



MASTER HANDI
Domaine : Sciences Technologie, Santé (STS)
Mention: Ingénierie et Cognition
Spécialité: Technologie et Handicap
Rapport de stage M2

Application assistée par ordinateur pour les malvoyants par acquisition et reconnaissance d'images statiques et dynamiques

Présenté par :
Enmar KHALIS

Directeur de stage : RAZAFINDRAMARY Donné
Lieu du stage : Supméca
Organisme : LISMMA



Coordinateur :
J. LOPEZ KRAHE

Paris, september/2013



SOCRATES *Community action programme
in the field of education*

Remerciements

Je tiens en premier lieu à remercier mon tuteur de stage Monsieur Donné RAZAFINDRAMARY de m'avoir permis de réaliser ce modeste travail au sein de son équipe au laboratoire LISMMA. Il a su me faire bénéficier de son expérience et de sa compétence. Je tiens à le remercier pour la confiance qu'il m'a toujours témoignée, en m'accordant à la fois une large autonomie et un soutien permanent.

J'exprime ma reconnaissance à M. le professeur Jaime LOPEZ-KRAHE qui m'a accordé l'occasion de réaliser mes études de qualité en master 2 au sein de la formation *Technologie et Handicaps*.

Je tiens enfin à remercier l'équipe enseignante de cette formation qui j'ai passé cette année tout en effectuant une formation de qualité et dans une ambiance sympathique.

Application assistée par ordinateur pour les malvoyants par acquisition et reconnaissance d'images statiques et dynamiques

Résumé	5
Introduction	6
1. Introduction générale	6
1.1 Systèmes de reconnaissance faciale	6
1.2 Difficulté de la vision par ordinateur	6
2 Présentation du laboratoire de recherche (LISMMA)	8
2.1 Statut et partenaires industriel et universitaire	8
2.2 Equipes en activité et les thématiques et les axes de recherches	9
3 Choix du stage	13
● Chapitre 1 Handicapés malvoyants	15
1-1 Définition et Généralités	15
1-2 Causes de déficience visuelle	16
1-3 Qui est affecté par la déficience visuelle?	17
1-4 Incidence de la déficience visuelle sur le développement	18
● Chapitre 2 MATLAB	24
2-1 Qu'est-ce que MATLAB?	24
2-2 Le système MATLAB	25
2-3 Interprétation VS compilation	26
2-4 Pourquoi choisir MATLAB?	27
2-5 Faiblesses de MATLAB	28
2-6 Mise en route MATLAB	28
2-7 Figures, courbes et graphiques	36
2-8 Traitement de l'image	44
2-8.1 Qu'est-ce que l'image numérique?	45
2-8.2 Traiter les images numériques avec des algorithmes informatiques	45
2-8.3 Formats d'image supportés par Matlab	45
2-8.4 Aspects du traitement de l'image	46
2-8.5 Types d'images numériques	46
2-8.6 Comment convertir entre différents formats	47
2-8.7 Commandes générales utilisées dans le traitement de l'image	49
● Chapitre 3 Algorithmes / Méthodes	57
3-1 Méthode de Viola et Jones	57
3-2 Cascade Object Detector	62

3-3 Local Binary Pattern (LBP)	64
3-4 Suivi de points dans une vidéo à l'aide de l'algorithme de Kanade-Lucas-Tomasi (KLT)	71
3-5 APPLICATION	75
● Perspectives	83
● Conclusions	85
● Bibliographique	86

Résumé

Les visages sont des objets spéciaux d'intérêt considérable. Développer des systèmes automatisés de détection et de reconnaissance des visages est utile dans une variété de domaines d'application, y compris **l'aide aux personnes ayant une déficience visuelle** et la gestion des collections d'images à grande échelle. Ce mémoire se concentre sur l'émulation de cette approche en utilisant différents types d'informations contextuelles pour améliorer la performance des différentes méthodes de détection de visage et reconnaissance faciale.

Dans ce travail, l'algorithme de reconnaissance de visage est présenté sur la base d'un mixte de loi Manhattan doublement tronqué et multi variée avec transformation discrète en cosinus (DCT) et le modèle binaire locale (LBP). Ici, l'image faciale d'entrée est transformée au domaine de configuration binaire locale. L'image de la configuration binaire locale obtenue est divisée en blocs qui ne se chevauchent pas. Puis à partir de chaque bloc, des coefficients DCT sont calculées et des vecteurs de caractéristiques sont extraits. Du fait que le vecteur de caractéristique suit une loi Manhattan doublement tronqué et multi variée, l'image du visage est ainsi modélisée. En utilisant l'algorithme Expectation Maximisation, les paramètres du modèle sont estimés. L'initialisation des paramètres du modèle se fait en utilisant l'algorithme K-means.. De plus, l'efficacité du système de reconnaissance faciale développé est évaluée en effectuant des expériences avec des bases de données d'image de visage du personnel. La performance mesure environ un tiers du taux d'erreur total (1/3). Finalement, une étude comparative de l'algorithme développé avec une partie du l'algorithme précédemment existant a révélé que ce système fonctionne mieux lorsque les informations locale et globale du visage sont utilisées.

Introduction

1 Introduction Générale

1.1 Systèmes de reconnaissance faciale

La reconnaissance faciale automatique est une problématique de vision par ordinateur intéressante avec beaucoup d'utilisation à but commerciale et de lois de répression de leurs applications. L'utilisation et le contrôle d'accès utilisateur, la surveillance de la foule, l'aide des personnes à facultés de vision affaiblies et le renforcement de l'interaction homme machine deviennent possible si un système de reconnaissance de visage efficace peut être mise en œuvre. Bien que la recherche dans ce domaine remonte aux années 1960, ce n'est que très récemment que des résultats acceptables ont été obtenus. Toutefois, la reconnaissance du visage est encore un domaine de recherche actif depuis, une approche où un modèle complètement réussi n'a pas été proposée pour résoudre le problème de la reconnaissance des visages.

L'insuffisance des systèmes de reconnaissance faciale est particulièrement évidente lorsqu'on les compare à notre propre capacité de reconnaissance faciale innée. Nous effectuons la reconnaissance faciale, une tâche visuelle extrêmement complexe, presque instantanément, et notre propre capacité de reconnaissance est beaucoup plus robuste que celle de n'importe quel ordinateur peut espérer avoir. Nous pouvons reconnaître une personne familière dans des conditions d'éclairage très défavorables, à partir d'angles ou de points de vue divers. Les différences d'échelle (un visage étant de près ou de loin), les horizons différents n'affectent pas notre capacité à reconnaître les visages et on peut même reconnaître les personnes avec seulement une fraction de leur face visible et même après le passage de plusieurs années. En outre, nous sommes en mesure de reconnaître les visages de plusieurs milliers d'individus que nous avons rencontrés au cours de notre vie.

1.1.2 Difficulté de la vision par ordinateur

- Le niveau de vision et de compréhension qui est instinctif pour nous (les humains) est encore hors de la portée de la création de la vision par ordinateur.

- La difficulté principale des problèmes de vision est que la quasi-totalité d'entre eux sont mal définis.

- La complexité des problèmes de vision par ordinateur est exacerbée par le fait que nous avons à faire avec d'énormes quantités de données. Une image typique d'échelle de gris a 640x480 pixels, chacun ayant 8 bits (256) des valeurs d'intensité (niveaux de gris). Par conséquent, la taille de l'image entière est $640 \times 480 \times 8 \text{ bits} = 2457600 \text{ bits}$. Tout calcul d'algorithme avec une grande complexité serait extrêmement lent en vision par ordinateur et nous devons donc faire un effort pour résoudre ces problèmes en utilisant des techniques de traitement très simples.

Cependant, même avec toutes ces contraintes, il est possible d'obtenir des résultats utiles en vision par ordinateur en réduisant la généralité d'un problème. Le problème du domaine d'application de la vision par ordinateur peut être limité à un environnement structuré et des hypothèses pourront être faites à propos de l'éclairage, des types d'objet, etc.

1.2 Présentation du LISMMA

Appellation: Laboratoire d'Ingénierie des Systèmes Mécaniques et des Matériaux (LISMMA)

2.1 Statut et partenaires industriel et universitaire

Le Laboratoire d'Ingénierie des Systèmes Mécaniques et des Matériaux (LISMMA) est une Unité Propre de Recherche ayant un statut EA 2336.

LE LISMMA est un organisme public de recherche, composé de plusieurs équipes de recherche réparties sur différents sites universitaires à l'instar de SUPMECA à Saint-Ouen, l'IUT de Tremblay-en-France et l'IUT de Montreuil.

Le Laboratoire traite de sujets de recherche pluridisciplinaires mais complémentaires relevant des domaines de la mécanique et des matériaux. Les recherches menées au **Laboratoire d'Ingénierie des Systèmes Mécaniques et des Matériaux (LISMMA)** concernent l'ingénierie des ensembles et des systèmes mécaniques. Elles ont un caractère pluridisciplinaire, recouvrant le génie mécanique, la mécanique, la mécatronique et le génie industriel. Les activités du LISMMA concernent : la physique des matériaux, la conception mécanique (partenariat avec Dassault systèmes pour le développement de CATIA), les systèmes logistiques, la tribologie (étude de tous les problèmes liés aux contacts) et la vibro-acoustique (vibrations, bruits, matériaux amortissants).

Les activités de recherches sont centrées sur la **conception et la conduite de systèmes complexes dans les domaines de l'ingénierie mécanique**. Elles se placent dans une approche systémique et concernent en priorité l'analyse et la modélisation des systèmes complexes. L'objectif est ainsi de développer de nouvelles approches méthodes et des outils adaptés à ces systèmes.

Le LISMMA à travers notamment Supméca a, par ses activités de recherche, développé un grand nombre de partenariats, tant universitaires qu'industriels, avec des projets financés au niveau régional, national, et européen.

Ainsi, les enseignants-chercheurs et personnels techniques du Laboratoire travaillent en forte imbrication avec le monde industriel au travers de divers partenariats et

collaborations (DASSAULT SYSTEMES, ALSTOM, ASTRIUM BOSCH, CNES, EADS, SAFRAN, RENAULT, PSA, VALEO, etc.).

Par ailleurs, à travers ses différents co-encadrements de thèses existants, le **LISMMA** peut se prévaloir d'un grand nombre de partenariats universitaires, nationaux (Université du Maine, Université de Technologie de Compiègne.. .), européen (Université de Technologie Chimique et de Métallurgie de Sofia), et internationaux (Université de Sherbrooke, Ecole Supérieure des Sciences et Techniques de Tunis...).



2.2 Equipes en activité et les thématiques et les axes de recherches

Le LISMMA regroupe aujourd'hui un peu plus de **75 personnes** qui font partie des différentes équipes, réparties sur les trois sites évoqués antérieurement dont **33 enseignants-chercheurs** (professeurs des universités et maîtres de conférences) et **31 doctorants**.

L'ingénierie numérique regroupe les outils et méthodologies associées à l'étude (Conception et simulation) du produit et de ses procédés de fabrication, et à la gestion des données et informations du cycle de vie du produit. Contexte se à la virtualisation du produit et de son processus de simulation pour accélérer les processus de conception.

