I. Introduction :

L'analyse du visage est une discipline en plein essor dans la communauté du traitement d'images et de la vision par ordinateur. Nous évoquons dans ce chapitre quelques définitions qui nous aiderons dans la démarche à suivre : l'image , le visage , l'émotion et l'expression faciale.

Du point de vue physiologique, une expression faciale résulte de la déformation des traits du visage provoquée par une émotion. L'essentiel de l'information d'une expression est contenu dans la déformation des traits permanents principaux du visage, à savoir les yeux, le nez, les sourcils et la bouche.

II. Définitions :

II.1. L'IMAGE

II.1.a. Définition d'une image:

C'est la représentation d'une personne ou d'un objet par la peinture, la sculpture, la photographie, etc. C'est aussi un ensemble structuré d'informations qui après affichage sur écran, ont une signification pour l'œil humain. Elle peut être écrite sous forme de fonction d'intensité lumineuse ou de couleur aux coordonnées spatiales. De cette façon l'image est exploitable par la machine, d'où la nécessité de sa numérisation.[23]

II.1.b. L'image numérique:

les images manipulées par l'ordinateur sont numériques (série de bits) qui représente l'image dont la surface est divisée en éléments de taille fixe appelés pixels, ayant comme caractéristique le niveau de gris ou de couleur.Une image numérique est ensemble de pixels contenant chacun différentes informations (intensité lumineuse, couleur…). Ces informations sont codées dans une grille échelonnée. Selon la nature de cette grille on peut obtenir différents types d'images. [24]

II.1.c. Caractéristiques d'une image numérique:

- **Pixel:** Contraction de l'expression anglaise " Picture Eléments ": éléments d'image, le pixel est le plus petit point de l'image, c'est une entité calculable qui peut recevoir une structure et une quantification. Si le bit est la plus petite unité d'information que peut traiter un ordinateur, le pixel est le plus petit élément que peuvent manipuler les matériels et logiciels d'affichage ou d'impression.
- **Dimension**: C'est la taille de l'image. Cette dernière se présente sous forme de matrice dont les éléments sont des valeurs numériques représentant les intensités lumineuses (pixel).
- **Résolution**: C'est la clarté ou la finesse des détails atteinte par un moniteur ou une imprimante dans la production d'image, sur les moniteurs d'ordinateurs, la résolution est exprimée en nombre de pixel par unité de mesure (pouce ou centimètre). On utilise aussi le mot

résolution pour désigner le nombre total de pixels affichable horizontalement ou verticalement sur un moniteur; plus grand est ce nombre, meilleure est la résolution.[25]

- **Niveaux de gris:** C'est la valeur numérique qui reflète l'intensité de la luminosité d'un point. Pour 256 niveaux de gris compris entre 0 et 255, chaque pixel sera codé sur 8 bits, et ses niveaux de gris seront obtenus après dégradation de la couleur noire, 0 représente le blanc et 255 représente le noir.
- **Image en couleur:** C'est une image ou chaque pixel est codé dans l'espace de couleur RGB (rouge, vert, bleu). Donc c'est une représentation dans un espace tridimensionnel de la valeur d'intensité lumineuse du pixel. Ce dernier sera codé sur trois octets, un pour chacune des couleurs.
- **Bruit:** C'est un phénomène de brusques variations d'un pixel par rapport à ces voisins suivant un certain seuil. Il existe quatre sources de dégradation induisant le bruit, qui sont: le bruit lié au contexte d'acquisition, le bruit lié au capteur, le bruit lié à l'échantillonnage, et celui lié à la nature de la scène.
- **Histogramme:** est une fonction qui donne la fréquence d'apparition de chaque niveau de gris (couleur) dans l'image. Pour diminuer l'erreur de quantification, pour comparer deux images obtenues sous des éclairages différentes, ou encore pour mesurer certaines propriétés sur une image, on modifie souvent l'histogramme correspondant. Il permet de donner un grand nombre d'informations sur la distribution des niveaux de gris (couleur) et de voir entre quelles bornes est repartie la majorité des niveaux de gris (couleurs) dans le cas d'une image trop claire ou d'une image trop foncée.[26]

II.1.d. Prétraitement et amélioration d'images :

L'image brute est une structure simple dans son organisation mais complexe dans son contenu. Ceci est dû à plusieurs facteurs, la grande quantité d'informations qui y réside, et le nombre important de processus de l'utilisateur qui tentent à introduire des anormalités. On peut citer comme élément de perturbation: l'éclairage qui peut être trop faible et rend l'image trop sombre, le bruit électronique inhérent au capteur qui engendre une imprécision dans la restitution des niveaux de gris et la qualité optique de l'objectif qui peut engendrer des aberrations.

- **Binarisation:** C'est la transformation de l'image en plusieurs niveaux de gris en une image binaire . Elle peut être réalisée en choisissant un certain seuil, et tous les pixels ayant des valeurs inférieures à celui-ci deviennent noirs, tous ceux de valeurs supérieures deviennent blancs.
- **Egalisation de l'histogramme :** procède de la manière suivante: On part du niveau zéro et on calcule la valeur cumulée des niveaux suivants dans l'histogramme original

jusqu'au moment où la somme est la plus près possible d'une certaine valeur.ces la solution pour résoudre les problèmes des régions sombres d'une image qui sont difficilement perceptibles.

- **Filtrage:** Tout ce qui est indésirable dans une image est considéré comme un phénomène de variation brusque d'un pixel qui peut provenir de l'éclairage, des dispositifs optiques et électroniques du capteur ou de l'image elle-même. Pour lutter contre les effets du bruit, il est nécessaire de faire des opérations de filtrage sur chaque pixel en tenant compte de son voisinage. On distingue deux types de filtrage: le filtrage linéaire où la transformation d'un pixel est le fruit d'une combinaison linéaire des pixels voisins, et le filtrage non linéaire où les pixels voisins interviennent suivant une loi non linéaire. **[27]**

II.2. Le visage

« Le visage est le lieu même de l'expression et de l'identité» **[28]**

Pour mieux connaitre le visage et ses gestualités, on va s'intéressé aux définitions suivantes:

- **Gesticulation faciale (mimique) :** une mimique faciale est un état du visage composé par un ensemble de configurations des muscles faciaux, Le sourire est par exemple une mimique faciale composé d'un certain nombre d'activation des muscles faciaux (mouvements des muscles zygomatiques Muscle de la pommette de la joue, qui se contracte notamment lorsqu'une personne rit ou sourit).
- **Emotion :** l'émotion est un des générateurs des expressions faciales qui se traduit via de nombreux canaux comme la position du corps, la voix et les expressions faciales. Une émotion implique généralement une expression faciale correspondante
- **une expression :**. Alors que les expressions dépendent des individus et des cultures, on distingue généralement un nombre limité d'émotions universellement reconnues.
- **Expression faciale :** une expression faciale est une mimique faciale chargée de sens représentant une émotion sans pour autant ressentir cette émotion. [29]

II.2.1. Description d'un Visage :

Faut-il peindre ce qu'il y a sur un visage ? Ce qu'il y a dans un visage ? Ou ce qui se cache derrière un visage ?

