



# Reconnaissance des formes

**Thème : Reconnaissance des visages**

Rédigé par : **BIAKOTA BOMBIA Herbert Céphas**  
**ANDRE Perrault**  
**MILORME Pierre Rubens**

*SOUS LA SUPERVISION DE :*  
**DR HO TUONG VINH**

**Année académique : 2016-2017**

## **Sommaire**

<b>Sommaire.....</b>	<b>2</b>
<b>Introduction.....</b>	<b>3</b>
<b>1. principe de la reconnaissance des visages.....</b>	<b>3</b>
<b>2. Etat de l'art des méthodes.....</b>	<b>3</b>
<b>2.1 Les algorithmes de reconnaissance de visage .....</b>	<b>4</b>
<b>2.1.1 Reconnaissance de visages par détection de contour.....</b>	<b>4</b>
<b>2.1.2 Reconnaissance de visages par réseaux de neurone.....</b>	<b>4</b>
<b>2.1.3 Reconnaissance de visages par SVM.....</b>	<b>4</b>
<b>2.1.4 Reconnaissance de visages par Analyse en composantes principales.....</b>	<b>4</b>
<b>2.1.5 Reconnaissance de visages par Analyse Linéaire Discriminante.....</b>	<b>4</b>
<b>3. Domaines d'application .....</b>	<b>5</b>
<b>I. introduction.....</b>	<b>5</b>
<b>II. Implémentation .....</b>	<b>5</b>
<b>III- Expérimentation et resultats.....</b>	<b>7</b>
<b>Conclusion.....</b>	<b>9</b>
<b>références scientifiques.....</b>	<b>10</b>

## Introduction

De nos jours, l'informatique est maintenant devenue un domaine indispensable. Pour ce fait, beaucoup de recherches et de connaissances ont été appliquées au domaine de la reconnaissance du visage, des formes, du traitement d'images, des statistiques. En plus, les capacités grandissantes des moyens informatiques et l'existence de bases de données de grande taille ont permis de mettre au point des algorithmes et des approches de plus en plus complexes.

Dans les années 90, le but de la recherche dans ce domaine se concentrait sur l'identification des visages enregistrés dans de bonnes conditions, dans des bases de taille relativement réduite. Les protocoles d'évaluation, qui en ont découlé, ont permis de mettre en place les premiers outils d'évaluation et de comparaison des algorithmes.

La reconnaissance des visages est la technique la plus commune et populaire parmi les techniques de reconnaissance biométrique. Elle reste la plus acceptable puis qu'elle correspond à ce que les humains utilisent dans l'interaction visuelle. Elle reçoit une attention accrue du fait de son caractère non invasif, au sens où elle ne requiert pas la coopération du sujet.

## 1 Le principe de reconnaissance des visages

La détection de visage a pour objectifs de décider si une image contient un ou plusieurs visages, et de donner leur localisation en d'autre terme elle consiste à trouver la position exacte du visage dans une image. Deux principales métriques caractérisent les performances d'un système de détection : **le taux de détection**, c'est-à-dire, le pourcentage de visages correctement détectés, et **le taux de faux positifs**, correspondant à des détections dans des régions ne contenant pas de visage.

## 2 Etat de l'art des méthodes

Durant cette dernière année plusieurs méthodes de reconnaissance de visages ont été proposées. Le but n'est pas ici de faire un résumé exhaustif de l'ensemble des méthodes déjà existantes dans le domaine.

La première catégorie regroupe les méthodes d'identification utilisant l'ensemble du visage comme l'information à traiter par le système. Ces méthodes sont souvent appelées méthodes globales ou holistiques et visent généralement à réduire l'espace de représentation du visage. On peut citer, parmi les plus connues, l'algorithme EigenFaces [1], FisherFaces [2], [3] ou encore l'analyse en composantes indépendantes (ACI) et l'évolution poursuite [4].

La deuxième catégorie regroupe les méthodes d'identification basées sur des caractéristiques particulières du visage (les traits). On les appelle les méthodes locales. Parmi les méthodes les plus répandues et appartenant à cette catégorie, on peut citer les méthodes basées sur le descripteur LBP [5], les méthodes basées sur les ondelettes de Gabor, la méthode EBGM [6] mais également la méthode présentée dans [7] où un nouveau descripteur de visage est proposé.