Trois thématiques de recherche

✓ Modélisation géométrique pour la maquette numérique

L'objectif général des projets de recherche en cours est à la fois d'améliorer les performances en termes d'analyse et de résolution de problèmes géométriques 3D et de généraliser notre modélisation aujourd'hui purement géométrique à l'ensemble des effets physiques usuels en ingénierie mécanique.

✓ Ingénierie des systèmes mécatroniques

Les systèmes mécatroniques sont des systèmes intégrant des technologies informatiques, mécaniques électroniques et automatiques. L'évolution de la technologie a donné naissance à ces systèmes complexes qui améliorent les rendements, ajoutent de la fonctionnalité, de l'autonomie dans de nombreuses applications. L'objectif général est donc d'utiliser la simulation et spécialement le prototypage virtuel dans le cycle de vie de conception d'un système, dans une optique de réduction des coûts de réalisation de prototypes réels et d'augmentation de la qualité des systèmes.

✓ Conception & optimisation des systèmes industriels

Les recherches portent sur la synchronisation des décisions de conception et de production, réalisée par la synchronisation de celles du PLM (Product Life Management) avec celles des ERP (Enterprises Ressources Products) et APS (Advanced Planning System : la logistique de conception), la définition de modèles de décisions adaptés à l'automatisation de la conception, la modélisation et la résolution de problèmes de conception intégrée produits/réseaux logistiques, et l'évaluation de la robustesse et de la stabilité de tels réseaux.

Les travaux menés dans cette thématique de tribologie et matériaux s'inscrivent dans le cadre des économies d'énergie par réduction des frottements, des gains de matière par minimisation de l'usure ou par optimisation des propriétés des

matériaux. Les 3 axes de recherche, développés dans cette thématique, associent des travaux expérimentaux sur moyens d'essais originaux, mais également des travaux théoriques comprenant modélisations et simulations.

Trois axes de recherche

✓ Fatigue de contact

Les études portent sur la prise en compte de la triaxialité des contraintes de contact dans les assemblages soumis à des amplitudes de sollicitations dans les directions normales ou tangentielles. Les derniers développements portent sur les apports introduits par l'utilisation de revêtements.

✓ Microgéométrie et contact

L'objectif est de parfaire la connaissance de l'impact des microgéométries sur le comportement tribologique du contact. Les domaines abordés sont les contacts statiques (étanchéité, transferts thermiques dans les interfaces) mais principalement les contacts lubrifiés travaillant en conditions sévères (régime de lubrification mixte). Les études actuelles concernent la prévision du frottement et de l'usure, y compris lors des phases de rodage.

✓ Comportement thermomécanique des matériaux

Cet axe regroupe des études orientées sur la caractérisation des effets de synergie existant entre sollicitations mécaniques et diverses actions thermochimiques. Les travaux sont centrés sur la problématique de transport ou de stockage d'hydrogène sous haute pression, et sur la tenue mécanique des joints soudés sous sollicitations mécaniques brutales (situation de type crash-test).

Ainsi, les aspects phénoménologiques du comportement élasto-inélastique — avec ou sans endommagement des matériaux — sont abordés expérimentalement et théoriquement. On s'intéresse à l'aspect théorique et à la modélisation du comportement des matériaux avec pour objectif de développer des modèles et des lois de comportement. Le domaine des grandes déformations expérimentales des structures en dynamique intermédiaire est également étudié. Soulignons ici comme applications et réalisations: (i) l'étude des systèmes d'absorption d'énergie ; (ii) une nouvelle méthodologie expérimentale relative à l'élaboration des grains ultrafins

(cette dernière est le fruit d'une collaboration avec le Laboratoire des Propriétés Mécaniques et Thermodynamiques des Matériaux de l'université Paris 13).

Les axes de recherche de l'équipe vibro-acoustique et structure concernent l'analyse vibratoire et acoustique des systèmes mécaniques et des matériaux qui les composent. Cette démarche est menée dans un souci de réduction des vibrations et du bruit: réduction lors de l'émission et réduction à la réception dans une optique fortement étayée par l'expérimentation.

✓ Trois thématiques de recherche

Amortissement des vibrations de structures

Deux principales techniques sont mises en œuvre par l'équipe:

– L'amortissement semi actif basé sur l'usage des patches piézo-céramiques shuntés par des circuits électroniques simples. Les recherches concernent l'évaluation du couplage électromécanique effectif, proposé comme l'indicateur de l'amortissement ajouté ainsi que la recherche de circuits électroniques induisant un amortissement shunté large bande (multimodale).

– L'amortissement passif, effectué par des matériaux viscoélastiques classiques, est pris en compte dans les problèmes d'instabilités provoquées par la présence de frottement sec dans le contexte de phénomènes de basses (broutement) et hautes (crissement) fréquences.

✓ Dynamique des structures

Cette thématique est centrée autour de l'analyse des comportements vibratoires des systèmes complexes en basses et moyennes fréquences: vibrations linéaires et non linéaires. Les quatre axes principaux étudiés sont les vibrations et bruits de contact qui génèrent des instabilités, la recherche de sources d'amortissement par micro-frottements, la détection et la quantification de l'endommagement dans les matériaux et structures, et enfin les mécanismes de propagation d'onde dans les milieux hétérogènes non cohésifs de type granulaires. L'originalité de cette thématique est non seulement de proposer des modèles physiques spécifiques, mais surtout de développer des outils numériques et expérimentaux adaptés: généralisation de la méthode des éléments continus pour étudier les moyennes fréquences, développement de visco-élasticimètres pour matériaux anisotropes et/ou fortement

amortissant, mesure de la dynamique du contact à l'échelle microscopique, et enfin mise en œuvre de capteurs et traitement des signaux spécifiques.

✓ Isolation acoustique et matériaux

Les réductions du bruit sont étudiées et mises en œuvre essentiellement par les 2 techniques suivantes: conception et caractérisation de matériaux spéciaux (amortissant, isolants, absorbants) et particulièrement des matériaux composites (sandwich monolithique perforé à âme en nid d'abeilles par exemple), et conception et caractérisation de complexes d'isolation par agencement de matériaux et structures divers. Des logiciels et des bancs de mesure spécifiques sont développés par l'équipe pour prévoir d'une part les propriétés vibro-acoustiques des matériaux et d'autre part celles des structures.

1.3 Choix du stage

La thématique de ce stage est à l'interface de l'axe de recherche de l'ingénierie numérique et celui de la tribologie et matériaux. Il est réalisé sous la direction de M. RAZAFINDRAMARY. Les développements de ce dernier sont articulés autour de deux axes de recherche: optimisation numérique appliquée aux modèles micromécaniques de comportement des matériaux et aux robotiques. En effet, cette démarche pluridisciplinaire associe le mécanicien des matériaux au spécialiste en génie de logiciel dans un but d'optimisation de la programmation des modèles, généralement lourds pour des applications complexes. De plus, une élaboration de logiciels avec des interfaces conviviales constitue également un autre volet du fruit de cette collaboration pluridisciplinaire.

Pour ces raisons, travailler sur une problématique basée sur des développements numériques est le fondement de mon stage avec, bien attendu, des applications plutôt axées sur l'acquisition et reconnaissance d'images statiques et dynamiques pour les malvoyants assistée par ordinateur. Ce thème et ses applications attirent l'attention de plusieurs chercheurs et fédérations au niveau national et international. Comme exemple, la fédération française SYROBO (regroupant nombre de petites et moyennes entreprises aux recherches et technologies innovantes) travaille actuellement dans le domaine du traitement vidéo, notamment sur un projet alliant

vidéo détection et géo localisation permettant de guider par la voix et le son une personne atteinte de déficience visuelle vers un objectif (ou un objet) donné. Ainsi, cette manière de travailler permet d'utiliser des caméras dont la résolution n'est pas la première qualité sur un système embarqué. Une technologie d'avenir donc, qui ne permettra peut-être pas aux non-voyants de voir, mais possiblement d'évoluer sans trop de problèmes dans un environnement inconnu [1].

Le stage a lieu au sein de l'équipe d'accueil sise à l'IUT de Tremblay-en-France pendant une période d'avril à septembre 2013.

Chapitre 1

Handicapés malvoyants

Les applications ciblées dans ce stage ont tendance à se concentrer sur l'acquisition et la reconnaissance des images statiques et dynamiques pour les personnes assistées qui sont visuellement dépréciées, où dans ce chapitre cette catégorie de personnes est bien définie et classée les causes de leur déficit, et l'impact de leur carence en leur développement, depuis la première analyse des besoins a commencé durant cette étude.

1-1 Définition et généralité

Un trouble de la vision désigne une déficience visuelle d'une personne qualifiée par un niveau d'incapacité visuelle plus ou moins importante causée par une maladie, un traumatisme, ou par des conditions congénitales, brutales ou progressives, qui requièrent une correction médicale, chirurgicale ou l'aide de verres correcteurs.

Cela conduit à [2]:

- La perte de l'acuité visuelle et l'incapacité de la personne à voir des objets aussi clairement comme une personne saine
- La perte du champ visuel impliquant l'incapacité d'un individu à voir une zone aussi vaste que la personne moyenne sans bouger les yeux ou tourner la tête.
- La photophobie - incapacité à regarder la lumière
- La diplopie - vision double
- La distorsion visuelle ou la distorsion des images
- Les difficultés de vision perceptuelle ou des difficultés de perception
- Ou une combinaison des caractéristiques ci-dessus

1-2 Causes de déficience visuelle

Les causes courantes qui entraînent une déficience visuelle ou une perte de la vision sont les suivants:

- Blessures aux yeux : blessures aux yeux se produisent lors de la lecture ou au travail ou en raison d'accidents peuvent entraîner une détérioration ou une perte de la vision, les blessures à la cornée sont plus particulièrement la principale cause de perte de vision ;
- Maladies héréditaires de cécité et de déficience visuelle ;

- La rétinite pigmentaire est la cause la plus fréquente de cécité héréditaire.

Infections des yeux

Parfois, lors de la grossesse, si la future mère a été affectée par une infection virale comme la rubéole, celle-ci peut transmettre l'infection au fœtus pendant sa grossesse. Ainsi, le nouveau né peut naître avec une cécité ou une déficience visuelle.

Le trachome des yeux provoqué par des micro-organismes contagieux appelé Chlamydia trichomatisme peut également endommager la vue d'œil. Cela se voit dans les pays développés et sous-développés via la qualité des eaux et les installations sanitaires délabrés.

L'amblyopie

Il s'agit essentiellement de troubles de la vision dans un œil en raison de l'absence de son utilisation dans la petite enfance.

Cela se voit dans le strabisme ou "œil paresseux (**amblyopie**)" puisque les deux yeux projettent différemment et envoient différents messages au cerveau. Ce dernier peut alors désactiver ou supprimer des images de l'œil le plus faible. Cela empêche le développement de l'œil le plus faible conduisant à l'amblyopie de celui-ci.