Pablo Picasso

Plusieurs caractéristiques sont utilisées pour la description d'un visage : sa forme et celle de ses composantes (forme des yeux, de la bouche, de la mâchoire), la texture du visage (couleur des yeux, de la peau, présence de fossettes, de rides).

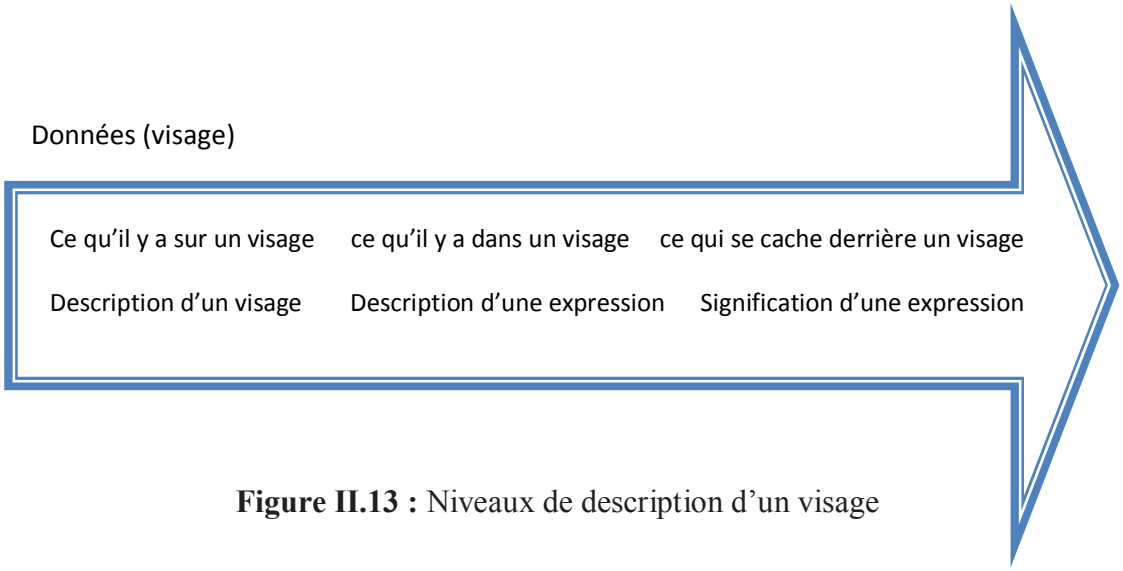


Figure II.13 : Niveaux de description d'un visage

Comme l'indique la citation de Picasso, lorsqu'on voit un visage, plusieurs niveaux de description sont possibles, allant du plus prés des données au plus abstrait. Cette distinction en trois niveaux peut être aussi applicable à la vision par ordinateur.

Nous pouvons tout d'abord décrire ce qu'il y a sur un visage, de façon à représenter ses traits, ses couleurs et ses aspérités. Tous les aspects du visage sont pris en compte, aussi bien les caractéristiques liées à l'identité de la personne que celles représentant son expression. C'est le premier niveau de description, que nous appellerons description d'un visage.

Nous pouvons ensuite décrire ce qu'il y a dans ce visage. Il s'agit alors d'extraire certaines caractéristiques. Evidement on s'intéressent aux caractéristiques de l'expression de la personne (sourire, haussement de sourcils, ...). C'est le second niveau de description, que nous appellerons description d'une expression.

Pour finir on peut décrire ce qui se cache derrière ce visage, de façon à interpréter et à donner une signification au visage expressif. Cela peut permettre d'indiquer l'émotion affichée par la personne (par exemple: elle est joyeuse). C'est le troisième niveau de description, que nous appellerons signification d'une expression. **[30]**

II.2.2. Psychophysiologie du visage :

Avant de s'intéresser à l'analyse automatique des expressions du visage, on doit connaître les particularités du visage : son anatomie et le lien qui existe entre le visage et les émotions.

II.2.2.a. Physiologie du visage :

Le visage est une zone importante du corps humain qui possède différents muscles. D'après une technique appelée L'électromyographie (EMG) qui mesure l'activité musculaire au cours du temps, l'activation de ses muscles, peut généralement être découpée en trois phases :

– la phase d'attaque (attack) ou (onset), qui correspond à la période pendant laquelle l'activité du muscle passe de la valeur nulle à sa valeur maximale,

– la phase de soutien (sustain) ou (apex), qui correspond à la période pendant laquelle l'activité stagne à son maximum,

– la phase de relâchement (relaxation) ou (offset), qui correspond à la période pendant laquelle l'activité du muscle baisse jusqu'au niveau initial.

Pour expliqué L'électromyographie (EMG) est une technique permettant de mesurer l'activité musculaire au cours du temps. Cette technique a permis de déduire que l'activation musculaire, et en particulier

II.2.2.b. Psychologie du visage

Les expressions faciales peuvent former une indication sur l'état émotionnel d'un individu : ce sont les expressions dites spontanées en contraste avec les expressions qui peuvent être forcées. Les deux types d'expression sont générées par deux zones distinctes du cerveau ,par exemple l'activation des muscles entourant les yeux accompagne en général un sourire. [31]

II.2.3. Extraction de la Forme du Visage :

La forme du visage décrit l'emplacement des principales composantes du visage que sont le front, les sourcils, les yeux, le nez, la mâchoire et la bouche. Cette forme peut être donnée par la localisation de points caractéristiques, par des maillages ou encore par des formes géométriques simples. Des procédures mathématiques permettent d'extraire les principales déformations et de réduire la dimensionnalité de ces vecteurs. Pour ce faire, plusieurs visages sont annotés et les procédures sont appliquées sur l'ensemble des vecteurs associés. Parmi ces méthodes, nous pouvons citer l'analyse en composantes principales (ACP) qui permet de transformer des variables corrélées en nouvelles variables décarrelées les unes des autres. Ces nouvelles variables sont nommées composantes principales. Les axes correspondants sont orthogonaux et correspondent aux principales déformations allant de la plus grande à la plus petite. Cette technique à souvent été utilisée pour décrire des visages.

