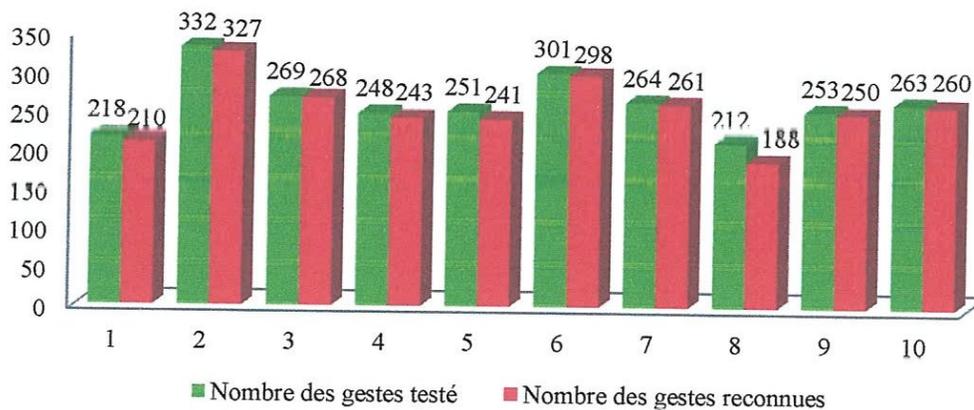


Figure III. 22: taux de reconnaissance cas 3

Chapitre III : Conception et implémentation

Partie II: Implémentation

II.1. Environnement de travail :

☛ Matérielle:

Dans ce travail nous utilisons un ordinateur qui a les références suivantes :

- ☞ Processeur : Intel®Core™i5-2430M CPU @2.40 GHz.
- ☞ Mémoire installée(RAM): 4.00 Go.
- ☞ Camera : HD Webcam Intègre 2 Méga pixels.

☛ logiciel :

Il y a plusieurs langages de programmation pour le développement des systèmes de reconnaissance de geste en temps réel on cite : JAVA, C#.NET, Python Nous utilisons C++ pour sa force et surtout pour les caractéristiques citées ci-dessous :

- **Très répandu** : Il fait partie des langages de programmation les plus utilisés sur la planète pour cela nous trouvons beaucoup de documentation. [44]
- **Rapide** : Très rapide même, ce qui en fait un langage de choix pour les applications critiques qui ont besoin de performances. C'est en particulier le cas des jeux vidéo, dans notre cas les systèmes de reconnaissance de geste qui doivent fonctionner en temps réel, et qui demande des calculs gourmands en temps et complexes. [44]
- **Portable** : Un même code source peut théoriquement être transformé sans problème en exécutable sous Windows, Mac OS et Linux. Vous n'aurez pas besoin de réécrire votre programme pour d'autres plates-formes. [44]
- **Nombreuses bibliothèques** : Pour le C++. Les bibliothèques sont des extensions pour le langage, un peu comme des plug-ins. en le combinant avec de bonnes bibliothèques, on peut créer des programmes 3D, réseaux, audio, fenêtrés, etc. [44]

Nous utilisons la bibliothèque *Open CV (version 2.4.5)* pour l'acquisition de flux vidéo et la bibliothèque **QT** qui est développée par NOKIA :

- ✚ Qt est en fait bien plus qu'une bibliothèque : c'est un Framework, qui contient un module GUI (Graphical User Interface), un module réseau, un module SQL,... etc.
- ✚ Qt est multiplateforme : on peut l'utiliser aussi bien sous Windows que Linux et Mac OS X. [44]

Chapitre III : Conception et implémentation



Figure III. 23 : Interface de QT.

II.2. Description de l'application :

Nous implémentons notre système de reconnaissance des gestes dans quatre sous application en temps réel pour remplacer les périphériques usuels (écran tactile, clavier, souris), dans l'interface de ce système l'utilisateur choisir un environnement d'application pour exécuter ces gestes de pointages et de manipulation directement sur une application réelle, la Figure III. 21 montre cette interface:



Figure III. 24: l'interface de notre application.

Le groupe outils (1) contient 5 boutons :

Chapitre III : Conception et implémentation

- Démarrer/pause : lancer le traitement pour le flux vidéo (25 images/seconde) et afficher le résultat de détection ou arrêter le traitement.
- Charger : charger une image pour le traitement.
- Traiter : traitement de l'image charger.
- Sauvegarder : sauvegarder l'image résultante.
- Fermer : quitter l'application.