Cataracte

Elle se définit par une opacification d'une partie ou l'ensemble de lentille de l'œil.

Normalement, la lentille est claire et laisse passer la lumière, celle-ci se concentre alors sur la rétine. La cataracte empêche la lumière de passer facilement à travers la lentille, ce qui entraîne une perte de vision.

Cette condition affecte habituellement les personnes âgées. Les symptômes se décrivent par la trouble de la vision ou vision floue, difficulté à voir dans les zones sombres et les lumières vives, les couleurs semblent fanées, double vision, etc.

La cataracte est la principale cause de cécité dans le monde (47,8%) par rapport à d'autres troubles oculaires.

Rétinopathie diabétique

Le diabète affecte les petits vaisseaux sanguins de la rétine. S'ils sont endommagés, il conduit à une détérioration de la vision.

Glaucome

Cette condition est due à une tension élevée dans les yeux. La tension accrue altère la vision en endommageant le nerf optique. Ceci peut être observé chez les adultes plus âgés et chez certains nouveaux nés et qui sont nés avec cette maladie.

Dégénérescence musculaire liée à l'âge

La dégénérescence musculaire liée à l'âge (DMLA) se traduit par une perte progressive de l'acuité visuelle due à l'atteinte de la macula qui est la partie la plus sensible de la rétine.

Le centre du champ visuel paraît flou ou opaque. Le patient est incapable de se concentrer clairement. Cela se produit principalement chez les personnes âgées.

Le risque de la DMLA augmente chez les personnes exposées à un excès de soleil et ceux qui fument trop.

SIDA et déficience visuelle

Cela est généralement causé par des infections virales de l'œil appelé rétinite cytomégalovirus ou CMV.

La proportion estimée de personnes atteintes du SIDA qui seront menés à développer des gammas rétinite CMV est de 20 à 40%.

Cancer des yeux

Le rétinoblastome est le cancer de l'œil le plus commun des enfants. Il y a entre 300 et 400 nouveaux cas diagnostiqués chaque année dans le monde.

1-3 Qui est affecté par la déficience visuelle?

Bien que la déficience visuelle puisse apparaître à tout moment de la vie, elle est plus fréquente chez les personnes âgées. Parfois, la déficience visuelle peut être héréditaire. Elle se manifeste à la naissance ou dans l'enfance. Les conditions courantes correspondent à la rétinite pigmentaire, génétique ou des anomalies du développement, etc. Ces enfants ayant une déficience partielle ou totale ont souvent un retard de développement notamment en motricité globale et en motricité fine. Les adultes ayant une déficience visuelle luttent avec les activités lucratives et les activités quotidiennes. La cécité est connue pour être l'un des handicaps les plus graves qui peut affecter un individu, sa famille et la société.

1-4 Incidence de la déficience visuelle sur le développement

Développement sensoriel

Pour un enfant né aveugle, les autres sens ont leur sensibilité intermittente et peuvent apparaître diminuées. L'enfant reçoit des fragments d'information incohérents, discrétisés et généralement non vérifiés.

L'ouïe est le seul sens à la disposition de l'enfant aveugle, mais l'enfant n'a aucun contrôle sur la présence ou l'absence de son dans son environnement. Le son sans une vérification visuelle n'est que bruit venant de nulle part. Ce n'est qu'après une quantité importante d'expériences tactiles, motrices et d'interactions auditives que le son peut avoir un sens. C'est alors seulement que le son peut fournir des informations sur l'emplacement, la cause ou la source.

Le son n'est pas la forte motivation de cette vision. L'enfant aveugle ne pourra atteindre ce niveau de liaisons entre le son seul et la gestion de l'espace et la localisation qu'au bout de 12 mois. L'exploration de l'environnement est habituellement retardée jusqu'à ce que l'enfant ait atteint ce point. Normalement, l'incitation à l'exploration tactile est fournie par des dimensions visuelles: couleur, motif, forme, emplacement. Ces dimensions ne sont pas disponibles à un enfant aveugle. Par conséquent, l'activité tactile ciblée est minime parce que l'environnement demeure inconnu et peu attirante.

Développement moteur

Mains

Bien que les mains soient un important organe sensoriel, un enfant aveugle a des retards de développement important dans sa capacité à utiliser ses mains fonctionnellement. Même à 5 mois, les mains d'un bébé aveugle seront en position fermés et maintenus à hauteur d'épaule. Il est incapable d'utiliser plusieurs doigts, aucun engagement sur la ligne médiane. À cet âge, un enfant voyant possède la coordination pour atteindre et transférer des objets d'une main à l'autre.

Ce retard dans l'utilisation des mains entraîne une motricité fine retardée et le retard de développement de la motricité globale. Sans vision, les mains et les yeux ne travaillent pas ensemble. Atteindre cette coordination (yeux mains) devrait se produire et cela prend beaucoup d'expérience et est atteinte beaucoup plus tard que la normale coordination yeux mains.

Corps

Un enfant aveugle obtient généralement le contrôle de sa posture approximativement au même âge que les enfants voyants grâce à la progression normale suivante:

- Assis seul momentanément
- Roulement de l'arrière vers l'estomac
- Assis seul régulièrement

- Posture droite debout et progresse par petit mouvement lorsque les mains sont tenues
- Se dresse seul
- Se tient à quatre pattes

Toutefois, les réalisations qui nécessitent une mobilité propre initiative ont beaucoup de retard:

- S'élever à l'aide des bras en position couchée
- Atteindre une position assise
- Tirer un support
- Marcher seul

Un enfant aveugle atteindra la main pour saisir un signal sonore (12 mois), il ne se déplace pas dans l'espace, que ce soit sur à quatre pattes ou en marchant sur les pieds.

La difficulté ou la réticence de l'enfant aveugle à se déplacer autour de l'environnement encourage un comportement passif comme les manières d'autostimulation.

Self Concept

L'enfant aveugle dépend de manière inhabituelle d'une personne voyante afin de l'aider à s'intégrer à l'environnement. Cette notion de dépendance doit être considérée comme un élément majeur dans le développement de l'enfant aveugle.

L'enfant aveugle a diminué de contrôle sur son environnement et il peut contrôler uniquement son monde intérieur. Comme il se retire dans ce monde, il diminue ses besoins d'interaction sociale.

Il ne peut pas comprendre qu'il y a un monde complexe en dehors de lui-même, qu'il est séparé de ce monde, qu'il peut à la fois agir sur celui-ci et devient le bénéficiaire de l'action.

Développement cognitif

Construction de son monde

L'enfant aveugle a une capacité limitée à coordonner et à organiser les éléments dans des niveaux plus élevés d'abstraction et à vérifier l'information. Par conséquent, il construit une réalité qui est différente de l'enfant voyant. Le processus de création de concepts définissant les attributs et les relations est plus problématique pour les aveugles et moins accessibles à l'orientation. L'enfant aveugle est constamment impliqué dans la résolution de problèmes. Mais ce

processus, qui est essentiel pour le développement futur, est plus difficile et moins gratifiant pour lui.

Permanence de l'objet

Un champ visuel stable est la condition d'un objet permanent et d'autres tâches conceptuelles. La permanence de l'objet ne peut pas être obtenue par un enfant aveugle jusqu'à ce qu'il ait la capacité d'atteindre des objets en fonction de son seul repère. Il est acquis presque un an plus tard que chez les enfants voyants.

Relation causale

Comme les résultats des actions ne peuvent pas être vus, l'enfant aveugle ne peut pas être motivé à l'action. Il ne peut pas comprendre sa capacité à provoquer des choses à se produire ou de conserver des agréables motivations.

Constance

Comprendre comment aligner des blocs ou orienter ses mains sur une page afin de dupliquer un modèle sera difficile. S'il n'a pas observé les objets dans diverses orientations, il est difficile de savoir qu'un objet est le même quelle que soit sa position dans l'espace.

Classement

Les possibilités limitées d'explorer les objets et de voir les similitudes sont reflétées dans les erreurs de classification du préscolaire pour un enfant aveugle. La conception du même et du différent peut évoluer que si les enfants identifient la variable distinctive sur laquelle se concentrer. Un enfant aveugle a peu de difficulté à généraliser à travers la taille, mais de nombreuses expériences avec une variété d'objets similaires ont été nécessaires pour accélérer la généralisation et les compétences associées.

Conservation

Un enfant aveugle présente des retards dans la conservation de la substance, poids, volume, longueur et liquide. Par conséquent, son développement social est clairement affecté par son incapacité.

Développement social

Rapports

Pour un enfant voyant, le sourire réciproque entre le nourrisson et sa mère est le début de l'attachement, de reconnaissance et de communication. L'enfant aveugle sourira à 2 mois pour la reconnaissance de la voix de sa mère, mais seulement le chatouillement permettra d'obtenir régulièrement un sourire.

Des années plus tard, l'enfant semblera avoir une implication émotionnelle ambivalente et semble désintéressé, non communicatif, et mal informés sur les rudiments dans le domaine des jeux avec ses pairs. Par conséquent, il peut être évité par ses pairs et rejeté ou surprotégé par des étrangers et des parents.

Dans l'ensemble, ses interactions sociales sont plus compliquées parce que les repères visuels subtils sont absents et les expressions faciales sont perdues.

Auto assistance

De nombreuses compétences d'auto assistance qui sont normalement acquises par l'observation sont retardées chez les enfants aveugles.

Mâcher, ramasser, les compétences d'auto alimentation peuvent être retardés 2 ans ou plus. Se brosser les dents est difficile à réaliser car l'enfant peut rejeter la texture et n'a pas la possibilité d'observer d'autres personnes effectuant les gestes de toilettage.

La peur de l'inconnu et l'incapacité à localiser, la salle de bain peut contribuer à la formation de toilette retardée.

Développement du langage

Imitation

Une grande partie de ce qui est appris par l'enfant sain est appris en imitant les autres.

L'ensemble de la communication, y compris la signature et la dactylologie, est ancrée dans le développement de l'imitation. L'imitation marque le début d'une signification symbolique pour un enfant.

L'enfant aveugle a besoin d'instruction planifiée et systématique dirigée contre le développement de l'imitation différée.

Utilisation d'un langage

L'enfant aveugle peut bafouiller et imiter tôt plus tôt qu'un enfant voyant, mais il peut également montrer un certain retard lors de la combinaison des mots pour exprimer ses besoins.

L'enfant aveugle utilise principalement la langue de s'exprimer ses besoins immédiats ou à décrire les activités courantes.

Il initie quelques questions et son utilisation des adjectifs sont rares.

L'enfant aveugle peut prendre dans les sons qui composent la langue, mais ne peut pas saisir le sens voulu par le locuteur. Ses expériences sensorielles ne sont pas

facilement codées dans le langage. Il peut stocker des phrases et des phrases dans sa mémoire et répétez-les hors de leur contexte.

L'enfant aveugle a souvent un langage qui est écholalie conservateur et sans signification

Le langage précoce de l'enfant aveugle ne semble pas refléter sa connaissance du monde en développement, mais plutôt sa connaissance de la langue des autres.