II.2.4. Extraction de la Texture du Visage :

La texture peut tout d'abord être décrite par le niveau de gris de l'ensemble des pixels de l'image contenant le visage. Afin de limiter l'information du visage l'image est préalablement recentrée sur le visage et rognée de façon à ce que les bords du visage touchent (ou presque) les bords de l'image. Parfois, le contour du visage est utilisé et seules les informations de texture à l'intérieur de cette forme sont prises en compte. **[32]**

II.2.5. Analyse du visage :

L'analyse de visage devenue un enjeu important de la vision. Modéliser la vision a toujours été un défi en informatique et en intelligence artificielle. Il apparaît qu'étudier le cerveau et la vision

humaine peut faire progresser en vision par ordinateur. N'importe qu'elle étape dans le processus de l'analyse de visages doit être précéder par l'étape de détection de visages.

Tout processus automatique de détection de visages doit prendre en compte plusieurs facteurs qui contribuent à la complexité de sa tâche, car le visage est une entité dynamique qui change constamment sous l'influence de plusieurs facteurs, et par la suit la démarche générale adoptée pour réaliser un tel système :

- ✓ **Monde physique :** Visages Humains.
- ✓ **Codage**: Acquisition d'images (Caméra,…etc), il consiste aussi à la digitalisation de l'image, et il comporte un risque de bruit.
- ✓ **Prétraitement**: Les étapes de prétraitement consistent à sélectionner l'information nécessaire à l'application. Cette sélection passe souvent par l'élimination du bruit dû aux conditions d'acquisition, par la normalisation des données, aussi bien que par la suppression de la redondance.
- ✓ **Analyse** : (appelée aussi indexation, représentation, modélisation ou extraction de caractéristiques), il faut extraire de l'image les informations qui seront sauvegardées en mémoire pour être utilisées plus tard dans la phase Décision.
- ✓ **Apprentissage** (si le système est basé sur l'apprentissage): consiste à mémoriser les représentations calculées dans la phase Analyse.
- ✓ **Décision** : Mesure de similarité. [33]

II.3. LES EMOTIONS

II.3.1. Définition :

Les émotions, représentes des états motivationnels. Elles sont constituées d'impulsions, de désirs ou, plus généralement, elles comportent des changements de motivation. Elles poussent l'individu à modifier sa relation avec un objet, un état du monde, un état de soi, ou à maintenir une relation existante malgré des obstacles ou des interférences .

Une émotion est une réaction psychologique et physique à une situation. Elle a d'abord une manifestation interne et génère une réaction extérieure. Elle est provoquée par la confrontation à une situation et à l'interprétation de la réalité. [34]

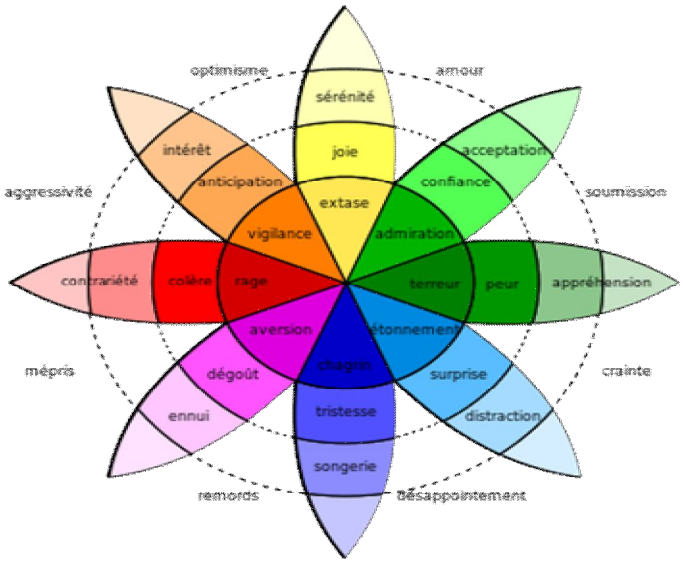


Figure II.14 : Roue des émotions de **Robert Plutchik**

Robert Plutchik a créé une roue des émotions pour illustrer diverses émotions motivatrices et nuancées. Il a proposé son modèle circomplexe en trois dimensions et son modèle de roue en deux dimensions en 1980 pour décrire les relations entre ces émotions.

II.3.2. Types d'émotion :

Les émotions sont des séquences courantes et vives qui interviennent tout au long de nos journées et en donnent le ton. Il existe trois types d'émotions : émotions primaires ou dites de base, émotions secondaires et émotions sociales.

II.3.2.a. Émotions primaires

Les émotions primaires sont déclenchées par des événements particuliers ou bien elles se manifestent dans des circonstances précises en provoquant des comportements spécifiques. Elles sont à la base de nos réactions qui ne sont pas seulement déterminées par notre jugement rationnel ou notre passé individuel mais aussi par notre passé ancestral (ancien).[35]

II.3.2.b. Émotions secondaires

Les émotions secondaires ont pour base, au départ, un processus de pensée et sont l'aboutissement de l'apprentissage des émotions primaires.