Le groupe seuil de la peau (2) contient des radios bouton pour choisir l'espace colorimétrique.

Le groupe application (3) pour choisir l'application :

- Diaporama : nous remplaçons les flèches droites et gouaches du clavier par un geste dynamique d'une main avec un seul doigt, qui est l'index. (déplacement à droite remplace la flèche droite, déplacement à gauche remplace la flèche gauche).
- Météo : nous manipulons une carte météo par des gestes de pointage et dynamique pour glisser et déplacer des objets (soleil, nuage, vent...), le doigt index pour le pointage sur notre objet puis en le glisser par le geste 4 et en arrêt le glissement par une main fermer (drag and drop) nous développons ce système pour simuler l'écran tactile.
- Souris : nous remplaçons la souris classique de l'ordinateur et quelque raccourci de clavier par un ensemble de gestes dynamiques et statiques qui sont illustrés dans le tableau ci-dessous.

							
Déplacer la souris	clique gauche	clique droite	double clique	sélection	copier	coller	arrêter l'action

Tableau III. 8: les gestes pour remplacer la souris et clavier.

- Paint : l'utilisateur peut dessiner avec le doigt index et déplacer le dessin avec le geste 4, agrandir et réduire par un geste dynamique 3 pour effacer le dessin par une main ouverte.

Le groupe paramètre (4) qui permet à l'utilisateur de changer l'intervalle de la couleur de la peau YCbCr.

Le groupe affichage (5) pour permet à l'utilisateur d'afficher :

- Image peau : afficher le flux vidéo du résultat de détection de la peau après le filtrage.

Chapitre III : Conception et implémentation

- Angle rotation : afficher la rotation inverse des bouts des doigts et les angles entre le centre et la main.
- Nombre d'objets : affiche dans le groupe objet (7) le nombre de visages et le nombre des mains avec leurs nombres des doigts ainsi que le nombre total des doigts existant. Le groupe vitesse de la souris (6) permet de changer la vitesse de déplacement de la souris, et un bouton aide pour apprendre le fonctionnement de l'application.

II.3. Testes et résultats :

Dans cette section nous présentons nos tests et leurs résultats :

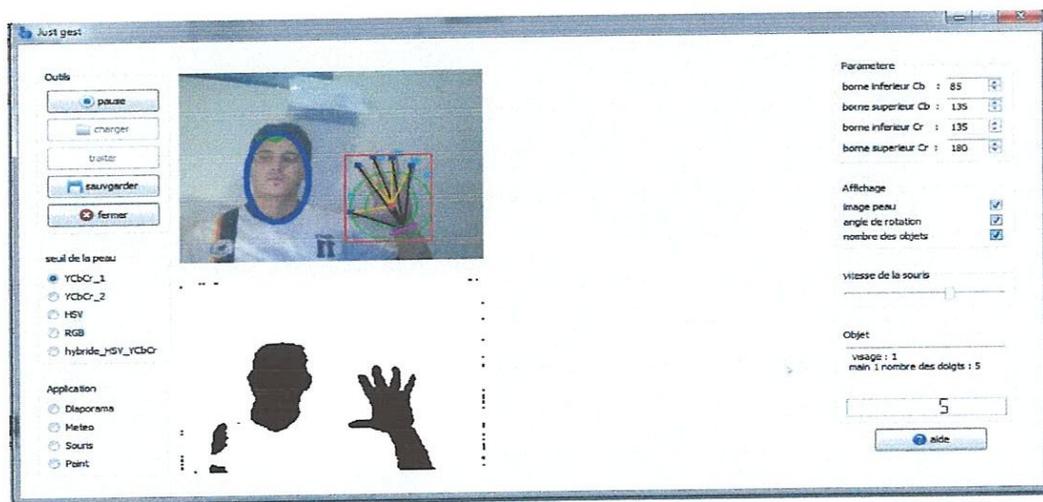


Figure III. 25 : fenêtre principale quand l'utilisateur clique sur démarrer

Quand l'utilisateur clique sur le bouton démarrer le système commence le traitement et affiche dans la même zone de flux vidéo des résultats, la boîte englobante de la main avec la couleur rouge, le centre de la main avec point rouge, le poigné avec la couleur rose, la paume de la main avec un petit cercle vert, un deuxième cercle plus grand utilisé pour améliorer la détection des bouts de doigts qui seront marquer des points de couleur bleue, la distance entre le centre et les bouts des doigts avec le jaune, les points de la couleur cyan sont les points qui représentent la rotation inversée des bouts des doigts, l'ellipse bleue définir le visage,

La zone de flux vidéo (noire et blanc) représente le résultat de traitement après la phase de filtrage.

