

2020-2021

---

**FONDEMENTS THÉORIQUES DE  
L'INFORMATIQUE**

Exposé

**LES MATHÉMATIQUES ET LA  
RECONNAISSANCE FACIALE**

Membres du groupe :

- 1- BABA Tchao Arsèkè
- 2- FIATSI Etsè
- 3- KOUDRAMASSAN Fo-Kossi Norbert

Enseignant :

Dr EDARH BOSSOU

## Table des matières

Introduction .....	4
1- La détection du visage.....	4
1.1- Objectifs [1] .....	4
1.2- Le balayage [2] .....	4
1.3- Image intégrale [3] .....	6
1.4- Sélection par boosting [2].....	7
2. La numérisation de visage .....	8
2.1- Objectif .....	8
2.2- Méthodologie [4] [5] .....	8
3- La comparaison.....	9
Conclusion.....	10
Bibliographie .....	11

## Liste des figures

Figure 1: Quelques fonctionnalités de type Haar [2] .....	5
Figure 2: Détection de certaines parties du visage [4] .....	5
Figure 3: calcul des images intégrales .....	6
Figure 4: Exemple d'un rectangle .....	7
Figure 5: Fonctionnement d'un classifieur fort [2] .....	7
Figure 6: Mise en cascade de plusieurs classificateurs faibles .....	7
Figure 7: Résumé de l'algorithme de Viola & Jones .....	8
Figure 8: Transformation en vecteur des images de références.....	9
Figure 9: Calcul du visage moyen .....	9
Figure 10: Calcul des informations propres .....	9

## Introduction

La reconnaissance faciale est un domaine qui a pris de l'ampleur ces dernières années grâce à de nouveaux algorithmes et à la puissance de calcul grandissante des machines. Ce document expose certaines des bases mathématiques sur lesquelles se base ce domaine.

Reconnaître un visage se fait en trois étapes :

- La détection d'un visage sur un flux multimédia ;
- Sa numérisation ;
- Et finalement sa comparaison avec d'autres images.

## 1- La détection du visage

Détecter un visage est la première étape du processus de reconnaissance faciale. En effet, les conditions dans lesquelles le programme devra être utilisé diffèrent grandement: reconnaître une personne dans une foule, reconnaître un visage sur une photo simple ou dans une vidéo... Il est donc nécessaire de pouvoir détecter sur le flux multimédia le/les images avant de lancer la méthode de reconnaissance faciale à proprement parlé.

L'une des méthodes principalement utilisées est celle de Viola & Jones que nous exposerons ici. Elle a été proposée en 2001 et a été améliorée depuis.

### 1.1- Objectifs [1]

La méthode de Viola & Jones consiste à balayer une image à l'aide d'une fenêtre de détection de taille initiale 24px par 24px (dans l'algorithme original) et de déterminer si un visage y est présent. Lorsque l'image a été parcourue entièrement, la taille de la fenêtre est augmentée et le balayage recommence, jusqu'à ce que la fenêtre fasse la taille de l'image. L'augmentation de la taille de la fenêtre se fait par un facteur multiplicatif de 1,25.

Le balayage, quant à lui, consiste simplement à décaler la fenêtre d'un pixel. Ce décalage peut être changé afin d'accélérer le processus, mais un décalage d'un pixel assure une précision maximale.

Cette méthode 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.

### 1.2- Le balayage [2]

Pour s'assurer que la partie de l'image présente dans la fenêtre de détection contient véritablement un visage, cette méthode utilise « Les fonctionnalités ou caractéristiques de type Haar »

Les fonctionnalités Haar sont essentiellement des collections de pixels de formes rectangulaires. La raison d'être des caractéristiques Haar est que si l'une d'entre elle est appliquée à une zone d'image, la soustraction entre la valeur des pixels de la région sombre et celle de la région claire donnera une certaine valeur « delta ».

L'illustration suivante montre un ensemble de caractéristiques.