Les émotions secondaires sont celles qui sont engendrées à l'évocation de souvenirs et arrivent à maturation à l'âge adulte.[36]

II.3.2.c. Émotions sociales

On parle aussi des émotions sociales qui sont inhérentes à la relation aux autres, comme la culpabilité, la honte, la jalousie, la timidité, l'humiliation, etc. Toutes ces émotions sont apprises

et sont constituées à partir des émotions primaires. L'éducation et la culture sont fortement impliquées dans l'acquisition des émotions sociales . [37]

Emotion	Déclencheurs et circonstances d'apparition	Comportement
Joie	Désir Réussite Bien être Accomplissement	Rapproche
Tristesse	Perte Deuil	Replis sur soi
Colère	Obstacle Injustice Dommage Atteindre à son intégrité physique ou psychique Atteindre au système de valeur	Attache
Peur	Menace Danger Inconnu	Fuite Sidération Evitement Parfois attaque
Dégout	Substance ou personne nuisible Aversion physique ou psychique Contre quelqu'un Rejet	
Surprise	Danger immédiat Inconnu Imprévu	Retrait Sursaut

Tableau II.1 – Les 6 principaux émotions [38]

II.3.3. Composantes d'une émotion :

D'après la définition de l'émotion, on peut noter la présence de composantes fondamentales: la composante comportementale, la composante physiologique et la composante cognitive/subjective.

II.3.3.a. Composantes physiologiques des émotions

Le comportement et l'état de l'organisme sont acétés par des réactions physiologiques périphériques qui accompagnent toute émotion. En effet, l'activité neurophysiologique établit une fonction adaptative attribuée à l'émotion. Il s'agit d'une activité sympathique qui permet à l'individu de réagir rapidement aux stimuli externes. Le système nerveux autonome (SNA) subit des modifications associées aux états émotionnels en commandant à nos viscères indépendamment de notre volonté.

II.3.3.b. Composantes comportementales des émotions

L'émotion peut être révélée par un ensemble de traits comportementaux par lesquels elle se révèle, tel que l'intonation de la voix, les pleurs, le sourire ou les mimiques faciales. La

fonction principale de l'expression émotionnelle est de générer un langage détectable par les autres individus. [39]

II.4. EXPRESSION FACIALE

L'expression faciale est un aspect important du comportement et de la communication humaine où le changement dans le visage, perceptible visuellement, dû à l'activation (volontaire ou non) de l'un ou de plusieurs des 44 muscles composant le visage.

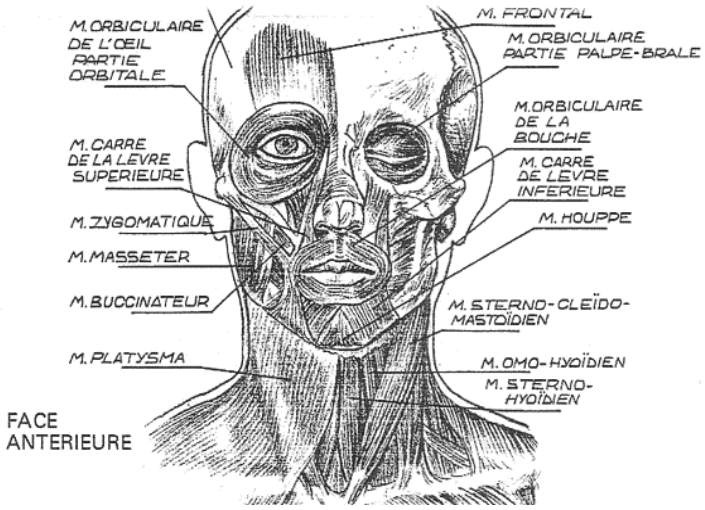


Figure II.15 : Principaux muscles faciaux (M représente le t muscle) [40]

II.4.1. Historique à propos des expressions faciales :

Les travaux fondamentaux dans le domaine de l'analyse des expressions du visage sont dus principalement à Charles Darwin, Guillaume Duchenne De Boulogne au XIXème siècle et plus récemment Paul Ekman.

Au XIXème siècle, **Guillaume Duchenne de Boulogne** est le premier à localiser individuellement les différents muscles faciaux par activation électrique. Il est un des premiers à livrer à la communauté scientifique un ensemble de photographies montrant l'activation des différents muscles faciaux.

Charles Darwin , est le premier à traiter de l'universalité des expressions du visage et à proposer une théorie évolutionniste sur la formation des expressions. L'argument principal est que les expressions des enfants et des nouveaux-nés existent aussi chez les adultes. D'après lui, l'expression des émotions est un processus nécessaire à la survie. Ainsi, les expressions non-verbales sont aussi importante que les interactions verbales dans le processus de communication humain.

Paul Ekman , psychologue, s'intéresse à partir du milieu des années 1960 aux expressions et émotions humaines. Il met en évidence l'universalité de certaines émotions et développe un outil

de codification des expressions du visage largement utilisé aujourd'hui. Il s'intéresse désormais à l'analyse des expressions de manière informatique.

Les expressions faciales jouent un rôle important dans la conversation humaine. Donc,l'interface homme machine du futur utilisera non seulement la souris, le clavier, mais aussi de nouvelles sources d'informations telles que les six expressions faciales (colère, dégoût, joie, peur, surprise, tristesse), la posture humaine ou parole [41].

II.4.2. Qu'est-ce qu'une expression faciale?

Les expressions du visage sont principalement générées par la contraction des muscles qui induisent des modifications temporaires des caractéristiques de forme du visage telles que le clignement des paupières, le haussement des sourcils, la forme de la bouche ou encore des modifications de la texture de la peau telle que l'apparition de rides ou de fossettes.



Figure II.16 : Mouvements faciaux globaux.

II.4.3. Pourquoi les expressions faciales existent-elles?

Trois points de vue différents :

- Les émotions sont au centre de l'explication des expressions faciales Sélection naturelle

  ⇒ émotions de base ⇒ expressions faciales universelles (colère, dégoût, joie, peur, surprise, tristesse), les étiquettes pourraient être remises en question car elles dépendent du langage et de la culture.
- Les expressions faciales sont des signaux de conversation qui dépendent des intentions de l'émetteur, du comportement du récepteur et du contexte de l'interaction, non des émotions.

- Les expressions faciales sont des activateurs et des régulateurs d'émotion dû à la plasticité du cerveau, changent la température de l'hypothalamus pour faciliter ou inhiber la production de neurotransmetteurs reliés aux émotions.

Une émotion peut causer, une expression faciale ou non et une expression faciale peut être causée par une émotion ou non [42].