Pronom personnel

Pour utiliser correctement le pronom personnel «je», un enfant doit avoir établi un sens de lui-même comme distinct de l'environnement. Le développement du concept du soi chez un enfant aveugle est retardé, il a tendance à confondre l'usage des pronoms personnels, d'étendre l'utilisation des deuxième et troisième pronoms personnels ou son propre nom pour faire référence à lui-même.

Chapitre 2

MATLAB

MATLAB est un langage de programmation se compose d'un environnement de développement et un ensemble de modules complémentaires et les bibliothèques qui étendent les fonctionnalités de MATLAB pour le développement de sa propriété importante est sa simplicité de calcul, ce qui permet d'analyser des images comme défi : les réglages en temps réel.

Dans ce chapitre, nous listons les principales caractéristiques d'une description de base de la langue que nous utilisons dans la demande de reconnaissance de visage.

2-1 Qu'est-ce que MATLAB?

Le nom de MATLAB est synonyme de matrice de laboratoire (MATRIX LABORATORY). MATLAB [3] est un langage de haute performance de calcul technique. Il intègre le calcul, la visualisation et la programmation dans un environnement facile à utiliser où les problèmes et les solutions sont exprimés dans une notation mathématique familière.

L'utilisation classique intègre :

- Les mathématiques et le calcul
- Le développement d'algorithmes
- La modélisation, la simulation et le prototypage
- L'analyse de données, l'exploration et la visualisation
- Le développement d'applications y compris le renforcement de l'interface graphique de l'utilisateur

MATLAB est un système interactif dont l'élément de base de données est un tableau qui ne nécessite pas de dimensionnement. Ceci permet de résoudre de nombreux problèmes, en particulier ceux avec des formulations matricielles et vectorielles. Ces problèmes sont résolus rapidement, au temps de temps qu'il faudrait pour écrire un programme dans un langage non interactif comme le langage C ou le Fortran.

MATLAB est constitué d'un environnement de développement, d'un langage de programmation, de modules et de bibliothèques qui étendent les fonctionnalités de

MATLAB pour les industries ou pour résoudre des problèmes. Le développement de l'environnement possède à la fois une ligne de commande et une interface graphique, ainsi que des navigateurs pour l'affichage de fichiers et un débogueur qui permet de tester le code entré en langage MATLAB. MATLAB est un langage de programmation orienté objet qui peut être utilisé pour écrire du code qui a pour but d'être utilisé dans l'environnement MATLAB.

Les bibliothèques et les modules intégrés varient mais peuvent contenir n'importe quelles équations communément utilisées pour des fonctions complexes qui compilent des données suivant un paramétrage prédéfini.

MATLAB possède son propre éditeur intégré :

- Soulignage syntaxique
- Mise en évidence d'erreurs
- Numérotation de lignes
- Auto formatage

2-2 Le système MATLAB

Le système MATLAB est constitué de cinq principales parties [4]:

Le langage MATLAB :

Il s'agit d'une matrice/gamme de langage de haut niveau avec des instructions de contrôle de flux, de fonctions, de structures de données, d'entrées/sorties et de fonctions de programmation orientée objet. Il permet à la fois une programmation « dans le petit » dans le but de créer rapidement un programme « jetable » et également de créer une programmation dite « à grande échelle » qui permet de créer des programmes complets et d'applications complexes.

L'environnement de travail MATLAB :

Il s'agit de l'ensemble des outils et des fonctionnalités qui sont utilisés par les utilisateurs ou les programmeurs MATLAB. Il comprend des fonctionnalités de gestion des variables dans l'espace de travail et permet d'importer et d'exporter des données. Il inclut également des outils pour le développement, la gestion, le débogage et le profilage des applications MATLAB.

Manipulation des graphiques :

Il s'agit du système de graphes des MATLAB. Il inclut un haut niveau de commande de visualisation des données en deux et trois dimensions. Il permet également le traitement de l'image, l'animation et des graphiques de présentation. Il inclut des commandes de niveau bas qui permettent de personnaliser complètement l'apparence des graphiques ainsi que de construire des interfaces graphiques complètes (GUIs) pour les applications MATLAB.

Librairie MATLAB de fonctions mathématiques :

Il s'agit d'une vaste collection d'algorithmes de calcul allant des fonctions élémentaires comme les fonctions somme, sinus ou cosinus et l'arithmétique complexe à des fonctions plus sophistiquées comme des matrices inverses, des matrices à valeur propre, les fonctions de Bessel ou encore des transformées de Fourier rapides.

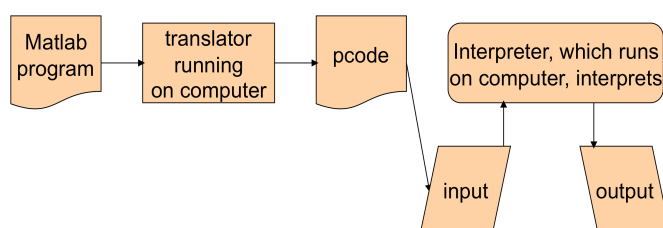
L'interface de programmation d'application MATLAB (API) :

Il s'agit d'une bibliothèque qui permet d'écrire des programmes en C et en Fortran qui interagissent avec MATLAB. Il inclut des fonctionnalités qui appellent des routines depuis Matlab (lien dynamique), appelant MATLAB comme un moteur de calcul et pour lire et écrire des MAT Files.

2-3 Interprétation VS compilation

En informatique, interpréter signifie exécuter alors que compiler signifie traduire dans le langage de l'ordinateur (aussi connu sous le nom de « machine ») :

- MATLAB est un langage interprété (comme Java)
- C++ est un exemple de l'un des langages compilés



Les variables déclarées VS les variables non déclarées

- Le langage compilé exige que chaque variable soit « à déclarer » avant qu'elle ne soit utilisée
- MATLAB (langage interprété) n'a pas besoin de déclaration

Les raisons des deux approches :

- Compiler met l'accent sur la vitesse du programme
 - La déclaration des variables donne au compilateur plus d'information pour l'optimisation
 - Le code source est traduit dans un code machine « natif » (correspond au langage de l'ordinateur)
- Interpréter souligne la vitesse de programmation
 - L'absence de déclaration de variable simplifie la programmation
 - L'interprète peut donner des informations de débogage détaillées

2-4 Pourquoi choisir MATLAB?

- Son élément des données de base est la matrice. Un simple entier est considéré comme une matrice d'une ligne et une colonne. Plusieurs opérations mathématiques qui travaillent sur des tableaux ou des matrices sont intégrées à l'environnement MATLAB. Par exemple, des produits croisés, point produits, les déterminants, les matrices inverses.
- Il n'y a pas vraiment de langages de programmation standards dans le domaine scientifique, mais MATLAB s'en rapproche.
- Un langage interprété (qui correspond au non compilé) Une bon langage interprété pour le travail numérique et la visualisation. Matlab utilise des fonctions précompilées et un langage basé sur les vecteurs pour faire un travail numérique et de visualisation sans avoir besoin d'être un langage compilé
- Les opérations vectorielles. L'ajout de deux tableaux en même temps ne nécessite que d'une seule commande, au lieu d'une boucle «for» ou «while».
- La sortie graphique est optimisée pour l'interaction. Il est possible de tracer des données très facilement, et ensuite changer les couleurs, tailles, échelles, etc., en utilisant les outils graphiques interactifs.

- La fonctionnalité Matlab peut être considérablement élargie par l'ajout de boîtes à outils. Ce sont des ensembles de fonctions spécifiques qui ont fourni des fonctionnalités plus spécialisées. Ex: lien Excel permet aux données d'être écrites dans un format reconnu par Excel, Statistiques Toolbox permet la manipulation statistique plus spécialisée de données.
- De nombreuses façons de résoudre des problèmes numériques
- Il a une énorme bibliothèque de fonctions, le travail numérique et la visualisation MATLAB sont par conséquent agrandis.

2-5 Faiblesses de MATLAB

- Il utilise une grande quantité de mémoire et sur les ordinateurs lents, il est très difficile à utiliser.
- Le langage interprété (et non compilé), le code interprété peut être beaucoup plus lent Beaucoup de fonctions et boîtes à outils pour qui influent sur la lenteur
- Il se situe "au dessus" de Windows, MATLAB obtient autant de temps CPU que Windows lui permet d'avoir. Cela rend les applications en temps réel très compliquées.
- Coût - chaque boîte à outils est une charge supplémentaire
- L'inconvénient est que Matlab perd beaucoup de son efficacité lorsque vous n'utilisez pas les fonctions précompilées, ou si vous travailler avec des données qui ne peuvent pas être mis dans des vecteurs et des matrices ou implique texte analyse

2-6 Mise en route MATLAB

Fenêtre principale de MATLAB

Le MATLAB IDE a 3 sous les parties principales:

- Répertoire
- Historique des commandes
- Commande Windows

Il a également:

- Une barre de menus
- Une barre d'outils

- Des raccourcis

Aide dans MATLAB

MATLAB fournit quelques moyens pour obtenir de l'aide ou plusieurs informations sur toutes ses fonctions et les services.

A partir de la ligne de commande,

- Typer Aide ou Aide Fn Name pour le texte d'aide dans la fenêtre de commande.
- Typer doc ou doc Fn Name pour un document détaillé dans le navigateur d'aide.
- Typer look for keyword pour rechercher des fonctions clé.
- Dans le menu principal, sélectionner le sous-menu Aide
- Autres options: une documentation complète, exemples, démos

Programmation MATLAB

- La fenêtre de commande est la principale zone de saisie et d'exécution des commandes.
- Utiliser les rédacteurs de MATLAB (et d'autres outils / GUI) pour entrer les commandes, créer des fonctions, etc.

Saisie des commandes

a) L'exécution de commande de MATLAB est : >>

b) À cette exécution, vous pouvez entrer des commandes tels que:

- Créer ou modifier une variable: par exemple $A = 3$
- Effectuer un calcul: par exemple $1 + 2$
- Appeler une fonction: par exemple $\max([1 \ 2 \ 3])$
- Obtenir de l'aide: par exemple `help`, `help max`
- Gérer l'espace de travail: par exemple `Qui`
- Enregistrer les variables de l'espace de travail: par exemple sauvegarder un fichier
A

Scalaires, variables et arithmétiques de base

- Les variables sont déclarées à tout moment dans votre code.

Il suffit d'utiliser l'opérateur = par exemple $A = 3$

- MATLAB n'a aucune notion de types de données - une variable scalaire peut être à un moment et un tableau ou une structure à un autre instant. Cela peut être ce qui est recherché, dans tous les cas une prudence est de mise.

