



DETECTION ET RECONNAISSANCE DES VISAGES BASEE SUR LES RESEAUX DE NEURONES ARTIFICIELS

R. ADJOUJ

Evolutionary Engineering and Distributed Information Systems Laboratory, EEDIS Computer Science Department University of Sidi Bel-Abbès, Algeria Tel/Fax: (213)-48-57 77 50 Email : AdjReda@yahoo.fr

Dr A. BOUKELIF

Laboratoire des Télécommunications et du Traitement Numérique de Signal, équipe de recherche des techniques vidéo Département Electronique, Université de Sidi Bel Abbès, Algérie Tél : (213)-48-57 82 81 Email : aboukelif@yahoo.fr

Résumé :

Ce document démontre comment un système de reconnaissance des visages peut être réalisé par un réseau de neurones artificiel. A noter que le processus d'apprentissage ne consiste pas à un seul appel à la fonction d'apprentissage. Par contre, le réseau de neurones doit faire l'apprentissage plusieurs fois sur une variété des images, dont elles contiennent des visages.

Dans ce cas, l'apprentissage du réseau sur différents ensembles d'images lui force à apprendre comment il se comporte vis-à-vis la variété des visages, un problème commun dans le monde réel.

Mots clés :

Détection des visages, vision par ordinateur, réseaux neurones artificiels (NNs), classification, apprentissage du réseaux neurones artificiels, comparaison des visages, reconnaissance des visages, détecteur de caractéristiques, extraction des caractéristiques, identification des visages, traitement des images, estimation de la pose du visage, émotion du visage.

Abstract :

This document demonstrates how a face recognition system can be designed with artificial neural network. Note that the training process did not consist of a single call to a training function. Instead, the network was trained several times on various input images, the images which contents faces.

In this case training a network on different sets of images forced the network to learn how to deal with variety of faces , a common problem in the real world.

1. Introduction

Actuellement, tous les pays sont intrigués par l'identification rapide des personnes criminelles ou d'autres, dans les zones piétonnes, les aéroports, les frontières et toutes sortes de bâtiments (Banques, administrations,...).

Les caméras de vidéo-surveillance ne sont pas capables de fournir l'identité des personnes

observées. Toutefois, l'opération d'identification est assurée par un opérateur humain.

L'efficacité d'un agent dans la reconnaissance des personnes s'avère longue et fastidieuse lorsque le nombre est assez important, ce qui provoque d'éventuelles erreurs.

D'où la nécessité de chercher un agent automatique pouvant compenser cette dégradation d'efficacité ou simplement de restreindre le champ de recherche, ce qui facilite la tâche de reconnaissance pour l'opérateur humain.

Ce système doit donc reconnaître rapidement des personnes à partir de leurs photos uniquement.

Cette reconnaissance est-elle possible par un ordinateur ?

Les réseaux de neurones (Neural Networks NN) ont pour objectif la classification et la prédiction. La plupart des algorithmes de reconnaissance de visages cherchent à extraire un code descriptif réduit du visage (entre 10 et 500 paramètres), puis à rattacher ce code à une classe statistique correspondant à un individu.

La méthode employée consiste à localiser les informations utiles pour le système c'est à dire localiser seulement le visage, extraire des caractéristiques à partir de ce visage, comparer ces caractéristiques avec d'autres qui existent dans la base de données images .Au cas où ces caractéristiques se correspondent alors le visage est identifié ou reconnu.

La réduction du champ de la recherche (qui est le rôle de la procédure de classification) limite le nombre des visages à parcourir en choisissant seulement ceux qui sont susceptibles de correspondre au visage cherché, c'est à dire ceux qui ont des caractéristiques qui répondent au critère de classification puis on procédera à l'identification.

L'authentification par reconnaissance de visages se heurte à plusieurs limitations et ne sera jamais assez fiable pour des applications de haute sécurité[8]:

Un humain moyen atteint un taux de reconnaissance de 99%. *Les NN implémentés sur ordinateur peuvent-ils le dépasser ?*

- Toute population comporte environ 1% de vrais jumeaux qui ont deux visages différents mais qui ont les mêmes distances métriques .
- Les humains peuvent parvenir à les reconnaître avec de l'expérience, qu'en est-il des ordinateurs ?
- Comment l'ordinateur pourra-t-il comprendre qu'il n'a affaire qu'à une photo d'une photo, ou pire, d'un masque ?