II.4.4. Les principaux expressions faciales universelles :

Suivant la théorie de paul Ekman qui affirme que les expressions du visage ne sont pas déterminées par la culture, mais qu'elles sont universelles , présentées biologiquement par une liste de 6 émotions de « base » qui sont:

- Joie
- Tristesse
- Colère
- Peur
- Dégoût
- Surprise

II.4.4.1. La joie (joy):



Figure II.17 : Micro-expression de joie

C'est peut-être banal, mais lorsqu'une personne ressent de la joie, elle sourit. L'inverse, c'est à dire essayer de faire croire à quelqu'un qu'on est joyeux, est beaucoup plus difficile

II.4.4.2. La tristesse (sad)



Figure II.18 : Micro-expression de tristesse

Il est parfois important de savoir repérer la tristesse derrière un sourire. Que ce soit pour avoir une oreille attentive ou bien pour repérer la détresse dans laquelle se trouve un proche.

Les indices : les paupières supérieures et les coins externes des sourcils s'affaissent. La personne donnera l'impression d'avoir le regard dans le vide. Egalement, les coins des lèvres ont tendance à s'affaisser.

II.4.4.3. La colère (angry)



Figure II.19 : Micro-expression de colère

La colère est en général facile a répérer. Mais parfois elle peut survenir d'un coup en pleine conversation banale.

Les Indices : les lèvres sont pressées l'une contre l'autre. Les sourcils se pressant vers le nez. Les rides entre deux sourcils sont apparentes et la plus grande caractéristique restant le regard fixe.

II.4.4.4. Le mépris (contempt)



Figure II.20 : Micro-expression de mépris

Les indices : un côté de la lèvre est légèrement tiré vers l'arrière (ou remonté). Ce mouvement de la lèvre peut être très subtile, apparaître et disparaitre très rapidement.

II.4.4.5. Le dégoût (disgust)

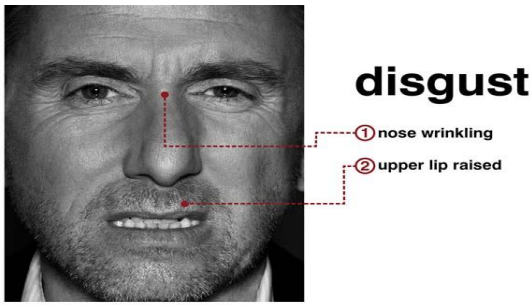


Figure II.21 : Micro-expression de dégoût

Les indices : la lèvre supérieure remonte laissant généralement apparaître les dents. Cet indice s'accompagne généralement d'un pli au niveau du nez.

II.4.4.6. La Peur (afraid)

C'est une expression que l'on connaît tous. Si vous voulez l'observer je vous invite tout simplement à surprendre l'un de vos amis. Sur son visage se dessineront pendant quelques millisecondes les expressions suivantes.

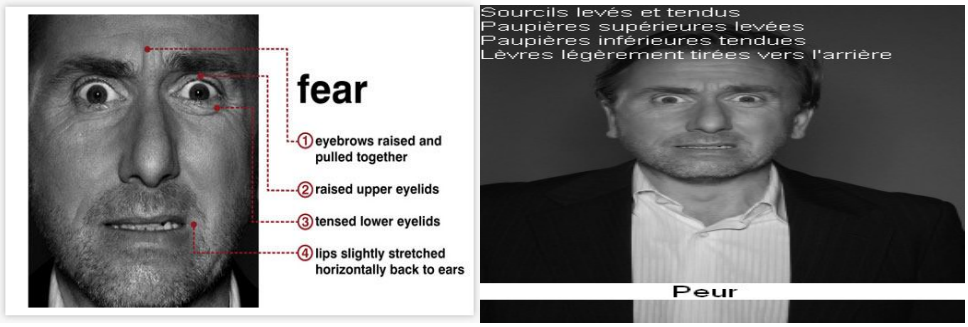


Figure II.22 : Micro-expression de peur

Les indices : Les lèvres s'étirent à l'horizontal vers les oreilles. Les paupières inférieures se tendent et les paupières supérieures sont remontées. Les sourcils remontent également .

II.4.4.7. La surprise (surprise)



Figure II.23 : Micro-expression de surprise

Les indices : la bouche est entrouverte, les yeux bien ouverts et les sourcils relevés. [43]

II.4.5. Formalismes de description des expressions faciales

La description des expressions du visage est un ensemble d'interprétations successives d'indices visuels. Un sens (dépendant du domaine d'application) est construit à partir de mesures de bas niveaux (présence ou non d'une composante, position éventuelle).

- **Attribut facial** : un attribut facial est une propriété élémentaire caractérisant un visage. La position des yeux est un attribut facial. La présence de barbe est un autre attribut. Les attributs faciaux directement visibles sont dit de premier ordre. Les attributs qui ne peuvent être mesurés qu'à partir d'autres attributs de premier ordre, sont des attributs de second ordre et ainsi de suite.
- **Indice visuel** : un indice visuel est une propriété élémentaire centrée observateur du visage : c'est un attribut facial qui est observé et visible. Certains attributs ne sont pas visibles chez certaines personnes (barbe, moustache, sourcils) ; certains ne sont visibles qu'à certains moments (un oeil peut être caché lors d'une rotation de la tête par exemple).[44]
- **Action faciale** : une action faciale est un ensemble d'indices visuels intégrés de manière temporelle. Le relèvement des sourcils est par exemple une action faciale composée d'un ensemble de positions successives des sourcils. le relèvement des sourcils par exemple consiste en une position actuelle des sourcils plus haute que sa position précédente. [45]

II.4.6. Differents Systemes et Techniques de Codage des Expressions du Visage

II.4.6.1. FACS(Facial Action Coding System)

En 1978, Ekman présente un système de codification manuelle des expressions du visage . Leurs travaux d'observation leur permettent de décomposer tous les mouvements *visibles* du visage en terme de 46 (AU) *Actions Unitaires* (*Action Unit* , qui correspond aux *actions faciales* définies plus haut) qui décrivent les mouvements élementaires des muscles. N'importe quelle mimique observée peut donc être représentée sous le forme d'une combinaison d'*Actions Unitaires* .

II.4.6.2. FACS+

Bien que FACS soit un système de description bénéficiant d'une grande maturité (environ vingt années de développement), il souffre cependant de quelques inconvénients

- ✓ **Complexité** : on estime qu'il faut 100 heures d'apprentissage pour en maîtriser les principaux concepts
- ✓ **Difficulté de manipulation par une machine** : FACS a d'abord été créé pour des psychologues, Certaines mesures restent floues et difficilement évaluables par une machine.
- ✓ **Manque de précision** : les transitions entre deux états d'un muscle sont représentées de manière linéaire, ce qui est une approximation de la réalité.