Les premiers systèmes de reconnaissance du visage sont commercialisés dès la fin des années 90 avec les produits des entreprises qui sont : FaceIt de Visionics, Visage Technology, FaceVACS de Plettac, FaceKey Corp, Cognitec Systems, Keyware Technologies, Passfaces de ID-

arts, ImageWare Software, Eyematic,,Interfaces Inc, BioID sensor fusion, Visionsphere Technologies, Biometric Systems, Inc, FaceSnap Recoder, SpotIt de face composite

Souvent une fenêtre détectée par le système est comptée comme valide par observation visuelle si la fenêtre inclut les yeux, le nez et la bouche [8]. Peu de travaux donnent une précision sur ce critère. Dans [9] par exemple, Lienhart avait évoqué qu'une localisation est correcte si la distance euclidienne entre les centres du visage détecté et du vrai visage est inférieure à 30% de la largeur du vrai visage, et la largeur du visage détecté est entre  $\pm 50\%$  de celle du vrai visage. Selon Garcia et Delakis [10] : un visage est considéré comme correctement détecté si la taille de la fenêtre englobante n'excède pas plus de 20% de la taille réelle de la région faciale, et qu'elle contient les yeux, le nez et la bouche. Dans [11] les auteurs considèrent une bonne localisation si les mesures de la position (à travers la position des yeux) et la taille (à travers la distance entre les yeux) du visage ne dépassent pas 30% des vraies valeurs. Jesorsky et al. [12] ont introduit une mesure d'erreur basée sur la distance de Hausdorff entre les vrais centres des yeux et ceux détectés.

## **2.1 Les algorithmes de reconnaissance de visage**

Plusieurs méthodes de reconnaissance de visage ont été proposées durant les trois dernières décennies. La reconnaissance du visage est devenue un challenge intéressant qui a attiré les chercheurs de différentes spécialités : reconnaissance de visages par détection de contour, reconnaissance de visages par réseaux de neurones, reconnaissance de visages par SVM

### **2.1.1 Reconnaissance de visages par détection de contour**

La détection de contour est une étape préliminaire à de nombreuses applications de l'analyse d'images. Les contours constituent en effet des indices riches, au même titre que les points d'intérêts, pour toute interprétation ultérieure de l'image.

### **2.1.2 Reconnaissance de visages par réseaux de neurone**

Un réseau neuronal est l'association, en un graphe plus ou moins complexe, d'objets élémentaires, les neurones formels. Les principaux réseaux se distinguent par l'organisation du graphe (en couches, complets. . . ), c'est-à-dire leur architecture, son niveau de complexité (le nombre de neurones, présence ou non de boucles de rétroaction dans le réseau), par le type des neurones (leurs fonctions de transition ou d'activation) et enfin par l'objectif visé : apprentissage supervisé ou non, optimisation, systèmes dynamiques.

### **2.1.3 Reconnaissance de visages par SVM**

Les **machines à vecteurs de support** ou **séparateurs à vaste marge** (en anglais *support vector machine*, SVM) sont un ensemble de techniques d'apprentissage supervisé destinées à résoudre des problèmes de discrimination et de régression. Les SVM sont une généralisation des classifieurs linéaires. Les séparateurs à vaste marge ont été développés dans les années 1990 à partir des considérations théoriques de Vladimir Vapnik sur le développement d'une théorie statistique de l'apprentissage

### **2.1.4 Reconnaissance de visages par Analyse en composantes principales**

En 1987 et 1990, Sirovich et Kirby [39] publient leurs travaux sur la reconstruction de visage à partir de vecteurs de projection par ACP (Analyse en Composantes Principales) ou KL (Karhunen-Loève). Ils montrent la validité de la réduction de dimensionnalité de l'espace des visages et aussi qu'une reconstruction de visage à partir de la nouvelle base est même meilleure sur le plan visuel, car elle corrige une grande partie des bruits ou des zones occultés lors de la capture. Cette approche a été reprise par Turk en 1991 [93] et appliquée à la reconnaissance des visages. Cette méthode reste d'actualité et même considérée comme une méthode de base. Elle est intégrée dans les deux systèmes de référence de FRGCv2 et de Biosecure .