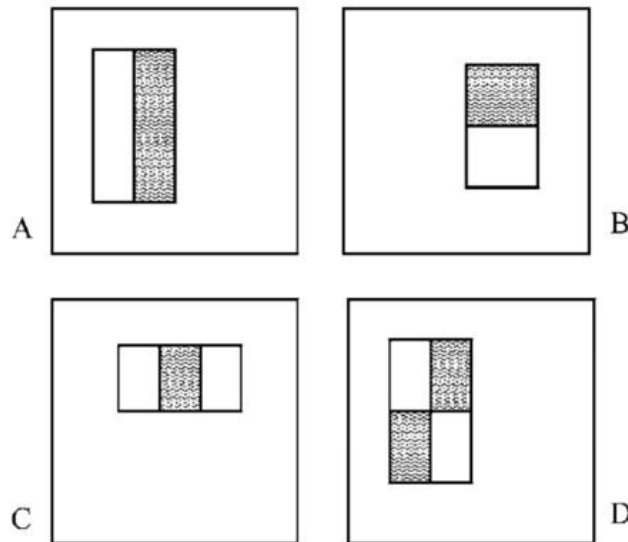


Figure 1: Quelques fonctionnalités de type Haar [2]

Ainsi, si par exemple nous avons envie de détecter une partie d'un visage - en l'occurrence un sourcil -, la teinte des pixels sur un sourcil dans une image sera plus sombre, et brusquement plus clair (à cause de la peau).

De même, si nous voulons détecter une bouche, la forme de la région des lèvres passe de la lumière à l'obscurité puis à la lumière. On pourrait donc obtenir la configuration suivante :

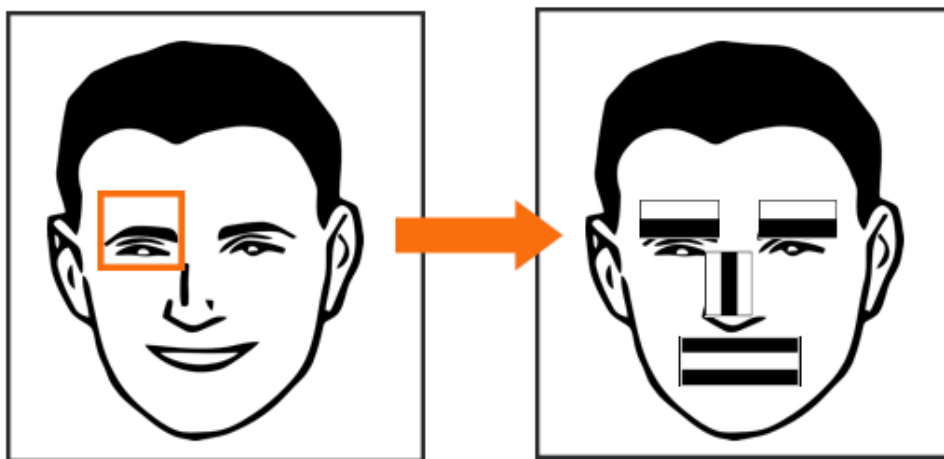


Figure 2: Détection de certaines parties du visage [6]

Placées bout à bout, ces différentes valeurs delta nous donnent les valeurs approximatives d'un visage. Puis, il faut maintenant créer un « classifieur » (un fichier XML) et l'entraîner à reconnaître les visages par deux situations: La première où une énorme quantité de cas positifs (de visages) lui sont présentés et la deuxième où, à l'inverse, une énorme quantité de cas négatifs (absence de visages) lui sont présentés. Concrètement, une banque d'images contenant des visages de personnes est passée en revue afin d'entraîner le classifieur. Ensuite, une banque d'images ne contenant pas de visages humains est passée.

Malheureusement, cette unique méthode est difficilement implémentable à cause des limites suivantes :

- Dans des scénarios réels, les images ne sont pas seulement des collections de pixels en noir et blanc. Les images sont généralement colorées (RVB) ou en niveaux de gris.
- La somme des valeurs de pixels pour tous les types d'entités dans toutes les images de l'ensemble de données peut être très coûteuse en calcul, en particulier en fonction de la résolution des images.
- Il existe plus de 160000 combinaisons de fonctionnalités possibles pouvant s'intégrer dans une image de 24 x 24 pixels et plus de 250 000 pour une image de 28 x 28

C'est pour contourner ces problèmes qu'intervient l'algorithme de Viola & Jones

### 1.3- Image intégrale [3]

C'est une représentation sous la forme d'une image, de même taille que l'image d'origine, qui en chacun de ses points contient la somme des pixels situés au-dessus et à gauche de ce point. Plus formellement, en prenant  $x$  (respectivement  $y$ ) comme les lignes (respectivement les colonnes), l'image intégrale est définie à partir de l'image  $i$  par :