II.4.6.3. MPEG4

La norme de codage vidéo MPEG-4 dispose d'un modèle du visage humain développé par le groupe d'intérêt Face and Body AdHoc Group. C'est un modèle 3D articulé.

Ce modèle est construit sur un ensemble d'attributs faciaux, appelés **Facial Feature Points** (FFP). Des mesures sur ces FFP sont effectuées pour former des unités de mesure (Facial Animation Parameter Units) qui servent à la description des mouvements musculaires (Facial Animation Parameters - équivalents des Actions Units d'Ekman).

II.4.6.4. Candide

Candide est un modèle du visage. Il est composé d'un modèle en fil de fer représentant un visage générique et d'un ensemble de paramètres :

- **Paramètres de forme** (*Shape Units*) : ces paramètres permettent d'adapter le modèle générique à un individu particulier. Ils représentent les différences inter-individus et sont au nombre de 12 :

1. Hauteur de la tête,
2. Position verticale des sourcils,
3. Position verticale des yeux,
4. Largeur des yeux,
5. Hauteur des yeux,
6. Distance de séparation des yeux,
7. Profondeur des joues,
8. Profondeur du nez,
9. Position verticale du nez,
10. Degré de courbure du nez (s'il pointe vers le haut ou non),
11. Position verticale de la bouche,
12. Largeur de la bouche. [46].

II.4.7. Historique les techniques récentes de REF :

-Hong-Viet [47] on travaillé en **2006** sur la **reconnaissance multimodale de gestes de communication non verbale**. Le contexte général est de présenter leurs contributions en reconnaissance automatique des expressions faciales à partir d'information vidéo. Premièrement, on doit extraire les caractéristiques de l'image du visage, puis on va les utiliser comme entrée du système de classification. Toutes les méthodes nécessitent une intervention manuelle pour la détection et la normalisation. Et toutes les méthodes vont mettre en correspondance les expressions faciales avec six expressions basiques.

-Ligang et al [48] ont étudiés en **2011** la **Géométrie vs Apparence pour discrimination posée par rapport version spontanée d'une émotion**. Les émotions spontanées des expressions

faciales diffèrent de celles qui se posent dans l'apparence (séquences et mouvements de tête). Les images fixes ne peuvent pas fournir des informations de mouvement directement. Cependant, indirectement, les distances entre les points clés sur un visage extraites d'une image fixe à l'aide de modèles à forme active peuvent capter une partie en mouvement et changements de pose. - **Rizwan et al** [49] ont aussi travaillé en **2012** sur **une méthode de reconnaissance des expressions du visage basée sur la perception.**.trois types d'expressions faciales se présente en conciderant que de petites régions du visage bien définies pour en extraire les caractéristiques. Les résultats de cette étude montrent que pour certaines expressions du visage une unique région est perceptuellement plus attractive que les autres.

II.4.8. Tableau comparatif des travaux récents sur la REF :

Source	Année	Base	Méthode	Taux de reconnaissance
Zilu Ying Beihang university, Beijing	2009	JAFFE	LBP et LE	70.48%
Ligang Zhang. Queensland University of Technology. Australia	2011	Le naturel d'expression visible et infrarouge du visage NVIE	FAP SIFT+ mRMS	76.1% **83.4 %**
AnishaHalder. Jadavpur University	2011		GT2FS	96.67%
Amit Konar1. Jadavpur University	2012		IT2FS	88.66%
Rizwan Ahmed Khan. Université de Lyon	2012	**CK (**Cohn-Kanade)	SVM C4.5 Decision Tree **RF**	97.3% 97.6% **99.5%**

Tableau II.2 : Comparaison des performances des algorithmes proposés sur les bases de visages. [50]

- Japanese Female Facial Expression(**JAFFE).**
- facial animation parameters(**FAP**).
- Scale-invariant feature transform(**SIFT**).
- minimal redundancy maximal relevance(**mRMS**).
- General Type-2 Fuzzy sets(**GT2FS**).
- Interval Type-2 Fuzzy Face-Space(**IT2FS**).
- Random Forest(**RF**)

II.5. Détection du visage par Viola Jones

La méthode de Viola et Jones est une méthode de détection de visages dans une image numérique, proposée par les chercheurs Paul Viola et Michael Jones. Elle fait partie des toutes premières méthodes capables de détecter efficacement et en temps réel des visages dans une

image. La méthode de Viola et Jones est une approche basée sur l'apparence, qui consiste à parcourir l'ensemble de l'image en calculant un certain nombre de caractéristiques dans des zones rectangulaires qui se chevauchent. Elle a la particularité d'utiliser des caractéristiques très simples mais très nombreuses. [51]

II.6. Extraction des traits faciaux par EigenFace

La plupart des approches de reconnaissance informatique des visages et de leurs expressions a été axée sur la détection des caractéristiques faciales individuelles telles que les yeux, le contour de la tête et la bouche, ou sur la définition d'un modèle de visage en position et en taille avec les relations entre ces caractéristiques.L'extraction des caractéristiques joue un rôle essentiel dans la phase de prétraitement. [52].



Figure II.24 : *Eigen faces*

II.7. Système de reconnaissance d'expression faciale existant :

II.7.1. Méthode basée sur la segmentation :

Ce système se présente sous forme de quatre phases essentielles : la segmentation, extraction de données, analyse et classification.

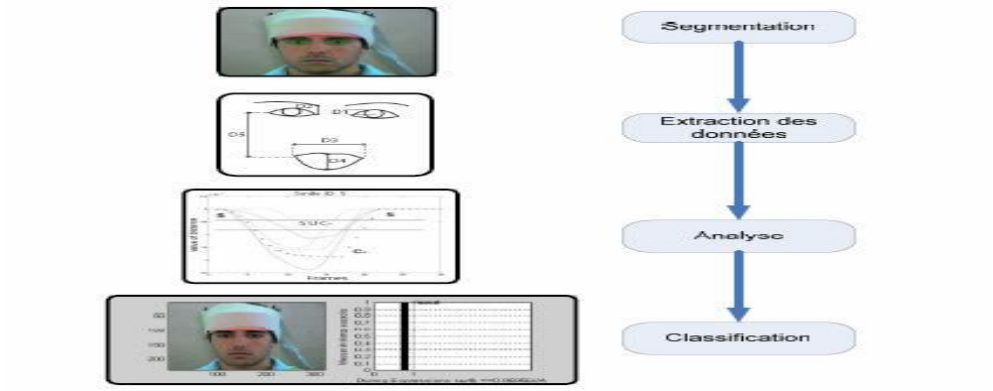


Figure II.25 : Les étapes dans le processus de reconnaissance des expressions faciales.