Le groupe (7) affiche dans leur zone le nombre des visages (1 visage), le nombre des mains et le total des doigts détectés dans les mains (1 main et 5 doigts).

Les figures suivantes représentent la carte météo et le paint.

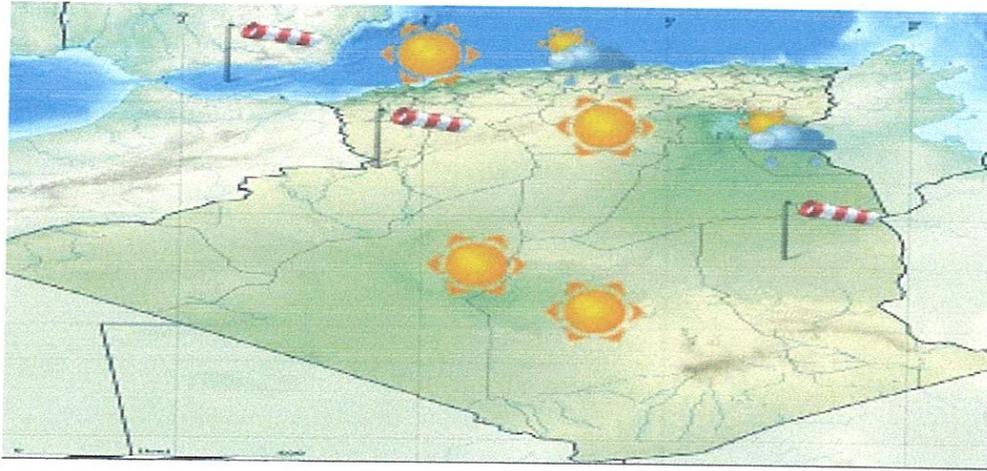


Figure III. 26: carte météo.

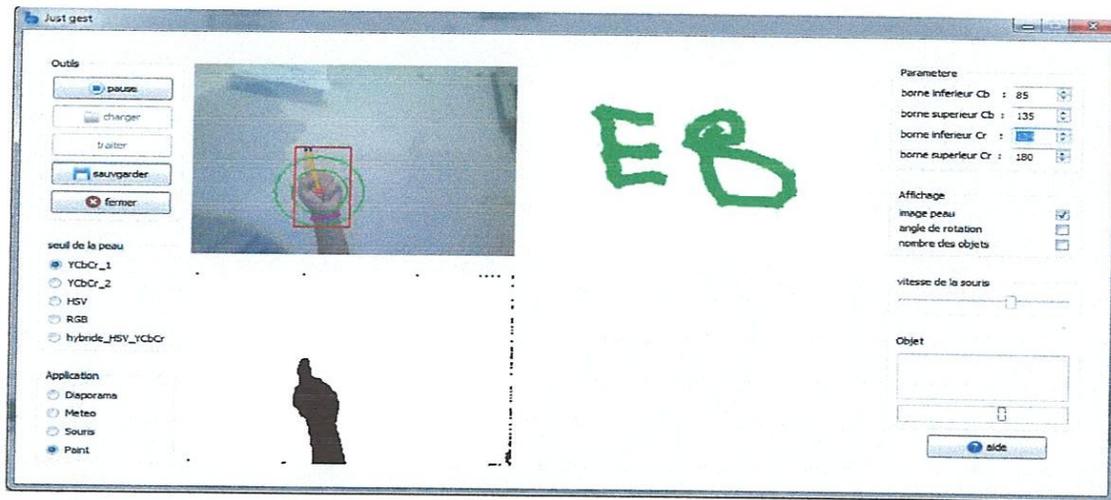


Figure III. 27: Paint

II. Conclusion :

Dans ce chapitre nous avons réalisé un système de reconnaissance de geste appliqué au Interface Homme-Machine pour remplacer le fonctionnement d'un écran tactile ou d'une souris, et qui fonctionne en temps réel.