2. les méthodes de reconnaissance et d'identification des visages :

Parmi les méthodes de reconnaissances des visages, on cite :

- L'approche à base de vues .
- L'approche fondée sur l'apparence
- L'approche des caractéristiques invariantes
- L'approche de Reconnaissance des visages basée sur les NNs[10]:

2.1. L'approche à base de vues :

Cette technique utilise une représentation vectorielle du visage, ce vecteur est le résultat d'une détection faite au préalable, et il contient les caractéristiques qui ont été énumérées : f_1, f_2, \dots, f_n , et que la représentation du vecteur contient en premier les mesures de f_1 puis de $f_2 \dots$ jusqu'à f_n .

Les mesures pour chaque f_i incluent son emplacement (x, y) et l'intensité ou la valeur de la couleur locale de ce point, c'est-à-dire pour chaque f_i on a un triplet (x_i, y_i, T_i) [6].

2.2. L'approche fondée sur l'apparence :

C'est une technique fondée sur l'apparence pour caractériser et rechercher des visages dans une base de données images. Cette méthode est invariante à la translation, à la rotation plane, et au changement d'échelle et elle peut capter l'information significative de l'image sans recours à la segmentation. Elle est également très bien adaptée au faible changement de point de vue et elle est également robuste au bruit et aux occultations.

2.3. L'approche des caractéristiques invariantes:

Cette approche consiste à enregistrer des caractéristiques dans la vue d'exemples qui ne changent pas, caractéristiques telles que la couleur ou les invariants géométriques, lorsque les paramètres "pose, expression et éclairage" changent. Cette approche n'est pas applicable dans la reconnaissance de visage alors qu'elle présente de bons résultats pour l'indexation d'articles d'épicerie emballés utilisant la couleur.

2.4. L'approche de Reconnaissance des visages basée sur les NN [10]:

Le professeur Tom Mitchell avec son équipe du laboratoire de l'intelligence artificielle de CMU, ont initié le projet de reconnaissance des visages en utilisant les NN artificiels.

Le professeur Tom Mitchell a abouti enfin à une architecture énorme, plus de 20 NN pour reconnaître 20 personnes, il en résulte une architecture colossale; les résultats sont favorables[10].

Avec cette méthode on a besoin d'incrémenter l'architecture du NN quand on veut reconnaître plus de personnes. A un moment donné, on sera limité par la technologie.

3. PRINCIPALES DIFFICULTES DES APPROCHES:

La performance d'un algorithme de reconnaissance de visages dépend principalement du nombre de paramètres statistiquement indépendants extraits, et de leur variance.

- La variance doit être faible au sein d'une même classe
- La variance entre les classes doit être grande

On se heurte alors à trois principales sources de variations au sein d'une même classe: l'éclairage, l'angle de vue ou la pose, et l'expression du visage. Généralement, ces trois paramètres engendrent des variations plus grandes que les variations entre classes distinctes, à paramètres constants!

Dans une première approche, on peut chercher à fixer les paramètres instantanés, en standardisant les modalités d'acquisition: expression neutre, sans lunettes, éclairage de face, etc. Une approche plus générale consiste à prendre tous ces paramètres en compte. Il en résultera un algorithme plus complexe, avec une phase d'apprentissage beaucoup plus importante (photos du même visage sous plusieurs éclairages, avec différentes expressions).

3.1. APPROCHE GEOMETRIQUE DE DETECTION ET DE RECONNAISSANCE DES VISAGES :

3.1.1. Localisation des visages:

Des visages sont issus d'images de photos d'identités; la position du visage est à peu près dans le même endroit. Lorsque le système localise les limites du visage automatiquement, il va distinguer entre ce qui se trouve à l'intérieur, qui est le visage lui-même "foreground", et ce qui est à l'extérieur des limites du visage qui est le "background". Après cette distinction, le système va opérer seulement sur le visage donc sur le "foreground".

3.1.2. Classification :

Une fois le visage localisé, le système peut le classifier puis extraire les points caractéristiques tels que les centres des yeux et la bouche, pour le faire passer au module d'identification. Le critère de classification est la proportion de la largeur du visage (l) sur sa hauteur (h) (Figure 1). Cette proportion implique la création de trois classes distinctes :

- La **première** contient les visages qui ont une forme ovale, tel que la proportion varie dans l'intervalle [0.5...0.64].
- La **deuxième**, la classe normale, tel que la proportion varie dans l'intervalle [0.66...0.7].
- La **troisième** classe celles des visages qui convergent vers la forme d'un cercle, tel que la proportion varie dans l'intervalle [0.71...1.00].