Dans l'étape de segmentation, le système va localiser les régions des yeux, des sourcils, et de la bouche.En fait, ce sont les contours des yeux, de la bouche, des sourcils. A partir des squelettes, on va choisir cinq distances pour caractériser les déformations.

- ***Segmentation des traits du visage***

Dans cette partie, on décrit comment on peut extraire le squelette d'émotion à partir d'une image. Une hypothèse indispensable est que l'utilisateur doit faire face à la caméra.Après l'utilisation de l'algorithme de segmentation de contours de l'œil, des sourcils et de la bouche, les contours de caractéristiques faciales sont extraits automatiquement. [53]

- ***Extraction de données caractéristiques***

Les caractéristiques considérées contiennent les contours des yeux, des sourcils et des lèvres. Finalement, on obtient les squelettes d'expressions.

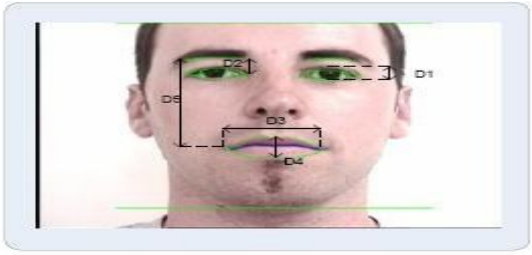


Figure II.26 : Extraction des contours et définition des 5 distances.

À partir du squelette d'expression, on peut déterminer les déformations de caractéristiques faciales. Cinq distances sont définies pour chaque squelette.[54]

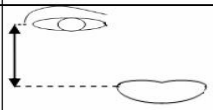
D1		Distance entre la paupière supériure et inférieure
D2		Distance entre le coin intérieur de l'oil et le celui du sourcil
D3		Largeur de bouche
D4		Ouverture de bouche en hauteur
D5		Distance entre le coin de la bouche et celui du coin extérieur de l'oeil

Tableau II.3 : Tableau de définition de distances.

*II.7.2. M*éthode par masque pour le visage

L'idée c'est d'utiliser un masque pour l'enregistrement dans une reconnaissance de visage. Ce système n'est pas nouveau. Certains masques existants, sont bien conçus pour le montage surface déformable selon l'une des actions musculaires. Toutefois, lorsqu'on travaille avec des images face planes, il est difficile d'atteindre d'un bon enregistrement pour tous les sommets densément placés. En outre, bien que ces masques sont à morphologie souple. Par conséquent, on utilise une procédure simplifiée et à masque triangulaire. Sur ce masque, les triangles gris correspondent à des régions des sourcils, yeux, nez et bouche. [55]

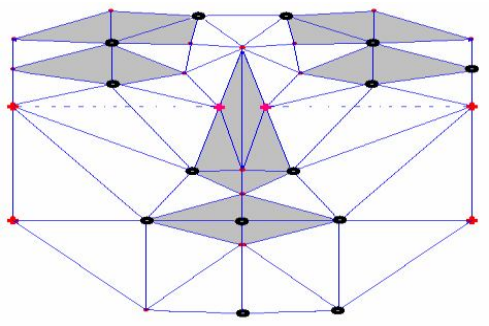


Figure II.27 : Masque pour le visage.

II.8. Synthèse des méthodes de reconnaissance des visages:

Les points forts et les limites de ces méthodes peuvent être résumés dans le tableau ci-dessous:

APPROCHES	AVANTAGES	LIMITES
Approches Géométriques	- Requiert moins d'espace mémoire - Plus rapide et simple a implémenter	-Dépendent étroitement de l'exactitudes d'extraction des caractéristiques -Prise en charge de la variation de la pose, de l'éclairage
Les Visages Propres	-Une représentation discriminante du visage -Réduction de la taille du problème	-Cout de calcul élevé -Détermination du nombre de visages propres (base de projection)
Les Templates Matching	-Robuste pour localiser les régions	-Complexité de calcul -Le nombre de templates -Espace mémoire important
Les Réseaux de Neurones	-Temps de classification -Résistant aux bruits et au manqué de fiabilité des données	-Sur-apprentissage -Extension a des problèmes plus larges -Apprentissage long
Les HMMs	-Algorithmes d'apprentissage très robuste	-Temps d'apprentissage et classification long -Choix des états se base sur l'intuition subjective
Les Machines à vecteurs de Support	-Capacité d'extraction des informations discriminantes et pertinentes du visage d'une manière entièrement automatique	-Temps d'apprentissage long et algorithme d'apprentissage compliqué

Tableau II.4 : Synthèse des méthodes des reconnaissances des visages[56]

III. La Classification

III.1. Définition

L'importance de la classification est connue dans divers domaines de la vie humaine. Elle permet d'avoir une structure compréhensible des objets pour les localiser dans la situation réelle. Elle sépare la zone de données en différentes catégories en fonction de leurs caractéristiques et de leurs qualités, de sorte que, chaque catégorie comprenne un ensemble d'objets similaires contre les autres.

III.2. Classification supervisée et non supervisée

Les méthodes de classification sont divisées en deux groupes selon l'information disponible sur les spécifications de classes:

- Classification non supervisée,
- Classification supervisée.

III.2.1. La classification non supervisée n'utilise pas une connaissance d'expert, et dans certains cas, il y a des informations qui manquent concernant le schéma de classes et de leurs attributs de sortie. Dans cette technique, des procédures statistiques sont utilisées pour estimer une similitude dans la zone de comparaison des données d'entrée.

III.2.2. La classification supervisée des informations suffisantes sur le caractère des classes étiquettes et des classes prédéfinies sont fournies.