- MATLAB est utilisé, comme tout autre langage, pour effectuer l'arithmétique de base
- MATLAB peut effectuer des opérations arithmétiques directement à la ligne de commande : par exemple $1 + 2$
- Les séries sont déclarées à l'aide d'apostrophe ' : par exemple $S = 's'$

Retour de résultats de calcul

- MATLAB peut retourner un calcul à une variable:

```
>> B = A + 3
```

```
B =
```

```
6
```

- MATLAB peut effectuer des opérations arithmétiques directement à la commande:
par exemple $1 + 2$

```
>> 1 + 2
```

```
ans =
```

```
3
```

Aucune variable permanente n'est affectée, mais les résultats des variables temporaires sont retournés. (Il s'agit d'une notion de base importante de MATLAB)

Les commandes terminées par un point virgule

- Le point-virgule termine généralement les énoncés dans MATLAB. Strictement parlant, ceci correspond aux retours à la ligne ou le début d'une nouvelle ligne
- Si le point virgule est omis, le résultat du calcul fait écho à la fenêtre de commande:
 - Idéal pour le développement et le débogage de code
 - ennuyeux et fastidieux dans les grandes étendues de code

```
>> 6+5
```

```
ans =
```

```
11
```

```
>> 6+5;
```

```
>> ans;
```

```
>> ans
```

```
ans =
```

```
11
```

```
>> A = 3
```

```
A =  
3  
>> A = 3;  
>> B = A + 3;  
>> B  
B =  
6
```

Matrices ou tableaux

MATLAB travaille essentiellement avec un seul type d'objet - matrice numérique rectangulaire avec des entrées éventuellement complexes:

- Toutes les variables représentent des matrices.
- Dans certaines situations, les matrices 1x1 sont interprétées comme des scalaires et les matrices avec une seule ligne ou une colonne sont interprétées comme des vecteurs.

La saisie de matrices

Les matrices peuvent être introduites dans MATLAB de plusieurs façons différentes:

- Entrer par une liste explicite d'éléments,
- Générer par des déclarations et des fonctions intégrées,
- Création d'un texte file avec un éditeur local,
- Charger des fichiers ou des applications de données externes

Par exemple,

- Créer une matrice 3x3 et l'affecter à une variable A.

Chacune des déclarations

```
A = [1 2 3; 4 5 6; 7 8 9]
```

And

```
A = [  
1 2 3  
4 5 6  
7 8 9 ]
```

- Les éléments d'une rangée d'une matrice peuvent être séparés par des virgules, ainsi qu'un espace.

A = [1,2,3; 4,5,6; 7,8,9]

Saisie des matrices plus grande

- optimisé dans un ASCII le avec un éditeur local, où les erreurs peuvent être facilement corrigées
- Le fichier doit être composé d'un tableau rectangulaire contenant uniquement les entrées de la matrice numériques. Si cette _le est nommé, disons, data.ext (où. ext est une extension de l'exception. Mat),
- La commande MATLAB data.ext de charge de lire ce _le pour les données variables dans l'espace de travail MATLAB. Par exemple :

```
>> Load A.dat
```

```
>> A
```

```
A =
```

```
1 2 3
```

```
4 5 6
```

```
7 8 9
```

```
A.dat:
```

```
1 2 3
```

```
4 5 6
```

```
7 8 9
```

Fonctions intégrées de création de matrice

Voici des exemples de fonctions qui créent des matrices simples:

- rand (n) va créer un n _ n matrice dont les entrées sont générées aléatoirement répartis uniformément entre 0 et 1, tandis que rand (m, n) va créer un m _ n de un.
- magic (n) va créer une matrice de nxn intégrale qui est un carré magique (lignes, colonnes et diagonales avoir somme commun).
- eye (n), eye(m, n) vont créer un carré n ou un m X n matrice d'identité.
- ones (n), ones(m, n) vont créer un carré ou n un m X n matrice de uns
- zéros (n), zéros (m, n) vont créer un carré ou n un m X n matrice de zéros.

Les données multimédia comme vecteurs, des matrices, tableaux, etc

Les données multimédia de base sont simplement représentées comme une matrice ou un tableau dans MATLAB:

Audio → tableau ou vecteur 1-D des amplitudes d'onde sonore

Image → tableau ou matrice 2-D des intensités de couleurs

Vidéo → tableau ou matrice 3-D : Chaque cadre d'une image 2-D, n images par seconde

Données audio en tant que vecteur

Lecture audio: MATLAB peut nativement lire des fichiers WAV et AU _les via wavread () et auread (). Par exemple

```
[y, Fs] = wavread ('handel.wav');
```

Un vecteur y est retourné avec la fréquence d'échantillonnage Fs du fichier d'entrée.

Traitement : Maintenant, nous avons l'audio dans un tableau ou un vecteur 1-D pour le traitement

Afficher la forme d'onde : une simple courbe (y) affiche les données audio.

Lancer le son : Un lecteur audio et Play (), le son est joué. Alternativement, vous pouvez utiliser le son (y, Fs) ou soundsc (y, Fs) pour jouer le son.

Exporter les données dans une fenêtre de commande :

y stocke les données comme un long tableau MATLAB 1-D. Donc, nous pouvons tout simplement taper:

```
>> y regarder pour les chiffres.  
>> [y,Fs] = wavread('handel.wav'); % read in wav file  
>> plot(y); % plot in figure  
>> p = audioplayer(y, Fs); % create audioplayer  
>> play(p, [1 n]); % play audio  
>> y % list elements  
ans =  
0  
-0.0062  
-0.0750  
-0.0312  
0.0062  
0.0381  
0.0189
```


-0.0250

-0.0312

Images comme matrices

les images (non compressé) sont représentés comme une grille de pixels (intensités pour les images en niveaux de gris).

Lecture / Ecriture des images dans MATLAB

- `imread ('filename')` lit un _le et l'assigne à une variable de tableau:
- `im= imread ('parrots.jpg');`
- MATLAB peut lire de nombreux formats de fichiers, y compris: JPEG, GIFF, TIFF, BMP, PNG
- image de couleurs dispose de 3 images plans (RVB), se fait un tableau 3-D dans MATLAB.
- `size ()` est utile pour obtenir les dimensions de l'image:
`[l m n] = size (im);`
- `who(var)` → énumère une variable donnée sous un long format.
- `imwrite ()` écrire une image sur un format de fichier spécial.
- Les images (et tous les graphiques) peuvent également être enregistrés directement depuis la fenêtre de la figure MATLAB.
- `imread ('filename')` lit un fichier et l'affecte à une variable de tableau:
`mshow (im);`

Vidéos comme matrices dans MATLAB

- MATLAB a une structure de film: il s'agit d'un tableau 2D de trames d'image dans le temps (cadence)
- MATLAB peut lire nativement les fichiers vidéo AVI avec :
`aviread («avi»)`il retourne un film
- `movie(tableau de film)` jouera la vidéo dans une figure.
- Comme pour les images et les audio, `size()` et qui sont des commandes utiles

Traitement vidéo avec mmreader

La commande «`mmreader`» est un outil plus puissant pour travailler avec les vidéo dans MATLAB. Il supporte plus de formats, autres que AVI (. Avi), par exemple

MPEG-1 (. Mpg), Windows Media Video (. Wmv,. Asf,. ASX) sous Windows, et MPEG-1 (. Mpg), MPEG-4 (. Mp4,. M4v), QuickTime (. Mov) sur Mac.

Un exemple : `mmr = mmreader('origdave.avi');`

 Create an mmreader object.

`vidFrames = read(mmr);`

Lit dans tous les cadres. Pour la vidéo en couleurs, vidFrames est un tableau à quatre dimensions: largeur x hauteur x chaines(3) x numberOfFrames
`numFrames = get(mmr, 'numberOfFrames');`

Pour obtenir un nombre de cadres

for k = 1:numFrames

`mov(k).cdata = vidFrames(:, :, :, k);`

`mov(k).colormap = [];`

end

Pour construire une structure de film pour contenir les données vidéo movie (mov, 1, `get(mmr, 'frameRate');`);

Lancer la vidéo

Indexation des matrices / tableaux

- des données d'entrée sous forme de matrice et de vecteurs peuvent être référencées avec des indices à l'intérieur de parenthèses.

Par exemple

A (2,3) représente l'entrée dans la deuxième rangée, la troisième colonne de la matrice A

Un autre exemple

x (3) représente la troisième coordonnée du vecteur x.

- Une matrice ou un vecteur n'acceptent que des nombres entiers positifs en indices

```
>> A = [1 2 3; 4 5 6; 7 8 9]
```

```
A =
```

```
1 2 3
```

```
4 5 6
```

```
7 8 9
```

```
>> A(2,3)
```

```
ans =
```

6

```
>> X = [1 2 3 4 5 6 7 8]
```

X =

```
1 2 3 4 5 6 7 8
```

```
>> X(3)
```

ans =

```
3
```

Opérations matricielles

Les opérations matricielles suivantes sont disponibles dans MATLAB:

Addition

Soustraction

Multiplication

Puissance

transposée conjuguée

transposée

division à gauche

division à droite

Ces opérations matricielles s'appliquent aux scalaires (matrices 1x1)

Règles de l'opérateur de Matrice

Les opérateurs de Matrice doivent obéir à des lois mathématiques de base:

- Si les dimensions des matrices sont incompatibles pour l'opération matricielle, un message d'erreur se produit:

Pour l'addition et la soustraction des matrices doivent avoir la même dimension.

La multiplication doit obéir à la règle de produit matriciel.

- Exception: le cas des opérations scalaire matrice (addition, soustraction et division, ainsi que pour la multiplication), auquel chaque entrée de la matrice opère par le scalaire.

2-7 Figures, courbes et graphiques

Types de courbes MATLAB

Fonctions graphiques en deux dimensions

- Graphiques linéaires
- graphiques à barres
- Graphiques de surface
- Graphiques de direction
- Triangulation radiale
- Diagramme de dispersion

Fonctions graphiques en trois dimensions

- Graphiques linéaires
- Graphiques de mailles et graphiques à barres
- Graphiques de la région et des objets constructifs
- Graphiques de surface
- Graphiques de direction
- Graphiques volumétriques

Qu'est-ce qu'un graphique MATLAB?

L'environnement MATLAB offre une variété de fonctions de données ainsi qu'un ensemble d'outils d'interface graphique pour créer et modifier des affichages graphiques de tracé. Les outils graphiques offrent plus de contrôle sur les propriétés graphiques et des options que les commandes tapées fournissent tels qu'annoter, obtenir, et définir.

Une figure est une fenêtre MATLAB contenant des affichages graphiques (généralement des courbes données) et des composants d'interface utilisateur. Vous créez des courbes explicitement avec la fonction de figure, et, implicitement, chaque fois que vous tracez un graphiques et aucun chiffre n'est actif. Par défaut, les fenêtres de figure sont redimensionnables et comprennent des menus déroulants et les barres d'outils.

Une courbe est un affichage graphique, vous pouvez créer dans une fenêtre de la figure. Les courbes peuvent afficher des données tabulaires, des objets géométriques, des objets de surface et de l'image, et les annotations tels que les titres, légendes et les barres de couleurs. Les chiffres peuvent contenir n'importe quel nombre de courbes. Chaque courbe est créé dans un 2-D ou un espace de données 3-D appelé axes. Vous pouvez créer explicitement des axes avec la fonction axes ou la fonction courbes.