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y')$$

*Figure 3: calcul des images intégrales*

Prenons un exemple. Nous souhaitons calculer la somme des pixels de la zone rectangulaire ABCD suivante. Grâce à l'image intégrale, nous connaissons la valeur de la somme des pixels en chacun des quatre points. Il suffit donc de faire :  $A-B-C+D$ . En seulement trois opérations nous avons réussi à calculer notre somme de pixels. Ainsi, on est en mesure de trouver la somme de pixels de n'importe quelle zone rectangulaire de l'image en seulement 3 opérations et 4 accès à l'image intégrale.

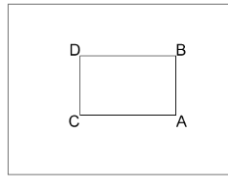


Figure 4: Exemple d'un rectangle

Une caractéristique pseudo-Haar à deux rectangles peut alors être déterminée en seulement 6 calculs (2 points sont partagés) à l'image, et une caractéristique à 3 rectangles en seulement 8.

#### 1.4- Sélection par boosting [2]

La sélection par boosting consiste à utiliser plusieurs classifieurs "faibles" (n'ayant que peu de critères : les valeurs delta) mis en cascade plutôt que d'utiliser un seul classifieur "fort". En effet, avec un seul classifieur dit "fort" qui se présenterait de la sorte :

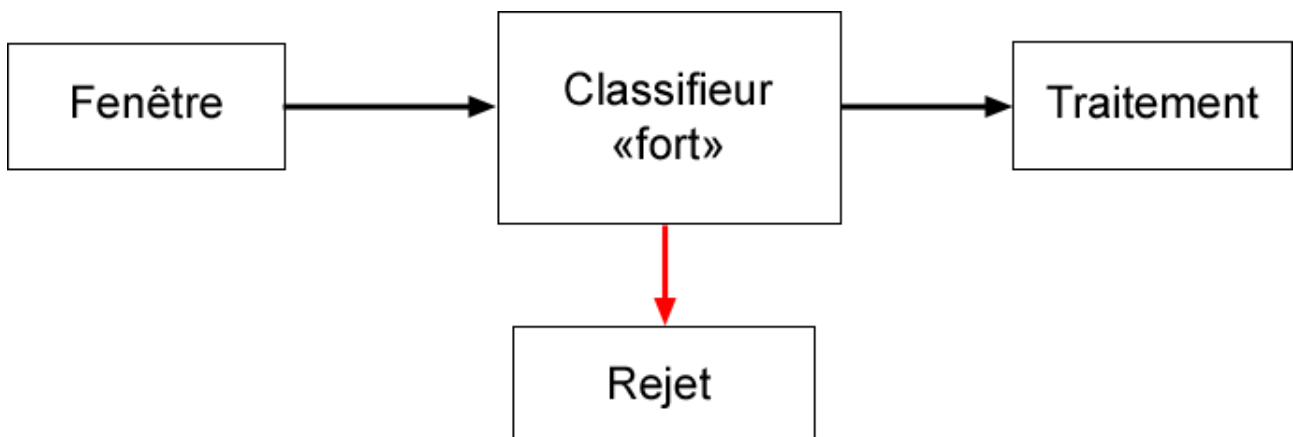


Figure 5: Fonctionnement d'un classifieur fort [2]

Il faudrait ainsi attendre que le classifieur ait analysé toute la fenêtre afin de savoir si un visage est présent dans l'image ou non. Une mise en cascade de classifieurs dont le critère de sélection serait moins sévère se présenterait de la sorte :