La classification des objets ou des modèles en plusieurs groupes est un but principal dans le processus de reconnaissance de formes. Il existe de nombreuses applications dans lesquelles la reconnaissance de modèle est utilisée comme une partie importante comme la reconnaissance de formes dans les systèmes de vision industrielle, les systèmes de reconnaissance de caractères optiques, les systèmes d'aide à la décision pour le diagnostic, les systèmes d'interaction homme-machine, les systèmes de reconnaissance vocale et ainsi de suite. **[57]**

III.3. Les Machines à vecteurs de support SVM

Les SVM sont des algorithmes de classification récemment développés par Vapnick [58]. L'idée de base des SVM est d'utiliser des fonctions noyaux et de projeter les données de l'espace d'entrée (non linéairement séparables) dans un espace de plus grande dimension appelé espace de caractéristiques , de façon à ce que les données deviennent linéairement séparables .Ces algorithmes tentent de minimiser les erreurs d'apprentissage tout en maximisant la marge séparant les données de classes. Le comportement du SVM est conditionné par le type de noyau utilisé dans la projection et les valeurs accordées à ces paramètres. [59]

III.3.1. SVM principe de fonctionnement général

III.3.1.1 Notions de base : Hyperplan, marge et support vecteur

Dans la présentation des principes de fonctionnements, on schématisera les données par des « points » dans un plan. Pour deux classes d'exemples donnés, le but de SVM est de trouver un classificateur qui va séparer les données et maximiser la distance entre ces deux classes. Avec SVM, ce classificateur est un classificateur linéaire appelé hyperplan.

Dans le schéma qui suit, on détermine un hyperplan qui sépare les deux ensembles de points.

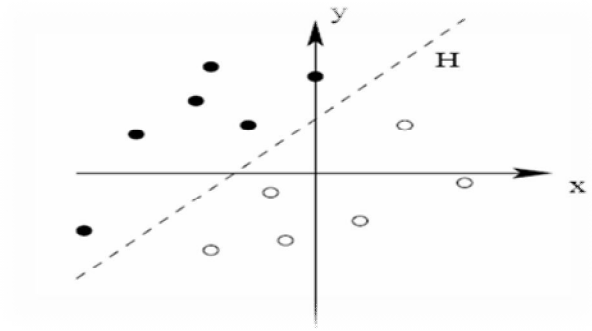


Figure III.28 :Hyperplan qui sépare les deux ensembles de points

III. 3.2. Un SVM multi-classes

L'étude théorique des systèmes d'apprentissage s'est concentrée principalement sur des classifieurs à deux classes. Cette étude ne concerne que des fonctions dont les sorties sont dans {−1, 1}. Par conséquent les principes d'induction de même que les bornes obtenues ne s'appliquent pas, du moins pas directement, à des ensembles de fonctions de $\mathbf{x}$ dans $\{1,2,\ldots,m\}, m > 2$, où m désigne le nombre de classes.

III.3.2.1. Une-contre-reste

L'architecture la plus ancienne, utilisée pour les machines à vecteurs supports multi-classes, est probablement la méthode une-contre-reste. Elle construit m classifieurs binaires à vecteurs supports où m est le nombre total des classes. L'apprentissage du $k^{\text{ème}}$ classifieur à vecteurs

supports s'effectue en considérant tous les exemples de la $k^{\text{ème}}$ classe dans la région positive et

tous les autres exemples dans la région négative. Le $k^{\text{ème}}$ classifieur à vecteurs supports s'obtient en résolvant le problème.

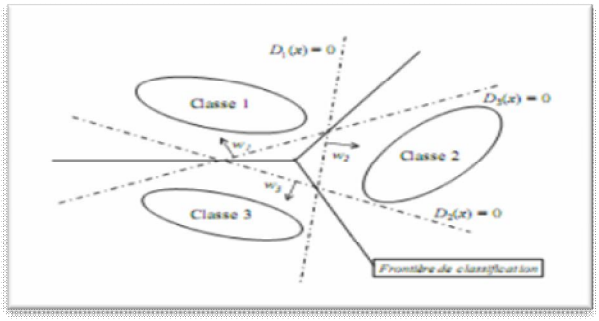


Figure III.29: les nouvelles frontières de classification.

III.3.2.2.Une-contre-une

L'approche une-contre-une consiste à construire un classifieur pour chaque paire de classes (k, s) définissant ainsi des fonctions de décision binaire . [60]

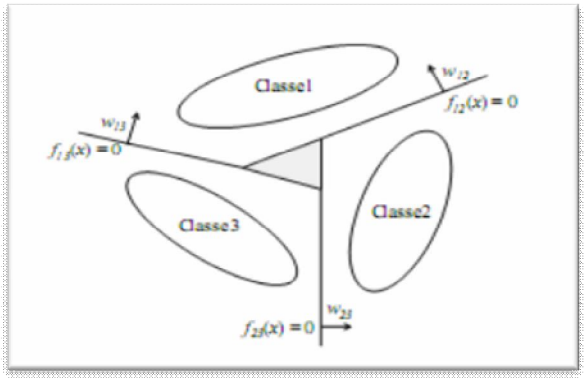


Figure III.30 :La région d'ambiguïté pour l'approche une-contre-une.

III.4. Avantages et inconvénients :

III.4.1. Avantages

- ✓ Les SVM possèdent des fondements mathématiques solides.
- ✓ Les exemples de test sont comparés juste avec les supports vecteur et non pas avec tout les exemples d'apprentissage.
- ✓ Décision rapide. La classification d'un nouvel exemple consiste à voir le signe de la fonction de décision f(x).

III.4.2. Inconvénients

- ✓ Classification binaire d'où la nécessité d'utiliser l'approche un-contre-un.
- ✓ Grande quantité d'exemples en entrées implique un calcul matriciel important.
- ✓ Temps de calcul élevé lors d'une régularisation des paramètres de la fonction noyau. [61]

IV. Conclusion :

Ce chapitre nous a permis de tirer un certain nombre d'enseignements sur des notions générales du visage humain et ses composants, processus de reconnaissance d'expressions faciales et la relation entre expression et émotions est ainsi précisé. Après nous avons présenté explicitement les six expressions universelles reconnus, un état de l'art des techniques proposées jusqu'à ce jour et méthodes et travaux récents à été présenté en faisant une synthèse sur la classification qui est une étape importante pour notre sujet en ce focalisant sur le SVM. Cela va nous aider pour l'élaboration du chapitre suivant.