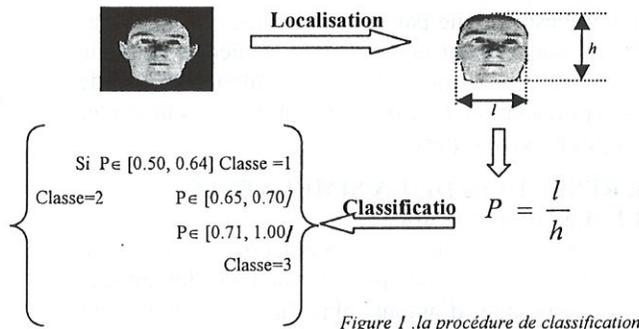


Figure 1 .la procédure de classification

3.1.3. Détecteur des caractéristiques ("Feature Finder") :

La procédure précédente permet de localiser le visage, c'est à dire à partir d'une image (représentant la photo d'identité d'une personne) avoir la sous fenêtre contenant le visage. Reste à extraire les différentes caractéristiques pour ce visage. Ces caractéristiques sont être essentiellement : (Figure 2)

- Un point centre de l'œil gauche → $P_1(x_1, y_1)$
- Un point centre de l'œil droit → $P_2(x_2, y_2)$
- Un point centre de la bouche → $P_3(x_3, y_3)$
- Un point centre du nez. → $P_4(x_4, y_4)$

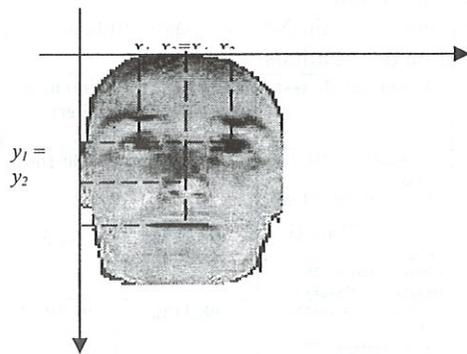


Figure 2 : Les points Caractéristiques visage

3.1.4 .Extraction des caractéristiques faciales du visage

♦ **Paramétrisation du visage:**

Dans la section précédente, on a montré comment le système peut extraire le centre de l'œil.Par ce même principe le système peut extraire quatre points du visage.

La paramétrisation est basée sur les relations métriques entre ces quatre points[4 -5]: .

Par paramétrisation, le visage est numérisé par un vecteur spécifiant les relations métriques existants entre ces différents points caractéristiques[4]:. Le système fournis essentiellement cinq paramètres (Figure 3) :

- D_1 : Distance entre l'œil gauche et l'œil droit : $D_1 = x_2 - x_1$.(figure 2)
- D_2 , D_3 : Distances entre les yeux et le nez :
- D_4 , D_5 : Distances entre les yeux et la bouche

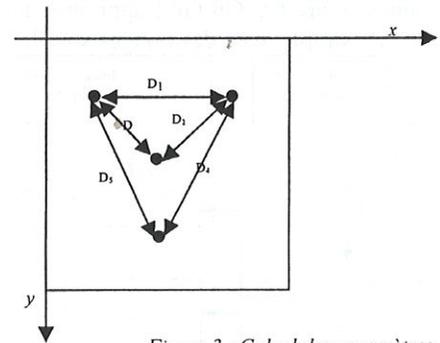


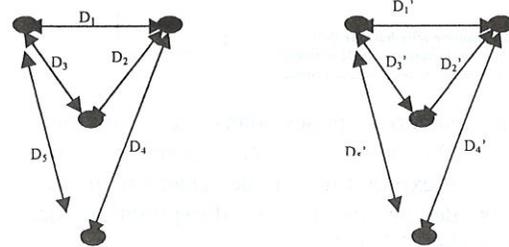
Figure 3 : Calcul des paramètres

3.1.5. Identification

L'identification est le noyau de ce système : C'est la comparaison des caractéristiques du visage de l'input avec celles des visages de la base de données. Cette comparaison est basée sur des notions géométriques telles que la proportion des distances, les euclidiennes. Si la proportion varie d'un rapport à l'autre, ce ne sont donc pas les mêmes distances alors le visage ne correspond pas au visage de l'input.(figure 4) .