### 2.1.5 Reconnaissance de visages par Analyse Linéaire Discriminante

L'Analyse Linéaire Discriminante (LDA, Linear Discriminant Analysis ou Fisher Discriminant Analysis) permet de rechercher, dans l'espace des données, les axes qui permettent de discriminer au mieux les différentes classes (plutôt que les axes qui décrivent au mieux les données, tel est le cas de l'ACP).

### 3 Domaines d'application

Les systèmes d'identification de personnes basés sur reconnaissance du visage deviennent de plus en plus répandus et trouvent des applications très variées dans la vie de tous les jours. Beaucoup d'algorithmes ont vu le jour depuis une trentaine d'années et sont essentiellement utilisés pour le contrôle d'individus dans des lieux publics tels que les aéroports ou privés comme certaines grosses entreprises. Mais la majorité de ces systèmes fonctionne dans un environnement dit contrôlé. C'est à dire que toutes les conditions pouvant détériorer la reconnaissance sont maîtrisées. Les individus sont coopératifs si bien que les visages sont acquis de face, sans expression ni variation d'illumination. Or, dans le cas de la vidéo surveillance, il est impossible de contrôler les conditions d'acquisition que ce soit au niveau du comportement des individus (pose, expression...) ou au niveau des conditions d'éclairage.

## Solution proposée

### I- Introduction

Suite différentes informations que nous avons pu recenser sur le travail qui nous a été donné sur l'état de l'art de la reconnaissance de visage, nous nous allons maintenant proposer une solution qui sur laquelle nous allons nous baser

#### L'acquisition de données

Nous avons décidé de faire tout d'abord une phase d'acquisition de données qui va nous permettre de détecter tous les visages des personnes dans les images s'il y en a et dont ils seront aussi le contenu de notre base d'apprentissage ainsi que celui de test [1].

#### L'extraction des caractéristiques

Pour cette partie, nous allons devoir utiliser le descripteur model binaire local parce qu'il est invariant aux transformations géométriques

#### La classification

Cette étape sera réalisée à l'aide de l'algorithme k plus proche voisin qui est une variante biologiquement inspirée des MLP, il possède différents types de couches et chaque couche différente fonctionne différemment des couches MLP habituelles [5] et [6]

### Évaluation et interprétation

Comme nous le savons, tout model en machine learning et deep learning doivent être évalué afin de savoir la performance du système. De ce fait il est aussi important d'interpréter les résultats que nous fournira notre système en basant sur la **matrice de confusion et le taux précision**

Afin de réaliser ce travail nous allons implémenter notre système sur python utilisant la librairie **opencv, Keras, scikit-learn, tensorflow** et les données de la base « **orl\_faces** ». [5]