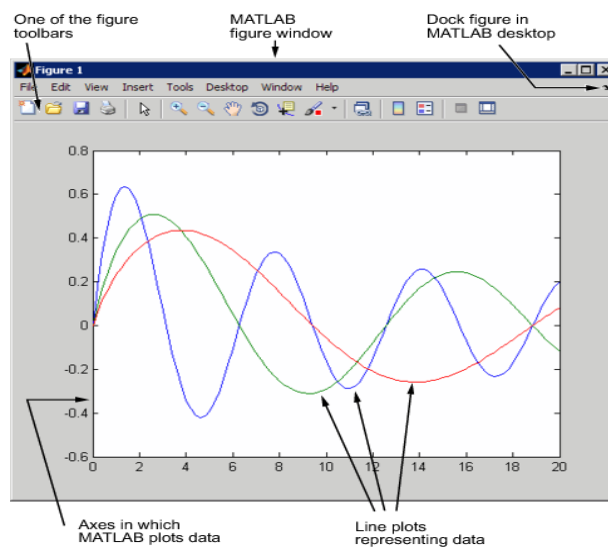
Un graphe est une représentation graphique des données au sein d'axes 2-D ou 3-D. La plupart des courbes réalisées avec MATLAB fonctions et interfaces graphiques sont donc graphiques. Si vous représentez une variable unidimensionnelle (par exemple, `rand(100,1)`), les indices du vecteur de données (dans ce cas, 1:100) deviennent assignés comme valeurs x, et les courbes le vecteur de données en tant que valeurs de y. Certains types de graphiques peuvent afficher plus d'une variable à la fois, d'autres non.

Anatomie d'un graphique

Les fonctions graphiques MATLAB et les outils orientent leur production vers une fenêtre de figure. Chaque figure est une fenêtre séparée que vous pouvez ancrer dans le bureau, et rassembler avec d'autres courbes dans un groupe de figure. Pour illustrer les éléments de base d'un graphique, exécutez le code [5] suivant pour créer une courbe d'une famille de courbes sinusoïdales :

```
x = 0:.2:20;
y = sin(x)./sqrt(x+1);
y(2,:) = sin(x/2)./sqrt(x+1);
y(3,:) = sin(x/3)./sqrt(x+1);
plot(x,y)
```

Le chiffre obtenu contient un ensemble d'axes 2-D. Ce graphique identifie les composants et des outils d'une fenêtre de la figure :



La fonction de traçage utilise un style de ligne par défaut et une couleur par défaut pour distinguer les ensembles de données tracées dans le graphique. Vous pouvez

modifier l'apparence de ces composants graphiques ou ajouter des annotations au graphe pour présenter vos données d'une façon particulière.

Vue d'ensemble du traçage MATLAB:

Processus de traçage

MATLAB offre une grande variété de techniques pour afficher graphiquement les données.

Des outils interactifs vous permettent de manipuler des graphiques pour obtenir des résultats qui révèlent le plus d'informations à propos de vos données. Vous pouvez également annoter et imprimer des graphiques pour des présentations, ou exporter des graphiques aux formats graphiques standard pour la présentation dans les navigateurs Web ou d'autres médias.

Le procédé de visualisation de données implique généralement une série d'opérations.

Création d'un graphique

Le type de graphe que vous choisirez de créer dépend de la nature de vos données et ce que vous souhaitez révéler sur les données.

Il y a deux façons de base pour créer des graphiques dans MATLAB:

- Utiliser des outils de traçage pour créer des graphiques de manière interactive.
- Utilisez l'interface de commande pour entrer des commandes dans la fenêtre de commande ou créer des programmes de traçage.

Il pourrait être très utile de combiner les deux approches.

Exploration des données

Une fois un graphique créé, vous pouvez extraire des informations spécifiques sur les données, telles que la valeur numérique d'un pic dans une courbe, la valeur moyenne d'une série de données, ou vous pouvez effectuer raccord données.

Modification des composants du graphe

Les graphiques sont composées d'objets, qui ont des propriétés que vous pouvez modifier. Ces propriétés affectent la manière dont les différentes composantes du graphique apparaissent et se comportent.

Par exemple, les axes permettant de définir le système de coordonnées du graphe ont des propriétés qui définissent les limites de chaque axe, l'échelle, la couleur, etc. Notez que les données utilisées pour créer une ligne graphique sont la propriété de la ligne. Il est par conséquent, possible de modifier les données sans réellement créer un nouveau graphique.

Annotation Graphiques

Les annotations sont le texte, des flèches, des légendes, et d'autres étiquettes ajoutées aux graphiques pour aider les utilisateurs à voir ce qui est important concernant les données.

Impression et exportation Graphiques

L'aperçu avant impression vous permet d'ajouter des en-têtes et pieds de page, une date, et ainsi de suite. La boîte de dialogue d'aperçu avant impression vous permet de contrôler la taille, la disposition et les autres caractéristiques du graphique (sélectionnez Aperçu avant impression dans le menu Fichier de figure).

L'exportation d'un graphe signifie créer une copie de celui-ci dans un format graphique standard de fichiers, tels que TIFF, JPEG ou EPS. Vous pouvez ensuite importer le fichier dans un traitement de texte, l'inclure dans un document HTML, ou le modifier dans un logiciel de dessin (sélectionner Setup Exporter dans le menu Fichier de figure).

Ajout et suppression de contenu d'une figure

Par défaut, lorsque vous créez un nouveau graphique dans la même fenêtre de figure, les données se substituent à celle du graphique qui est actuellement affichée, le cas échéant. Vous pouvez ajouter de nouvelles données à un graphique de plusieurs façons :

- Vous pouvez supprimer manuellement toutes les données, des graphiques et des annotations du chiffre actuel en tapant CLF dans la fenêtre de commande
- en sélectionnant Effacer figure dans le menu Édition de la figure

Enregistrement de graphiques qui seront rechargés dans MATLAB

Il y a deux façons de garder des graphiques qui vous permettent d'économiser le travail que vous aviez investi dans leur préparation :

- Enregistrer le graphique comme une figure de fichier (sélectionnez Enregistrer dans le menu Fichier de figure).
- Générer un code MATLAB qui peut recréer le graphique (sélectionnez Générer M-fichier dans le menu Fichier de figure).

Fig-Files.

Les Fig-fichiers sont dans un format binaire qui enregistre un chiffre dans son état actuel. Cela signifie que tous les objets graphiques et les paramètres de propriété sont stockés dans le fichier lors de sa création. Vous pouvez recharger le fichier dans une session MATLAB différente, même si vous utilisez MATLAB sur un autre type d'ordinateur. Lorsque vous chargez un fichier figure, MATLAB crée une nouvelle figure dans le même état que celui que vous avez enregistré. Notez que les états de tous les outils de figure (ie, tous les éléments des barres d'outils) ne sont pas enregistrés dans un fichier figure, seul le contenu du graphique est enregistré.

Code généré.

Vous pouvez utiliser le générateur de code de MATLAB M-code pour créer du code qui recrée le graphique. Contrairement à un FIG-File, le code généré ne contient pas de données. Vous devez passer des données appropriées à la fonction générée lorsque vous exécutez le code.

Composants de Graphique

MATLAB affiche des graphiques dans une fenêtre spéciale. Pour créer un graphique, vous devez définir un système de coordonnées. Par conséquent, chaque graphique est placé avec ses axes, qui sont contenues dans la figure.

La représentation visuelle réelle des données est réalisée avec des objets graphiques comme des lignes et des surfaces. Ces objets sont attirés dans le système de coordonnées défini par les axes, qui est automatiquement créé par

MATLAB spécifiquement pour accueillir la plage des données. Les données réelles sont stockées comme des propriétés des objets graphiques.

Outils des graphes

Le graphe est équipé d'ensembles d'outils qui fonctionnent sur les graphes. Le menu des outils de graphe permet d'accéder à de nombreux outils du graphe.

Accès aux outils

Vous pouvez consulter ou supprimer les barres d'outils et les outils de traçage depuis le menu Affichage, comme le montre l'image suivante. Activez ou désactivez les barres d'outils vous avez besoin. Ajout d'une barre d'outils juste au-dessous de la plus basse.

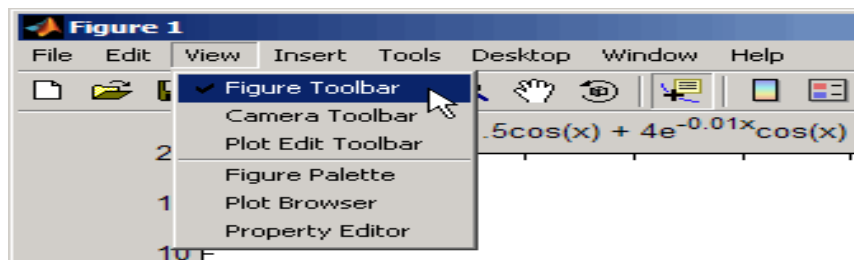
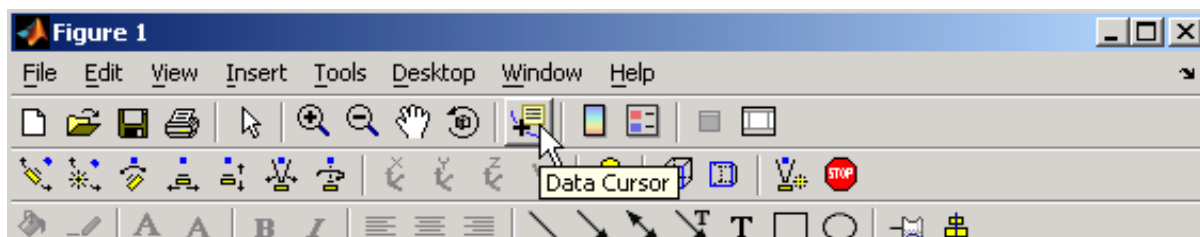


Figure Barres d'outils

Les barres d'outils de Figure permettent d'accéder facilement à de nombreuses fonctions de modification de graphiques.

Il existe trois barres d'outils. Lorsque vous placez le curseur sur un outil particulier, une zone de texte apparaît avec le nom de l'outil. L'image suivante montre les trois barres d'outils s'affiche avec le curseur sur l'outil curseur de données.

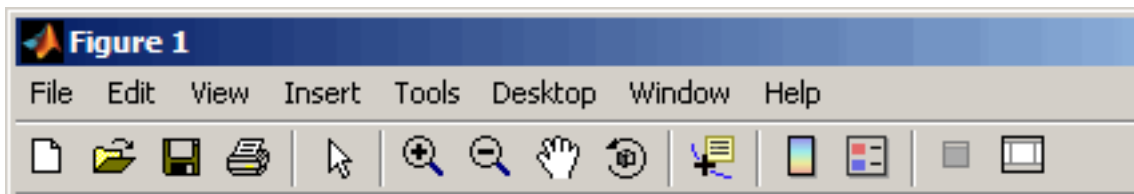


Outils de traçage

Les outils de traçage sont attachés aux figures et créer un environnement propice à la création de graphiques. Ces outils vous permettent d'effectuer les opérations

suivantes :

- Choisir parmi une grande variété de types de graphes.
- Modifier le type de graphique qui représente une variable.
- Voir et définir les propriétés des objets graphiques.
- Annoter des graphiques avec du texte, des flèches, etc.
- Créer et organiser des sous figure secondaires dans la figure.
- Faire glisser et déposez les données sous forme de graphiques



Il est également possible de démarrer les outils de traçage de la commande MATLAB, en utilisant la commande :

Plot Tools

Les outils de traçage sont constitués de trois composantes GUI indépendantes:

- Palette de Figure (figurepalette): Définir et organiser des sous figures secondaires, accéder à des variables d'espace de travail pour le tracé ou l'édition, et ajouter des annotations.
- Navigateur de figure (plotbrowser) : Sélectionner les objets dans la hiérarchie du graphe, la visibilité de contrôle, et ajouter des données aux axes.
- Éditeur de propriétés (PropertyEditor) : Modification des propriétés clés de l'objet sélectionné. Cliquez sur Plus de Propriétés pour accéder à toutes les propriétés de l'objet avec l'inspecteur des propriétés.