Notre travail se concentre sur le mouvement de la paume de la main et les doigts, avant de reconnaître la posture, les contraintes théoriques et matérielles affectant notre travail, l'ont orientés vers une approche par apparence.

Les différentes étapes abordées concernent la détection de la main basée sur la couleur de la peau, extraction de caractéristiques invariantes, la comparaison de descripteurs de forme pour la reconnaissance de posture, et le suivi bidimensionnel du mouvement des doigts et de la main. Les algorithmes et l'architecture proposée à été programmée et à donner de très bons résultats, que ce soit d'un point de vue temps ou performance tout en respectant les

Chapitre III : Conception et implémentation

principes des interfaces homme-machine usuelles, et aussi en surmontant tous les contraintes imposé par un système de vision par ordinateur travaillant en temps réels, dans un environnement complexe, et surtout en essaient de ne pas imposer des contrainte de travail sur l'utilisateur pour le rendre plus à l'aise.

Conclusion générale :

Dans ce mémoire, nous nous sommes intéressés à la réalisation d'un système: Détection, suivi bidimensionnel et reconnaissance des gestes pour les interfaces Homme-Machine augmentées, qui se base sur l'utilisation des gestes reconnues pour interagir avec l'ordinateur.

Notre but c'est de remplacer les fonctionnalités de la souris par les gestes de la main, la réalisation de ce travail s'est faite à travers l'utilisation de certaines étapes qui ont englobé différentes méthodes, la première étape : la segmentation de la main à travers la détection de la couleur de la peau, ensuite nous avons évoqué l'extraction des caractéristiques de la main et visage puis on a utilisé ces caractéristiques extraites pour l'étape de classification qui par son rôle résulte le geste reconnu pour qu'on l'utilise dans étape d'interface gestuelle, on na teste notre application dans quelle que domaine réel.

Les problèmes qui on a pris en compte dans la réalisation de notre application sont : Problème de variation de l'éclairage qui rendre la détection de la couleur de la peau faible, la complexité de l'environnement importe quelle objets peut entrer dans le champ de caméra, le conflit de visage et main avec bras nu qui porte une fausse détection de main, puis le comprimé de temps réel-performance, la facilite de manipulation pour l'utilisateur.

Les difficultés techniques, extraction des caractéristiques universelles pour rendre le taux de la reconnaissance élevé, pour se la dans notre perspectives utilisation d'un capteur infrarouge pour donner des meilleurs résultats dans l'étape de la segmentation.

Parmi nos proche perceptive est de l'application de notre système de reconnaissance gestuelle pour la langue des signes arabe et l'intégré dans les réseaux sociaux pour aider les sourds-muets pour communiquer avec les personnes normaux.

Bibliographie :

- [1] Jérôme Martin, « Reconnaissance de gestes en vision par ordinateur ». Pour obtenir de grade de Docteur en L'INPG, Spécialité Imagerie Vision Robotique, le 13 juillet 2000.
- [2] C.Cadoz , « le Geste Canal de Communication homme-machine », La Communication «Instrumentale» » Techniques et sciences Informatique, 1994.
- [3] F.K.H.Quek.Toward «Vision –Based Hand Gesture Interface», Dans Virtual Reality Software and Technology, 1994.
- [4] P. A. Harling et A. D. N. Edwards, « Hand tension as a gesture segmentation cue»,. In Proc. of Gesture Workshop, 1996.
- [5] Bruno, Bossard, « Problèmes posés par la reconnaissance de gestes en Langue des Signes », thèse de doctorat, 2000.
- [6] Vogler C, Metaxas D.N,Toward, «Scalability in ASL Recognition: Breaking Down Signs into Phonemes», ActesdeGEsture Workshop'99, Springer, LNAI 1739, 211-224,1999.
- [7] Vogler C, Metaxas D.N, « Parallel Hiden Markov Models for American Sign Language Recognition», Actes de IEEE International Conference on Computer Vision, 116-122,1999.
- [8] Simon Conseil, « Suivi Tridimensionnelle de La main et Reconnaissance de Gestes Pour Les Interface Homme-Machine ». Thèse pour obtenir le grade de Docteur de l'université paul cézanne, faculté des sciences et techniques, 13 mars 2008.
- [9] Paramod K, Mishra, Mokhtar M, Hassan,Robust, « Gesture Recognition Using Gaussian Distribution for Features Fitting», International journal of Machine Learning and Computing, Vol. 2, 3, June 2012.
- [10] Pedro T, Jorge L, P.Barreto , « Hand Gesture Recognition Using Color and Depth images enhanced with Hand angular Pose Data», Hamburg, Germany, IEEE International Conferenceon Multisensor Fusion and Integration for intelligent Systems (MFI) September 13-15,2012.
- [11] M.G.Ligozart, « Reconnaissance et compréhension de gestes, application à la langue des Signes ».These pour l'obtention du titre de Docteur de l'uviversité de Paris-XI, 1996.