## II- Implémentation

Suite aux directives qui nous ont donné sur le choix de l'implémentation basé sur les réseaux neurones convolutionnels, nous avons pu essayer malheureusement nous nous sommes confrontés à quelques difficultés basés sur la non maîtrise du concept qui l'accompagne. Ainsi nous avons pu réaliser ce travail avec la méthode KNN (k plus proche voisin).evoquée ci haut dans la solution proposée. L'implémentation du système s'est faite en python, en utilisant la librairie opencv2 sous l'IDE Anaconda. Nous avons utilisé le fichier « haarcascade\_frontalface\_alt.xml » proposé par opencv pour la détection des visages puis appliqué l'algorithme de viola et jones. Le programme implémenté est nommé reconnaissance\_visage.py et pour le compiler il faut utiliser Jupyter puis se positionner dans le répertoire ou se trouve le fichier et importer le contenant le code source à l'aide de la commande **%load reconnaissance\_visage.py** et ensuite taper sur la combinaison des touches **shift et enter**. Ainsi pour la réalisation de notre travail nous avons défini deux fichier txt : apprentissage.txt et test.txt ». Le fichier « apprentissage.txt » décrit la base d'apprentissage et « test.txt » la base de test. Chaque ligne de ces deux fichiers contient deux informations, le chemin vers une image suivi de la classe associée à cette image , les deux informations étant séparées par un point virgule « ; ». La figure ci-dessous montre un extrait du fichier apprentissage.txt.

```

170 base_apprentissage/s15/2.pgm;33
171 base_apprentissage/s2/5.pgm;34
172 base_apprentissage/s2/3.pgm;34
173 base_apprentissage/s2/4.pgm;34
174 base_apprentissage/s2/1.pgm;34
175 base_apprentissage/s2/2.pgm;34
176 base_apprentissage/s14/5.pgm;35
177 base_apprentissage/s14/3.pgm;35
178 base_apprentissage/s14/4.pgm;35
179 base_apprentissage/s14/1.pgm;35
180 base_apprentissage/s14/2.pgm;35
181 base_apprentissage/s40/5.pgm;36
182 base_apprentissage/s40/3.pgm;36
183 base_apprentissage/s40/4.pgm;36
184 base_apprentissage/s40/1.pgm;36
185 base_apprentissage/s40/2.pgm;36
186 base_apprentissage/s1/5.pgm;37
187 base_apprentissage/s1/3.pgm;37
188 base_apprentissage/s1/4.pgm;37
189 base_apprentissage/s1/1.pgm;37
190 base_apprentissage/s1/2.pgm;37
191 base_apprentissage/s7/5.pgm;38
192 base_apprentissage/s7/3.pgm;38
193 base_apprentissage/s7/4.pgm;38
194 base_apprentissage/s7/1.pgm;38
195 base_apprentissage/s7/2.pgm;38
196 base_apprentissage/s28/5.pgm;39
197 base_apprentissage/s28/3.pgm;39
198 base_apprentissage/s28/4.pgm;39
199 base_apprentissage/s28/1.pgm;39
200 base_apprentissage/s28/2.pgm;39

```

Figure 1: Extrait du fichier apprentissage.txt

### III- Expérimentations et résultats

Nous avons réalisé nos expérimentations sur les données de la base « The ORL Data of faces » et sur une base d'images réalisée par nous-mêmes. La base ORL est constituée de 400 images de 40 personnes à raison de 10 images par personne. Notre base d'apprentissage sera constituée de 200 images obtenues en prenant les 5 premières images de chacune des 40 personnes de la base. La base de test sera constituée des 200 images restantes. Notons que le bloc de prétraitement n'est pas utilisé lors de nos expérimentations avec cette base car les images de la base sont déjà prétraitées.

La base d'images réalisée par nous-mêmes est composée des images de 8 personnes (3 images par personne). 32 images sont utilisées pour la base d'apprentissage (2 images par personne) et 2 images pour la base de test (2 images par personne). Pour cette base nous utilisons le bloc de prétraitement car les images ne sont pas de même taille, elles sont également en couleurs et avec des niveaux d'illumination différents.

- Pour notre première expérimentation nous appliquons le descripteur LBP uniquement en considérant 8 voisins autour de chaque pixel, un rayon égal à 1 et des valeurs de K pour l'algorithme du KNN comprises entre 1 et 3. Les résultats obtenus sont consignés dans le tableau ci-dessous:

Distances/ K	K = 1	K = 2	K = 3
CHI 2	88%	83%	80,5%

Bhattacharyya	91,5%	87%	86%
---------------	-------	-----	-----

Tableau 1 : Résultats de la première expérimentation nombre de voisins =8 et rayon =1 (base ORL)

On remarque que les meilleurs résultats sont obtenus pour  $K = 1$  avec un pic à 91,5% pour la distance d'intersection entre les histogrammes et que notre méthode nous permet d'obtenir 90,5% de taux de reconnaissance en moyenne. Notre implémentation permet donc d'atteindre un taux élevé de reconnaissance de visages. Toutefois ce résultat est inférieur à ceux obtenus dans la littérature 96% à 99% mais il est à préciser que contrairement au cas présentés dans l'état de l'art nous n'associons la méthode LBP à aucune autre méthode ici.