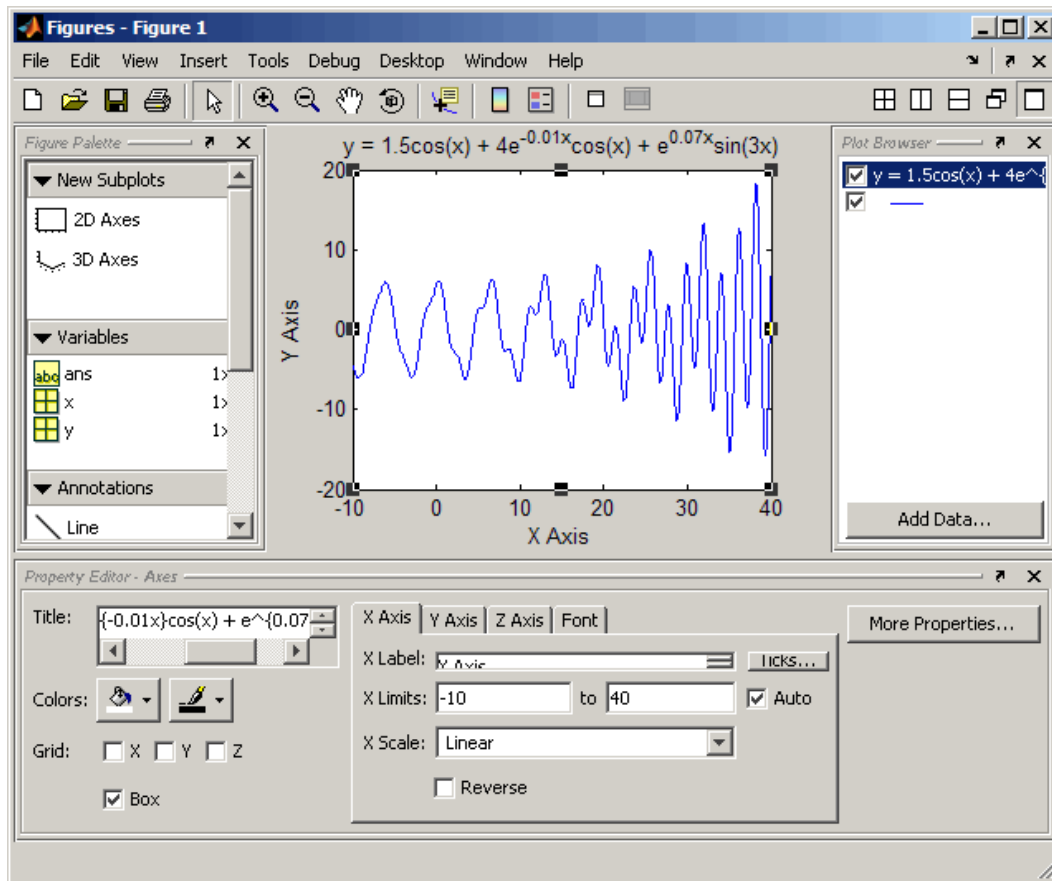
Vous pouvez également contrôler ces composants à partir de la fenêtre de commande MATLAB, en tapant la commande suivante :

Figurepalette

Plotbrowser

PropertyEditor

L'image suivante montre une figure avec les trois outils de traçage permis



2-8 Traitement de l'image

2-8.1 Qu'est-ce que l'image numérique?

Une image peut être définie comme une fonction à deux dimensions, $f(x, y)$, où x et y sont des coordonnées spatiales, et l'amplitude de f à tout couple de coordonnées (x, y) est appelée le niveau d'intensité ou de gris de l'image à ce point. Lorsque x , y et les valeurs d'amplitude de f sont toutes les quantités finies, discrets, nous appelons l'image une image numérique. Le domaine du traitement numérique de l'image se réfère au traitement d'images numériques au moyen d'un ordinateur numérique. A noter que l'image numérique se compose d'un nombre fini d'éléments, dont chacun possède un emplacement et une valeur particulière. Ces éléments sont considérés comme des éléments d'image, les éléments de l'image, Pels et de pixels. Pixel est le terme le plus largement utilisé pour désigner les éléments d'une image numérique.

2-8.2 Traiter les images numériques avec des algorithmes informatiques

Le traitement de l'image numérique est l'utilisation d'algorithmes informatiques pour créer, traiter, transmettre et afficher des images numériques. Les algorithmes de traitement d'images numériques peuvent être utilisés pour:

convertir les signaux provenant d'un capteur d'image en images numériques

- Améliorer la clarté, supprimer le bruit et d'autres artefacts
- Extraire la taille, l'échelle ou le nombre d'objets dans une scène
- Préparer les images pour l'affichage ou l'impression
- Compresser les images pour la communication à travers un réseau

Vous pouvez effectuer un traitement numérique de l'image dans MATLAB avec Image Processing Toolbox, qui fournit des algorithmes numériques de traitement d'image, des outils et un environnement complet d'analyse des données, la visualisation et le développement d'algorithmes

2-8.3 Formats d'image supportés par Matlab

Les formats d'image suivants sont pris en charge par Matlab:

- BMP
- HDF
- JPEG
- PCX
- TIFF
- XWB

2-8.4 Aspects du traitement de l'image

- Amélioration de l'image: Traitement de l'image de sorte que le résultat est plus approprié pour une application particulière. (définition des contours ou à une déconvolution la mise au point d'une image, souligner les bords, améliorer le contraste de l'image, ou raviver une image, suppression du bruit).
- Restauration de l'image: Cela peut être considéré comme inverser les dommages causés à l'image par une cause connue. (Suppression de flou causé par le mouvement linéaire, la suppression des optiques de distorsions).

- **Segmentation de l'image:** Il s'agit de diviser une image en plusieurs parties constitutives, ou d'isoler certains aspects d'une image. (Trouver des lignes, des cercles ou des formes particulières dans une image, à une photographie aérienne, l'identification des voitures, des arbres, des bâtiments ou des routes.

2-8.5 Types d'images numériques

- **Binaire:** chaque pixel est juste noir ou blanc. Comme il n'y a que deux valeurs possibles pour chaque pixel (0,1), nous avons seulement besoin d'un bit par pixel.

- **Grayscale:** chaque pixel est une nuance de gris, normalement de 0 (noir) à 255 (blanc). Cette gamme signifie que chaque pixel peut être représentée par huit bits, soit exactement un octet. D'autres gammes en niveaux de gris sont utilisées, mais ils sont généralement une puissance de 2.

- **Vrai couleur ou RVB:**

C'est un autre format pour les images de couleur. Il représente une image avec trois matrices de dimensions correspondant au format d'image. Chaque matrice correspond à l'une des couleurs rouge, verte ou bleue, et donne une instruction de la quantité de chacune de ces couleurs un certain pixel devrait utiliser, comme une image est une "pile" de trois matrices; représentant le rouge, le vert et le bleu des valeurs pour chaque pixel. Cela signifie que pour chaque pixel correspond à 3 valeurs.

- **Image indexée**

C'est une façon pratique de représenter des images en couleur. Une image indexée enregistre une image sous forme de deux matrices. La première matrice a la même taille que l'image et un nombre pour chaque pixel. La deuxième matrice est appelé le plan de la couleur et de sa taille peut être différente de l'image. Les chiffres de la première matrice est une instruction de ce numéro à utiliser dans la matrice de panel de couleur.

- **Multi frame images**

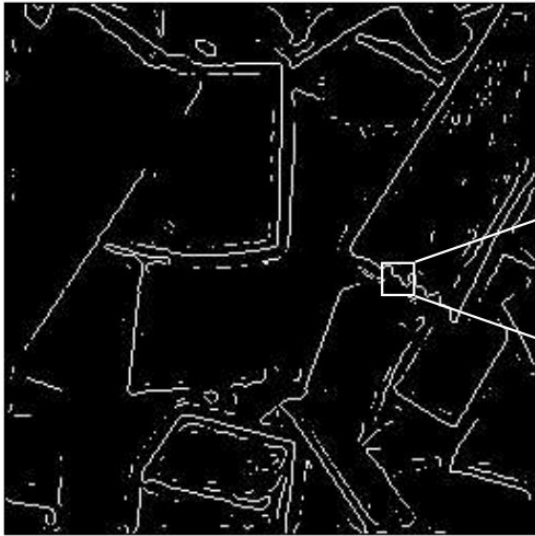
Dans certaines applications, nous voulons étudier une séquence d'images. Ceci est très fréquent dans l'imagerie biologique et médicale où vous pourriez étudier une séquence de tranches d'une cellule. Pour ces cas, le format multi image est un moyen commode de travailler avec une séquence d'images.

2-8.6 Comment convertir entre différents formats

Le tableau suivant montre comment convertir entre les différents formats indiqués ci-dessus. Toutes ces commandes nécessitent la boîte à outils de traitement d'image. Conversion de format d'image :(dans les parenthèses) vous tapez le nom de l'image que vous souhaitez convertir.