- [12] R.Gautam , « A study on hand gesture recognition technique».department of electronics and communication engenering, june 2012.
- [13] Pham Van Viet, « Détection de visages 2D dans des poses frontales et non-frontales », Mémoire de fin d'études, Institut de la Francophonie pour l'Informatique, 17/9/2010.
- [14] Thomas Burger, « Reconnaissance Automatique Des Gestes De la Langue Française Parlée Complétée », le 26/10/2007.
- [15] Rafiqul Z. Khan, Noor A. Ibraheem., « Hand Gesture Recognition: A Literature Review», International Journal of Artificial Intelligence & Applications And Information Science,(IJAIA), Vol. 3, 2012.
- [16] Rafiqul Z.K, Noor A Ibrahim , « Survey on Gesture Recognition for Hand Image Postures », Computer and information Science, Vol. 5, May 2012.
- [17] Alain Boucher, « traitement d'image, segmentation », cours.
- [18] Bencheriet Chemesse- Ennehar , Boualleg Abd El halim, Tebbikh Hicham, «Segmentation de la Couleur de Peau par Seuillage Selon Différents Espaces de Couleur »,LAIG, Université 8 Mai 45 de Guelma, 3èmes Journées Internationales sur l'Informatique Graphique, JIG'2007.
- [19] Pooja Sharma : 1Dept. of Computer Science, Jagannath University, Jaipur, Rajasthan, India, Dr. Veena Yadav : Dept. of Computer Science, Global Institute of Engg. & Tech., Jaipur, Rajasthan, India, « Discrimination between Skin and Non- Skin Pixels in Image Using the Range of HSV Color Space », International Journal of Computer Science and Technology, ISSN: 2229-4333 (Print) IJCST Vol. 4, Iss ue 1, Jan - March 2013.
- [20] Samir Hallaci, « Cours de Vision par Ordinateur », 1^{er} année Master université de 8 mai 1945 guelma, 2013.
- [21] S.Chitra , « Comparative Study for Two Color Spaces HSCBCR and YCBCR in Skin Color Destection » ,Applid mathematical Scienses,Vol 6,no. 85,4229-4238,2012.
- [22] Lingrand Diane, « Cours de Traitement D'images », Rapport de recherche ISRN I3S/RR-2004-05-FR