La figure ci-dessous nous présente la matrice confusion obtenue pour la reconnaissance de visage avec la distance d'intersection et  $K = 1$ .

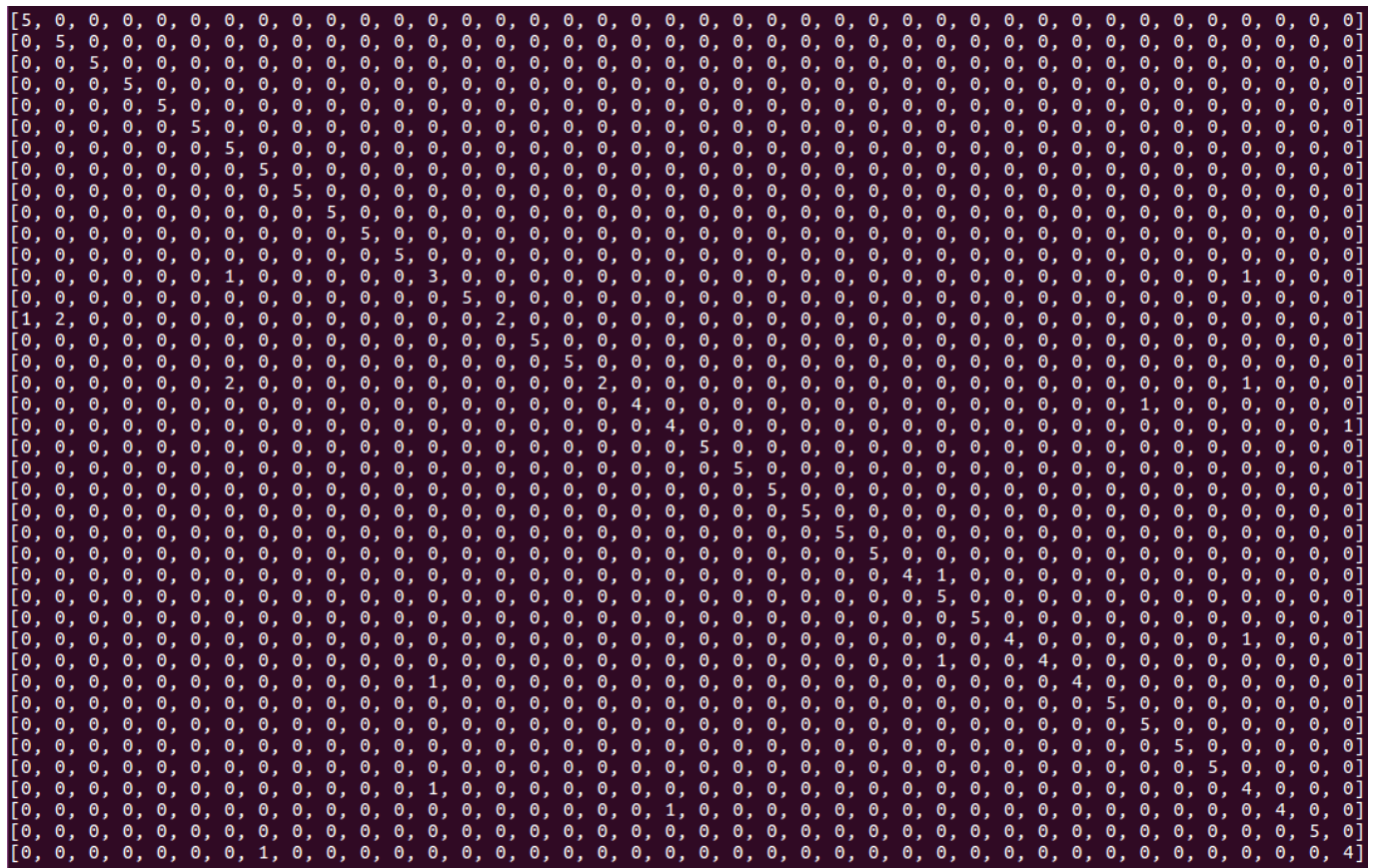


Figure 2: Matrice de confusion obtenue avec la distance d'intersection,  $K = 1$ , nombre de voisins =8 et rayon =1 (base ORL)

Après analyse de la matrice nous remarquons que les images qui ont été le plus mal classées sont celles des classes 13 et 15. Nous présentons donc plus bas quelques images de ces deux classes. Notons que pour identifier les dossiers correspondant aux différentes classes, il suffit de se référer au fichier « test.txt » accompagnant le code source.