Visage Input

Un visage de la base de



La proportion $\frac{D_1}{D_1'} = \frac{D_2}{D_2'} = \frac{D_3}{D_3'} = \frac{D_4}{D_4'} = \frac{D_5}{D_5'} = \alpha$ (α Constante)

Figure 4 : Correspondance entre le visage input et un visage de la base de données.

3.2. APPROCHE DE DETECTION ET DE RECONNAISSANCE BASEE SUR LES RESEAUX NEURONES ARTIFICIELS :

Le système est composé de cinq modules :module de détection, le module d'estimation de pose du visage, le module de reconnaissance d'émotion du visage et le module de détection de la présence ou absence des lunettes de soleil sur le visage. L'architecture du

système ainsi que les principales étapes sont illustrées dans la figure 5.

3.2.1. NN de reconnaissance :

Le NN est implémenté avec une couche cachée de 100 neurones et une couche pour l'input de 960 neurones qui reflètent le nombre total des pixels (32 x 30 pixels) et une couche d'output avec 20 neurones codés de façon à représenter les 20 personnes (figure 6). On fait l'apprentissage, puis le test du NN simple avec des images simples sans

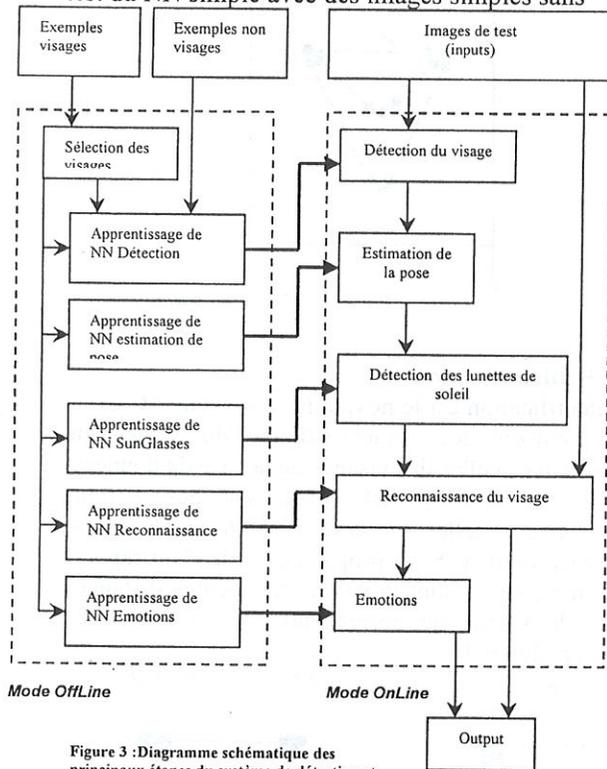


Figure 3 : Diagramme schématique des principales étapes du système de détection et reconnaissance basé sur les NN développés.

variations majeures prises dans des conditions favorables. On aboutit au NN après plusieurs tentatives d'exploration et de changement des paramètres de ce dernier et d'exploration des architectures des NNs [2].