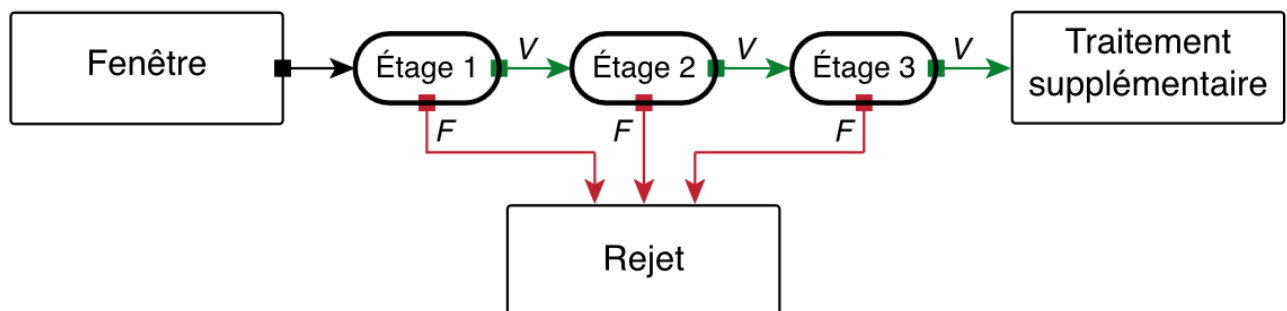


Figure 6: Mise en cascade de plusieurs classifieurs faibles

Ainsi dès que l'un des étages estime qu'il n'y a pas de visage, la fenêtre est rejetée et l'algorithme passe à la suite ce qui permet un gain de temps considérable.

Voici ainsi le résumé de l'algorithme :

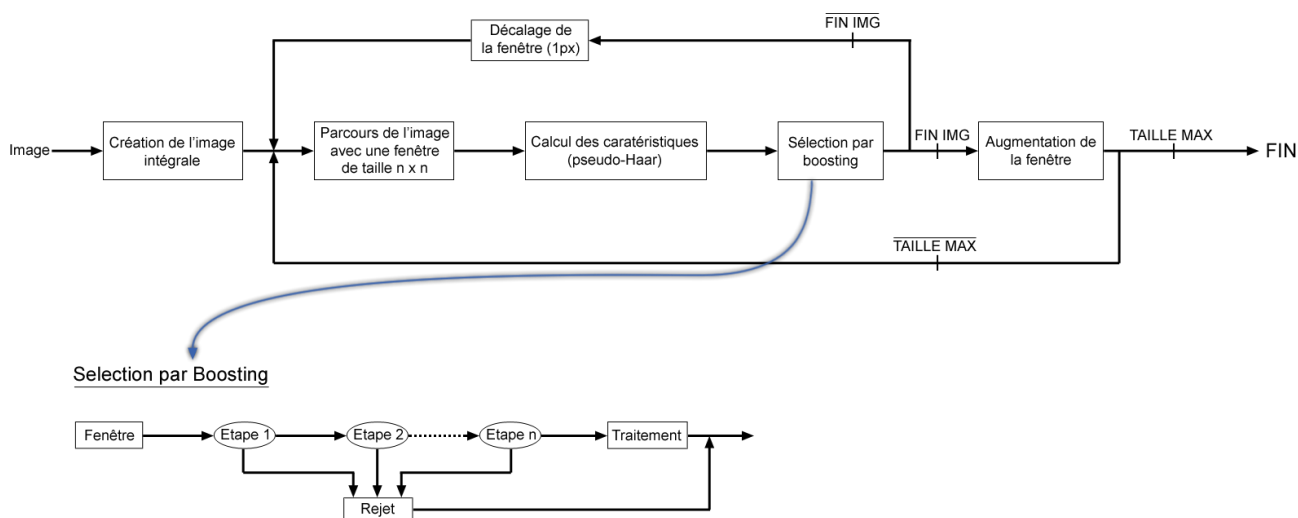


Figure 7: Résumé de l'algorithme de Viola & Jones

## 2. La numérisation de visage

### 2.1- Objectif

Maintenant que la zone contenant l'image est connue par le programme, il faut numériser le visage, c'est-à-dire l'enregistrer sous un format lisible par l'ordinateur afin de faire les comparaisons. Cela est fait par la méthode d'Eigenfaces ou des faces propres (Eigenfaces venant de l'allemand Eigen qui signifie propre et de l'anglais faces qui signifie face ou visages).

De manière simple, elle vise à diminuer la dimension de l'espace de travail pour simplifier les données et leur interprétation. Le but est ainsi de prendre en compte les informations importantes qui permettront de reconnaître un visage parmi d'autres avec un bon taux de réussite.