**Figure 3:** Images provenant dans l'ordre des classes 13 et 15 (base ORL)

L'analyse de ces différentes images fait ressortir que les individus avaient presque tous des expressions faciales (sourire) et pour celui de la classe 13 nous notons aussi la présence de lunettes. Nous concluons donc que ces éléments sont des sources d'erreur pour la classification des images car ils modifient les caractéristiques des visages. L'autre facteur à prendre en compte également est l'orientation des visages parce qu'elle modifie la structure du visage et peut être source d'erreurs.

Pour notre deuxième expérimentation nous appliquons le descripteur LBP uniquement en considérant 8 voisins autour de chaque pixel, un rayon égal à 1. La valeur de K pour le KNN est fixée à 1. La base de données utilisée est celle constituée par nous-mêmes. Les résultats obtenus sont consignés dans le tableau ci-dessous :

Distances/ K	K = 1
CHI 2	18,75%
Bhattacharyya	18,75%

**Tableau 2 :** Résultats de la deuxième expérimentation nombre de voisins =16 et rayon =2 (base personnelle)

À ce niveau les résultats obtenus diffèrent de ceux des expériences précédentes. Le meilleur taux n'est que de 18,75% et est obtenu avec la distance de corrélation des histogrammes. Ce taux plus faible peut s'expliquer par le fait que les images de notre base n'ont pas été prises dans les mêmes conditions d'éclairage, avec le même arrière-plan et n'ont pas été prétraitées de façon professionnelle comme pour la base ORL. Le fait que le meilleur taux soit obtenu avec une distance différente de celle des cas précédents nous montre que l'on ne peut pas généraliser sur la meilleure distance à prendre en compte lors de la comparaison des histogrammes. Il est nécessaire d'essayer chacune d'elles et de retenir à chaque fois la meilleure suivant la nature des images qui sont traitées.

La figure ci-dessous nous présente la matrice de confusion obtenue pour 18,75% de taux de reconnaissance.

[0, 0, 1, 1, 0, 0, 0, 0]
[0, 0, 1, 0, 0, 0, 0, 1]
[0, 0, 0, 0, 0, 0, 0, 2]
[0, 0, 0, 2, 0, 0, 0, 0]
[0, 0, 1, 1, 0, 0, 0, 0]
[0, 0, 1, 1, 0, 0, 0, 0]
[0, 0, 1, 0, 0, 0, 1, 0]
[0, 1, 0, 1, 0, 0, 0, 0]

**Figure 4:** Matrice de confusion obtenue avec la distance de Bhattacharyya,  $K = 1$ , nombre de voisins = 8 et rayon = 1 (base personnelle)

Nous remarquons que beaucoup d'images sont mal classées sous celles de la classe (classes4) se



## Conclusion

De tout ce qui précède, dans ce présent rapport nous avons pu définir d'une manière générale la reconnaissance des visages, chercher des articles scientifiques qui parlent de la des reconnaissances des visages avec quelque concepts qui l'accompagnent. Bien que le principe de reconnaissance des visages reste mais nous avons pu relater différentes méthodes qui ont permis la mise en place des applications de reconnaissance des visages qui sont utilisées dans plusieurs domaine

## Références scientifiques

- [1] M. Turk, A. Pentland, "*Eigenfaces for recognition*", Journal of cognitive Neuroscience, vol. 3, pp. 71-86, 1991.
- [2] P. Belhumeur, J. Hespanha, D. Kriegman, "*Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection*", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, pp. 711-720, July 1997.
- [3] Aleix M. Martinez et Avinash C. Kak: *PCA versus LDA*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(2) :228–233, 2001. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.28>.
- [4] Chengjun Liu et Harry Wechsler: *Evolutionary Pursuit and Its Application to Face Recognition*. IEEE Trans. Pattern Anal. Mach. Intell., 22 :570–582, June 2000, ISSN 0162-8828.