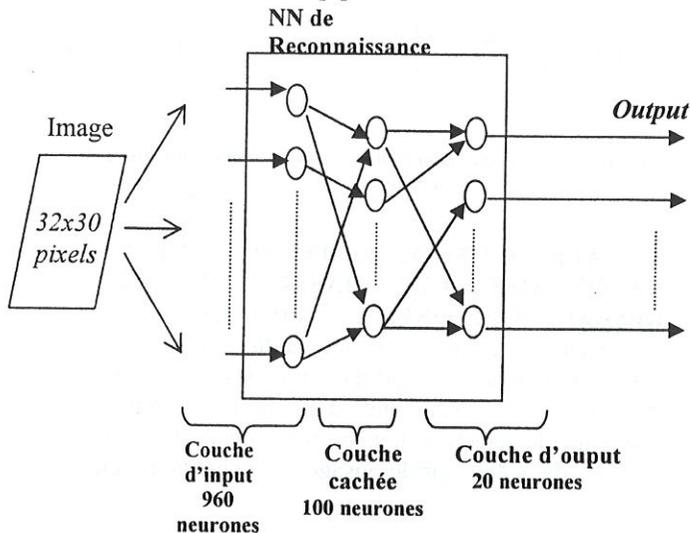


Figure 7 : Réseaux de neurones de reconnaissance

3.2.3. Description de l'architecture :

Le NN de reconnaissance est un réseau de neurones à 2 couches ; la 1^{ère} couche correspond à la couche cachée avec 100 neurones, la 2^{ème} couche aux sorties avec 20 neurones. La fonction de transfert utilisée dans chaque neurone de cette architecture c'est la Log-Sigmoïd La fonction Log-Sigmoïd est choisie parce que le rang des outputs du NN est de 0 à 1. Elle est parfaite pour l'apprentissage des valeurs booléennes comme outputs [7]. La couche cachée a 100 neurones; ce nombre est désigné par expérience [3]. Les neurones des 2 couches sont massivement connectés au mode Feedforward qui utilise l'algorithme de Backpropagation (rétro-propagation) pour régler ses poids synaptiques [2-11].

4. RESULTATS DE LA SIMULATION

4.1. La reconnaissance :

Le NN a montré des meilleurs résultats de reconnaissance surtout pour l'ensemble des images d'apprentissage d'autant plus que l'ensemble ne comporte que 38 images; le NN a reconnu toutes les images de l'ensemble (tableau 1).

Architecture (module)	Ensemble de test	Taux de Reconnaissance	Erreur de Reconnaissance
Le NN de Reconnaissance	217 images		
	Les images de MIT 38 images	97,36 %	01/38
	Les images de CMU : 80 Images (20 x 4 Visages) 20 personnes avec 4 images différentes (4 émotions)	93,75%	05/80
	Les images "Divers" : 39 images	97,74%	02/39

Tableau 1 : Résultats sur la reconnaissance.

4.2. NN de détection

Après l'apprentissage du NN, on passe au teste, et à la récupération des résultats :

L'architecture (module)	Ensemble de test	Taux de détection	Erreur de détection
Le NN de détection	Les images de CMU 100 Images (80 avec visages +20 sans Visages)	100 % *	00/100
	Les images de MIT 58 images (38 avec visages +20 images sans Visages)	98,27%	01/58
	Les images divers (59 images) (39 avec visage + 20 images sans visages)	98,31%	01/59 **

Tableau 2 : Résultats sur la détection.

4.3. Résultats du test de performance du NN estimation de pose

Une fois que le NN Émotion a fait l'apprentissage, vient le rôle des exemples de test et de validation de l'architecture (voir Tableau 3).

4.4. Apprentissage du NN Sunglasses

Systèmes	Ensemble de test	Taux de Détection	Erreur de détection
Le module NN de SunGlasses (détection des lunettes de soleil) dans l'approche neuronale	Les images de CMU 60 Images de visages (40 porte SunGlasses + 20 sans SunGlasses)	96,67%	02/60
	Les images de MIT 38 images. (ils sont tous sans SunGlasses, lunettes de soleil)	94,74%	02/38
	Les images diverses 39 images (ils sont tous sans lunettes de soleil, Sunglasses)	20,51% ***	31/39

Systèmes	Ensemble de test	Taux de reconnaître d' Émotion	Erreur de reconnaître d'Émotion			
Le module NN De reconnaissance l'Émotion du visage dans l'approche neuronale	Les images de CMU 80 Images de visages (20 en état neutre, 20 heureux, 20 tristes, 20 en colère)	70 %	24/80			
			neutre	heureux	triste	en colère
			14/20	02/20	05/20	03/20
	Les images de MIT 38 images. (il sont tous en état neutre)	*** (le système a reconnue qu'ils sont tous en colère)	38/38			
			neutre	heureux	triste	en colère
			/	/	/	38/38
	Les images diverses 39 images	61,54%	15/39			
			neutre	heureu	triste	en colère
			/	09/39	03/39	03/39

Tableau 3 : Taux et erreur de reconnaissance de l'émotion du visage dans l'image d'input.