### 2.2- Méthodologie [4] [5]

Cette technique de reconnaissance utilise la méthode d'analyse en composantes principales ou la méthode de décomposition en valeurs singulières. Dans ce cadre, l'approche consiste à représenter un visage (celui à faire reconnaître) comme étant la combinaison linéaire d'un ensemble d'images, ces dernières formant une base de référence. Les  $m$  images de référence ou d'apprentissage sont chargées sous forme de matrices de dimensions  $l \times h$  (largeur\*hauteur en pixels) puis transformées en vecteurs de la manière suivante :



$$I_i = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1N} \\ a_{21} & a_{22} & \dots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \dots & a_{NN} \end{bmatrix}_{N \times N} \xrightarrow{\text{concatenation}} \begin{bmatrix} a_{11} \\ \vdots \\ a_{1N} \\ \vdots \\ a_{2N} \\ \vdots \\ a_{NN} \end{bmatrix}_{N^2 \times 1} = \Gamma_i$$

Figure 8: Transformation en vecteur des images de références

Ensuite, on déduit le visage moyen de l'ensemble des M images d'apprentissage. Il traduit l'ensemble des caractéristiques communes à tous les visages. Cela donne :

$$\Psi = \frac{1}{M} \sum_{i=1}^M \Gamma_i$$

Figure 9: Calcul du visage moyen

Puis, le visage moyen est soustrait de chaque visage. Ce faisant, il ne reste que l'ensemble des informations propres de chaque visage. Chaque visage présente ainsi des informations uniques sur le visage considéré.

$$\Phi_i = \Gamma_i - \Psi$$

Figure 10: Calcul des informations propres

$\Gamma_i$  Le vecteur image de référence

$\Phi_i$  Le vecteur des caractéristiques propres de l'image.

Ensuite, on fait la multiplication entre ce vecteur, et son vecteur transposé. On obtient ainsi les Eigenfaces qui seront utilisés pour la comparaison.

### 3- La comparaison

L'identification d'une personne consiste à trouver l'image qui lui ressemble le plus parmi les images de référence, ou si elle correspond à une image préenregistrée. La méthode est identique : retirer de l'image à trouver les caractéristiques moyennes puis calculer les pondérateurs avec l'ensemble des Eigenfaces connues. On calcul ensuite la

distance (méthode des moindres carrés) entre les pondérateurs des images de références et ceux de l'image à tester.

En fonction du résultat, on détermine si l'image est la même ou si elle est différente (on choisit généralement celles dont les pondérateurs sont inférieurs à 2).

## Conclusion

Ce document nous a permis de voir les différentes applications des mathématiques tout au long du processus de reconnaissance faciale. Dans chacune des trois étapes de son fonctionnement, les mathématiques interviennent, tant pour l'extraction de données que pour les différents calculs.

## Bibliographie

- [1] [En ligne].  
[https://fr.wikipedia.org/wiki/M%C3%A9thode\\_de\\_Viola\\_et\\_Jones#:~:text=La%20m%C3%A9thode%20de%20Viola%20et%20Jones%20est%20une%20m%C3%A9thode%20de,des%20objets%20dans%20une%20image..](https://fr.wikipedia.org/wiki/M%C3%A9thode_de_Viola_et_Jones#:~:text=La%20m%C3%A9thode%20de%20Viola%20et%20Jones%20est%20une%20m%C3%A9thode%20de,des%20objets%20dans%20une%20image..) [Accès le 10 02 2020].
- [2] [En ligne].  
<https://www.firediy.fr/article/face-tracking-implementation-de-la-methode-de-viola-jones-en-c>. [Accès le 10 02 20].
- [3] [En ligne].  
[https://fr.wikipedia.org/wiki/Image\\_int%C3%A9grale](https://fr.wikipedia.org/wiki/Image_int%C3%A9grale). [Accès le 11 02 2020].
- [4] [En ligne].  
<https://prezi.com/cecdphlhpznm/detection-et-reconnaissance-des-visages-par-methodes-de-pca/>. [Accès le 11 02 2020].
- [5] [En ligne].  
<https://silanus.fr/sin/?p=785#:~:text=Les%20eigenfaces%20sont%20un%20ensemble,la%20reconnaissance%20du%20visage%20humain>. [Accès le 11 02 2020].
- [6] [En ligne].  
<https://medium.com/@aaronward6210/facial-detection-understanding-viola-jones-algorithm-116d1a9db218>. [Accès le 11 02 2020].