- [5] Timo Ahonen, Abdenour Hadid et Matti Pietikainen: *Face Recognition with Local Binary Patterns*. Dans *Lecture Notes in Computer Science*, tome 3021/2004, pages 469–481, Berlin, Heidelberg, 2004. Springer-Verlag.
- [6] Laurenz Wiskott, Jean Marc Fellous, Norbert Krüger et Christoph von der Malsburg: *Face recognition by elastic bunch graph matching*, pages 355–398. CRC Press, Inc., Boca Raton, FL, USA, 1999, ISBN 0-8493-2055-0. <http://portal.acm.org/citation.cfm?id=320684.320713>.
- [7] Ngoc Son Vu et Alice Caplier: *Face Recognition with Patterns of Oriented Edge Magnitudes*. Dans Kostas Daniilidis, Petros Maragos et Nikos Paragios (rédacteurs) : *Computer Vision ECCV 2010*, tome 6311 de *Lecture Notes in Computer Science*, pages 313–326. Springer Berlin / Heidelberg, 2010, ISBN 978-3-642-15548-2. [http://dx.doi.org/10.1007/978-3-642-15549-9\\_23](http://dx.doi.org/10.1007/978-3-642-15549-9_23).
- [8] R.-J. Huang, "*Detection strategies for face recognition using learning and evolution*", Ph.D. thesis, George Mason University, Fairfax, Virginia, 1998.
- [IV2.fr] Le site web du projet IV<sup>2</sup>: <http://iv2.ibisc.fr/>.
- [9] R. Lienhart, A. Kuranov, V. Pisarevsky, "*Empirical analysis of detection cascades of boosted classifiers for rapid object detection*", Proceedings of the 25th DAGM-Symposium, Magdeburg, Germany, pp. 297–304, 2003
- [10] C. Garcia, et M. Delakis, "*Convolutional Face Finder : A Neural Architecture for Fast and Robust Face Detection*", IEEE Transactions on Pattern Analysis and Machine Intelligence, 26(11), pp.1408-1423, 2004.
- [11] B. Fröba, C. Küblbeck, "*Robust face detection at video frame rate based on edge orientation features*", in: Proceedings of the 5th IEEE International Conference on Automatic Face and Gesture Recognition AFGR, pp. 342–347, 2002.
- [12] O. Jesorsky, K. Kirchberg, R. Frischholz, "*Robust face detection using the hausdorff distance*", proceedings of the 3rd International Conference on Audio- and Video-Based Biometric Person Authentication (AVBPA), pp. 90–95, 2001.
- [39] Kirby and M Sirovich. Application of the karhunen-loeve procedure for the characterization of human faces. Pattern Analysis and Machine Intelligence, 12 :103\_108, Jan 1990.
- [93] M. Turk and A. Pentland. Eigenfaces for recognition. journal of Cognitive Neuroscience, 3 :71\_86, 1991.
- [1] M. Turk, A. Pentland, "*Eigenfaces for recognition* ", Journal of cognitive Neuroscience, vol. 3, pp. 71-86, 1991.
- [2] P. Belhumeur, J. Hespanha, D. Kriegman, "*Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection* ", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, pp. 711-720, July 1997.

[5]<http://www.pyimagesearch.com/2016/08/01/lenet-convolutional-neural-network-in-python/>

[6]<https://translate.google.com.vn/translatehl=fr&sl=en&u=https://pythonprogramming.net/convolutional-neural-network-cnn-machine-learning-tutorial/&prev=search>

<http://www.FaceIt.com>

<http://www.viisage.com>

<http://www.plettac-electronics.com>

<http://www.facekey.com>

<http://www.cognitec-systems.de>

<http://www.keywareusa.com/>

<http://www.id-arts.com/>

<http://www.iwsinc.com/>

<http://www.eyematic.com/>

<http://www.bioid.com>

<http://www.visionspheretech.com/menu.htm>

<http://www.biometrica.com/>

<http://www.facesnap.de/htdocs/english/index2.html>

<http://spotit.itc.it/SpotIt.html>