- [23] Mokhtar M. Hasan , and Pramod K. Mishra, « hand gesture modeling and recognition using geometric features », Computer Science Department, Banaras Hindu University, Varanasi, Uttar Pradesh, India, Canadian Journal on Image Processing and Computer Vision Vol. 3 No. 1, March 2012.
- [24] Anis Ben Ishak «Sélection de variables par les machines à vecteurs supports pour la discrimination binaire et multiclasse en grande dimension», thèse de doctorat, le 6 septembre 2007.
- [25] Ferroudji Karim, « Classification des particules par ultrason et réseaux de neurones artificiels (Application aux emboles) », Présenté en vue de l'obtention du diplôme de Magister en Electronique OPTION : Micro-onde pour Télécommunication, Dans le cadre du projet de coopération INSERM (France)/DPGRF (Algérie) , Année universitaire : 2009/2010.
- [26]: Paulo Cortez, Miguel Rocha, José Neves, University of Minho, Portugal, « Time Series Forecasting by Evolutionary Neural Networks » , 2006.
- [27] Badr Benmammar « Intelligence Artificielle et Systèmes Multi-Agents », 16 jan 2012.
- [28] Komal Sheth, Nitish Gadgil, P. R. Futane, « A Survey of Gesture Recognition Systems for Indian Sign Language Recognition », International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-2, Issue-6, August 2013.
- [29] G. R. S. Murthy, R. S. Jadon , «A Review of Vision Based Hand Gestures Recognition », International Journal of Information Technology and Knowledge Management, Vol. 2(2), pp. 405-410, 2009
- [30] Cristina Manresa, Javier Varona, Ramon Mas and Francisco J. Perales, « Hand Tracking and Gesture Recognition for Human-Computer Interaction », 07122 – Palma de Mallorca – España, Electronic Letters on Computer Vision and Image Analysis 5(3):96-104, 2005
- [31] X. Zabulis, H. Baltzakis, A. Argyros. « Vision-based Hand Gesture Recognition for Human-Computer Interaction », Institute of Computer Science Foundation for Research and Technology - Hellas (FORTH) Heraklion, Crete, Greece.
- [32] Matilde Gonzalez Preciado, « Computer Vision Methods for Unconstrained Gesture Recognition in the Context of Sign Language Annotation », these de doctorat, universitè de Toulouse France , le 24/09/2012

- [33] Rafiqul Zaman Khan, and Noor Adnan Ibraheem, « Comparative study of hand gesture recognition system », Department of Computer Science, A.M.U., Aligarh, India, 2012.
- [34] Prajakta M.Patil, Y.M.Patil; « Robust Skin Colour Detection And Tracking Algorithm »; International Journal of Engineering Research & Technology (IJERT), ISSN: 2278-0181, Vol. 1 Issue 8, October – 2012.
- [35] L.Bréthes, P.Menezes, F.Lerasle , M.Briot , « Segmentation couleur et condensation pour le suivi et la reconnaissance de gestes humains », avenue Colonel Roche, 31077 Toulouse Cédex 4,2003.
- [36] Archana S, Gajanan K. Kharate Department of Electronics and Telecommunication Engineering Matoshri College of Engineering and Research Centre University of Pune Nasik 422 105, India, « Hand Segmentation Techniques to Hand Gesture Recognition for Natural Human Computer Interaction », International Journal of Human Computer Interaction (IJHCI), Volume (3) : Issue (1) : 2012
- [37] José Manuel Palacios, Carlos Sagués, « Eduardo Montijano and Sergio Llorente, « Human-Computer Interaction Based on Hand Gestures Using RGB-D Sensors », ISSN 1424 8220, Published: 6 September 2013.
- [38] P. Jenifer Martina, P. Nagarajan, P. Karthikeyan, Dept of ECE, PSNA College of Engineering and Technology Dindigul, Tamilnadu, India, « Hand Gesture Recognition Based Real-time Command System », IJCSMC, Vol. 2, Issue. 4, April 2013.
- [39] E. Stergiopoulou, N. Papamarkos. « Hand gesture recognition using a neural network shape fitting technique » Elsevier Engineering Applications of Artificial Intelligence, vol. 22(8), pp. 1141– 1158, doi: 10.1016/j.engappai.2009.
- [40] Mokhtar M. Hasan, Pramoud K. Misra, « Brightness Factor Matching For Gesture Recognition System Using Scaled Normalization », International Journal of Computer Science & Information Technology (IJCSIT), Vol. 3(2) 2011.
- [41] Simej G. Wysoski, Marcus V. Lamar, Susumu Kuroyanagi, Akira Iwata, « A Rotation Invariant Approach On Static-Gesture Recognition Using Boundary Histograms And Neural Networks », IEEE Proceedings of the 9th International Conference on Neural Information Processing, Singapura 2002.

[42] Ray Lockton, Balliol College. «Hand gesture recognition using computer vision», 4th year project report, Oxford University,.

[43] Fan YANG, Michel PAINDAVOINE. «Détection et localisation de visages dans une scène : implantation parallèle sur un réseau de DSPs», Laboratoire Le2i, aile de l'Ingénieur - Mirande, Université de Bourgogne, France,2000.

[44] Mathieu Nebra et Matthieu Schaller « Programmez avec le langage C++ », Licence Creative Commons 6 2.0, 28/07/2013.