4.5. Résultats de test de performance du NN Estimation de pose:

Systèmes	Ensemble de test	Taux d'estimation de la pose	Erreur d' Estimation de la pose			
Le module NN d' Estimation de la pose du visage dans l'approche neuronale	Les images de CMU 80 Images de visages (20 en face, 20 en haut , 20 à gauche,20 à droite,)	90,00%	08/80			
			En Face	En haut	A gauche	A droite
			00/20	00/ 20	03/20	05/20
	Les images de MIT 38 images. (il sont tous en vue de face)	73,68%	10/38			
			En Face	En haut	A gauche	A droite
			00/38	07/38	00/38	03/38
	Les images divers 39 image (il sont tous en vue de face)	30% ***	26/39			
			En Face	En haut	A gauche	A droite
			00/39	20/39	01/39	05/39

Tableau 5 : Taux & l'erreur d'Estimation de la pose du visage dans les deux ensemble de CMU et d' MIT.

5. Comparaison entre les caractéristiques des deux approches (Géométrique/Neuronale)

Approche géométrique classique	Approche neuronale
<ul style="list-style-type: none"> • Pré-traitement : <ul style="list-style-type: none"> -Rendre l'image en Noir & blanc . (valeurs des pixels soit 1,soit 0) -Basée sur des photos en 2D ou 3D, transformées après en 2D,pour qu'elles seront acceptables par le système. • Détection du visage : <ul style="list-style-type: none"> -Manuelle. -Automatique :recherche du visage par le parcours pixel par pixel (les pixels de valeurs = 1) . (temps assez important de recherche) -Le traitement de bas niveau se réalise en noir et blanc (image binaire) • Reconnaissance : <ul style="list-style-type: none"> -1^{er} : Extraire les points caractéristique du visage détecté. -2^{ième} :calcule des distances qui relie ces points caractéristiques. -3^{ième} :calcule des rapports de ces distances . -4^{ième} :comparaison des rapports calculés avec les rapports des visages stockés dans la base de données visages, l'erreur la plus faible restitue le visage comme le plus proche du visage d'<i>input</i>, alors ce dernier est reconnu. • Remarques : <ul style="list-style-type: none"> -Une seule photo de la personne doit être disponible au niveau de la base de données images pour qu'il soit reconnu. -Aucune contrainte n'est posée au préalable sur les photos de l'<i>input</i>. <ul style="list-style-type: none"> -Identification non liée à la texture¹ du visage. -La base de données des images est donc théoriquement illimitée². -Applicable pour des photos contenant des personnes. -Les dimensions géométriques peuvent présenter une incohérence au niveau de l'identification. 	<ul style="list-style-type: none"> • Pré-traitement : <ul style="list-style-type: none"> -Pas de contraintes sur les valeurs des pixels. (l'image en niveau de gris, et même en couleur est acceptée). -Nécessite un ensemble des images d'apprentissage exhaustif et de nombre important, ainsi un ensemble de test pour valider l'architecture du NN conçu. • Détection du visage : <ul style="list-style-type: none"> Automatique : Se fait automatiquement dans une seule tâche, par le "NN Détection", et par simulation du NN en mode <i>Online</i> en temps négligeable, et après apprentissage bien sûr en mode <i>offLine</i> (par contre un temps assez important dans ce second mode). • Reconnaissance : <ul style="list-style-type: none"> Se fait automatiquement dans une seule tâche, par le "NN reconnaissance", et par simulation du NN en mode <i>Online</i> (après apprentissage bien sûr en mode <i>offLine</i>) • Extension du système vers d'autre possibilités : <ul style="list-style-type: none"> -Estimation de la pose en profondeur (par NN estimation pose) . -Détection de la présence ou l'absence des lunettes de soleil sur le visage (par NN SunGlsses) - savoir l'émotion de la personne (par NN Emotions) • Remarques : <ul style="list-style-type: none"> -Le système nécessite une interprétation des outputs des NN pour avoir les résultats . - Si nous avons plus de personnes à reconnaître, donc plus de neurones, ce qui implique une architecture ascendante du NN. - Image de dimension différente de 32x30 pixels, implique redimensionnement du vecteur d'<i>input</i> des NNs.

¹ Couleur du pixel ou son niveau de gris

² les limites physiques (matérielles) telles que la capacité du disque dur ou de la mémoire vive.

6. CONCLUSION

Cet article montre que le problème de détection et de reconnaissance des visages peut être résolu efficacement en utilisant une approche implémentée avec les NN (réseaux neurones) artificiels, et basée sur les modèles des visages, ces modèles sont considérés comme des exemples d'apprentissage pour le NN (réseaux neurones). L'approche neuronale ainsi que l'approche géométrique (classique) qui sont décrites dans cet article présentent des avantages et des limites, mais elles aident l'opérateur humain énormément dans sa tâche... Elles pourront être utiles dans des applications, telles que [2] :

- Identification des gens suspects à partir de leurs photos d'identité ou leurs portraits,
- Identification des voyageurs (contrôle de l'immigration), afin de prévenir les actes terroristes,
- Au niveau des établissements de haute surveillance, l'autorisation n'est donnée qu'aux personnels du service identifiés au préalable (la photo est soit prise par une carte personnelle ou par acquisition du visage à partir d'une caméra de surveillance),
- Dans le sondage, pour éviter la redondance des personnes,
- Recherche des enfants disparus (à partir d'une caméra de surveillance placée dans une zone piétonnière).

Parmi les perspectives dans le domaine d'identification des visages, des chercheurs américains explorent l'idée de description de l'apparence physique de la personne, notamment le visage, à partir de l'ADN trouvé par exemple sur le lieu du crime.

Une fois le visage est reconstruit il est passé au sous module d'identification pour le reconnaître (à un certain pourcentage).

En plus, parmi les perspectives dans le domaine d'identification et authentification des personnes, des travaux sont focalisés sur l'idée de réalisation d'un système de haute sécurité qui se base sur l'authentification automatique par identification et reconnaissance des personnes à partir des images [1]. Ces images contiennent le visage, l'empreinte digitale, le geste effectué par le corps, l'iris de l'œil, l'écriture manuscrite, ainsi que le signal vocal.

L'objectif de ces travaux consiste à améliorer l'efficacité et le taux de l'authentification, en combinant les résultats de reconnaissance et d'identification, et en utilisant les concepts de l'intelligence artificielle, tels que les réseaux neurones...

Une fois les différents composants de reconnaissance de ce système établis, une comparaison des taux d'authentification sera effectuée. ceux-ci seront plus performant vue le nombre de techniques combinées dans ce système.

7. Bibliographie :

- [1] Karl Baumgartner, Stefano Ventura, Phd, rapport de recherche : "Biométrie", Ecole supérieure d'informatique et de Télécommunication, suisse « ESIT », juillet 2002.
- [2] Réda Adjoudj, Thèse de MAGISTER, Spécialité Informatique, Option Intelligence artificielle intitulée : « Détection & Reconnaissance des Visages En utilisant les Réseaux de Neurones Artificiels », Université de Djillali Liabès, département d'informatique, SBA , Algérie, Octobre 2002.
- [3] Mounir boukadoum, Phd, Spécialité utilisation des NN dans l'industrie , Conférence intitulée : "utilisation des réseaux neurones artificiels pour résoudre les problèmes complexes , qui ne suit aucun modèles", à l'université Abou-bakr-Belkaïd de Tlemcen, par Mr Mounir Boukadoum , Mai 2002.
- [4] Adjoudj Réda, "Détection & reconnaissance des visages par une méthode géométrique", Rapport de recherche , Université de Djillali Liabès, département d'informatique, de SBA, Algérie.
- [5] Adjoudj Réda, "Système automatique d'identification & reconnaissance des visages ", Mémoire de fin d'étude d'ingénieur , Université de Djillali Liabès, département d'informatique, SBA, Juin 2000.
- [6] Henry A.Rowley , "Neural Network-based face detection", PhD thesis, Carnegie Mellon University, Pittsburgh, USA, May 1999.
- [7] Howard Demuth, Mark Beale, "Neural Network Toolbox User's Guide" For Use with MATLAB® by The MathWorks, Inc.1998.
- [8] John DAUGMAN, " Face and Gesture Recognition": Overview, IEEE PAMI, vol.19, no.7, July 1997.
- [9] Jeffrey S. Norris, " Face Detection and Recognition in Office Environments ", Submitted to the Department of Electrical Engineering and Computer Science in Partial Fulfillment of the Requirements for the Degree of Master of Engineering in Electrical Engineering and Computer Science at the Massachusetts Institute of Technology , M.I.T, USA, May 21, 1999.
- [10] Tom Mitchell, Home work 3, "neural net & Face images" CMU, October 1997.
- [11] Jérôme Lacaille, " Réseaux de neurones" DIAM – CMLA, ENS de Cachan, septembre 1996.