Opération	Commande
Convertir du format l'intensité/l'index/ le format RVB au format binaire.	dither()
Convertir du format intensité au format indexe.	gray2ind()
Convertir du format indexé au format intensité.	ind2gray()
Convertir du format indexé au format RVB.	ind2rgb()
Convertir d'une matrice régulière au format innéiste via une mise à l'échelle.	Mat2gray()
Convertir d'un format RVB au format intensité.	Rgb2gray()
Convertir d'un format RVB au format indexé.	Rgb2ind()

Image Binaire



1	1	0	0	0	0
0	0	1	0	0	0
0	0	1	0	0	0
0	0	0	1	0	0
0	0	0	1	1	0
0	0	0	0	0	1

Image avec niveaux de gris (grayscale)



230	229	232	234	235	232	148
237	236	236	234	233	234	152
255	255	255	251	230	236	161
99	90	67	37	94	247	130
222	152	255	129	129	246	132
154	199	255	150	189	241	147
216	132	162	163	170	239	122

Image colorée



49	55	56	57	52	53	64	76	82	79	78	78	66	80	77	80	87	77
58	60	60	58	55	57	93	93	91	91	86	86	81	93	96	99	86	85
58	58	54	53	55	56	88	82	88	90	88	89	83	83	91	94	92	88
83	78	72	69	68	69	125	119	113	108	111	110	135	128	126	112	107	106
88	91	91	84	83	82	137	136	132	128	126	120	141	129	129	117	115	101
69	76	83	78	76	75	105	108	114	114	118	113	95	99	109	108	112	109
61	69	73	78	76	76	96	103	112	108	111	107	84	93	107	101	105	102
Red						Green						Blue					

2-8.7 Commandes générales utilisées dans le traitement de l'image

[6]

- `imread` : Lire une image
- `figure` : crée un personnage à l'écran.
- `imshow(g)` : qui affiche la matrice `g` sous forme d'image.
- `pixval on` : Active les valeurs de pixel dans notre figure.
- `impixel(i, j)` : la commande retourne la valeur du pixel (i, j)
- `iminfo` : Information sur l'image

Les types de données

Data type	Description	Range
<code>int8</code>	8-bit integer	-128 — 127
<code>uint8</code>	8-bit unsigned integer	0 — 255
<code>int16</code>	16-bit integer	-32768 — 32767
<code>uint16</code>	16-bit unsigned integer	0 — 65535
<code>double</code>	Double precision real number	Machine specific

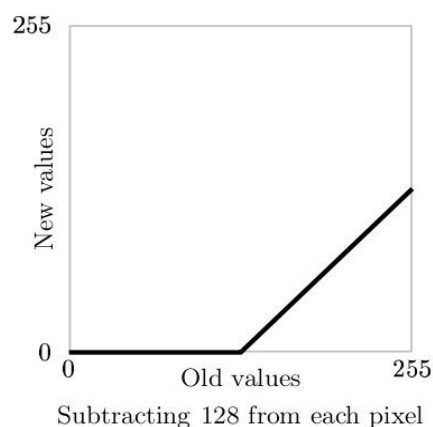
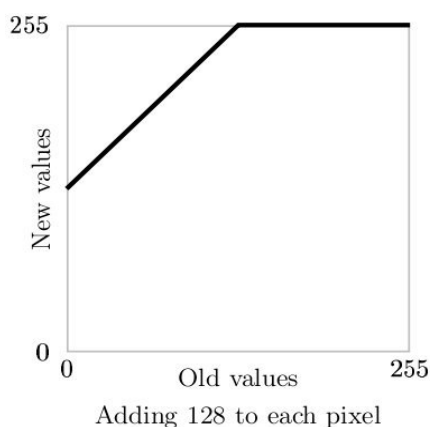
Résolution spatiale

- La résolution spatiale est la densité de pixels sur l'image: plus la résolution, spatial est grande, plus de pixels sont utilisés pour afficher l'image.
- Réduire de moitié la taille de l'image: Il le fait en prenant tous les rangs et toutes les colonnes, laissant ainsi uniquement les éléments de la matrice dont la rangée et la colonne sont de même indices.
- Doubler la taille de l'image: tous les pixels sont répétés pour produire une image avec le même format que l'original, mais avec la moitié de la résolution dans chaque direction.

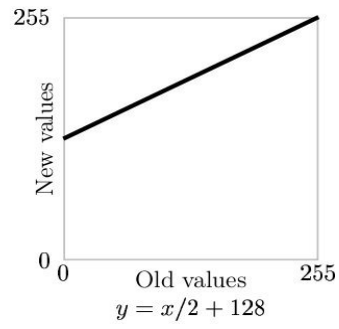
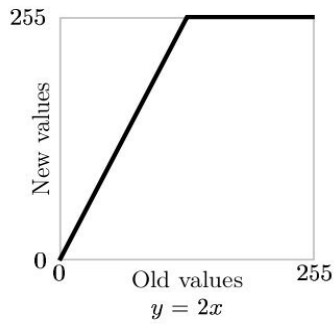
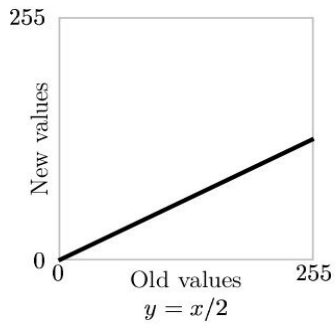
Opérations arithmétiques

- Ces opérations agissent en appliquant une simple fonction $y = f(x)$ pour chaque valeur de gris dans l'image.
- Les fonctions simples incluent l'ajout ou la soustraction d'une valeur constante pour chaque pixel: $y = x \pm C$ (iMADD, imsubtract)
- En multipliant chaque pixel par une constante: $y = C \cdot x$ (immultiply, imdivide)
- Complément: Pour une image à niveau de gris le complément est son négatif photographique.

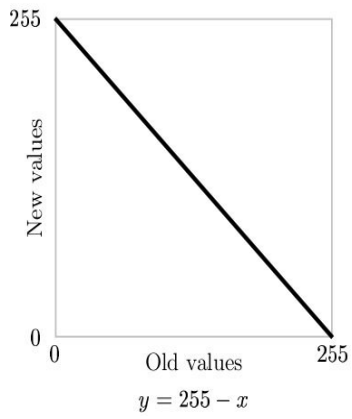
Addition soustraction



Multiplication division



Complément



Addition



Image: J



Image: J+20

Soustraction



Image: J Image: J-50

Multiplication

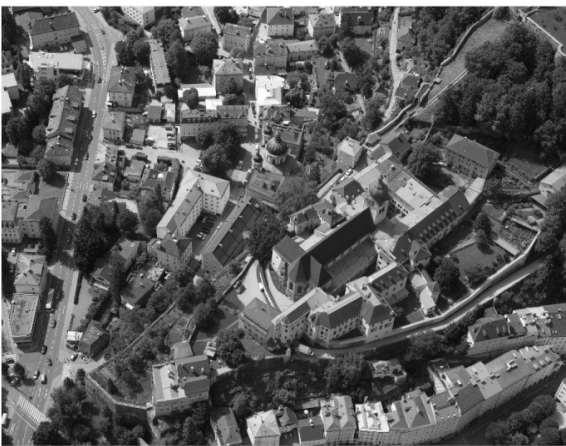


Image: J

Image: J*3

Division



Image: J

Image: J/2

Complément



Image: J



Image: 255-J

Histogrammes

- Étant donné une image en niveaux de gris, son histogramme est constitué de l'histogramme de ses niveaux de gris; ceci correspond à un graphique indiquant le nombre de fois où chacun des niveaux de gris est utilisé dans l'image.
- Nous pouvons en déduire beaucoup de choses sur l'apparence d'une image à partir de son histogramme.
- dans une image sombre, les niveaux de gris seraient regroupés à l'extrémité inférieure
- dans une image uniformément lumineuse, les niveaux de gris seraient regroupés à l'extrémité supérieure.
- Dans une image bien contrastée, les niveaux de gris seraient bien répartis sur une grande partie de la gamme.
- Problème: Étant donné une image pauvrement contrastée, le but étant d'améliorer ce contraste, en étalant l'histogramme. Il y a deux façons de le faire :

Étirement d'histogramme (étirement des contrastes)

- Image peu contrastés de gamme $[a, b]$
- Nous pouvons étirer les niveaux de gris dans le centre de la plage en appliquant une fonction linéaire

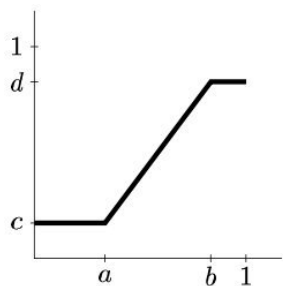
- Cette fonction a pour effet d'étirer les niveaux de gris [a, b] aux niveaux de gris [c, d], où

a < c et d > et b conformément à l'équation:

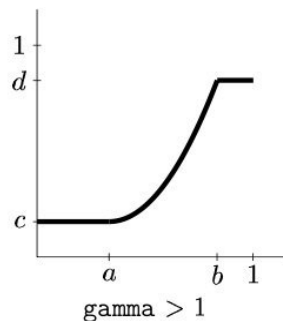
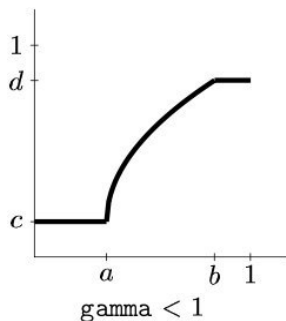
$$j = ((c-d) / (b-a)). (j-a) + c$$

- imadjust (I, [a, b], [c, d])

- les valeurs de pixel inférieures à c sont tous convertis en c, et des valeurs de pixel supérieures à d sont tous convertie en d



$$y = \left(\frac{x - a}{b - a} \right)^\gamma (d - c) + c.$$



Égalisation d'histogramme

- Le problème avec la méthode précédente d'étirement d'histogramme est qu'elles nécessitent des données d'entrée fournies par un utilisateur.
- L'égalisation d'histogramme, est une procédure entièrement automatique.
- Supposons que l'image a L niveaux de gris différents 0,1,2, ..., L-1 et que le niveau de gris i se produit ni fois dans l'image. Supposons également que le nombre total de pixels dans l'image soit n, de sorte que n₀ + n₁ + n₂ + ... + n_{L-1} = n. Pour transformer les niveaux de gris afin d'obtenir un meilleur contraste de l'image, on change de niveau de gris i à:

$$\left(\frac{n_0 + n_1 + \dots + n_i}{n} \right) (L - 1).$$

- et ce nombre est arrondi à l'entier le plus proche.

- un nombre à peu près égal de pixels est associé à chacun des niveaux de L, de sorte que l'histogramme de l'image de sortie soit sensiblement plan.

Seuillage

- Seuillage simple: une image en niveaux de gris est transformé en une image binaire en choisissant d'abord un niveau T de gris de l'image originale, puis de transformer chaque pixel noir ou blanc selon si sa valeur de gris est supérieur ou inférieur à T.
- Un pixel devient blanc si son niveau de gris est $> T$
- Un pixel devient noir si son niveau de gris est $\leq T$
- Double seuillage: Ici, nous choisissons deux valeurs T1 et T2 et nous appliquons la fonction de seuillage comme suit :
- Un pixel devient blanc si son niveau de gris entre T1 et T2
- Un pixel devient noir si son niveau de gris est autre

Filtrage spatial

- Déplacer un «masque»: un rectangle (généralement avec des côtés de longueur impaire) ou une autre forme de l'image.
- Une nouvelle image dont les pixels ont des valeurs de gris calculée à partir des valeurs de gris sous le masque.
- La combinaison et la fonction de masque est appelée filtre.
- La fonction linéaire de toutes les valeurs de gris dans le masque, le filtre est appelé un filtre linéaire.
- Le filtrage spatial nécessite 3 étapes:
 1. Positionner le masque sur le pixel courant,
 2. Rassembler tous les produits d'éléments de filtre avec les éléments correspondants de la zone,
 3. Ajouter l'ensemble des produits.
- Ceci doit être répété pour chaque pixel de l'image.
- filter2 (filtre, image, forme)

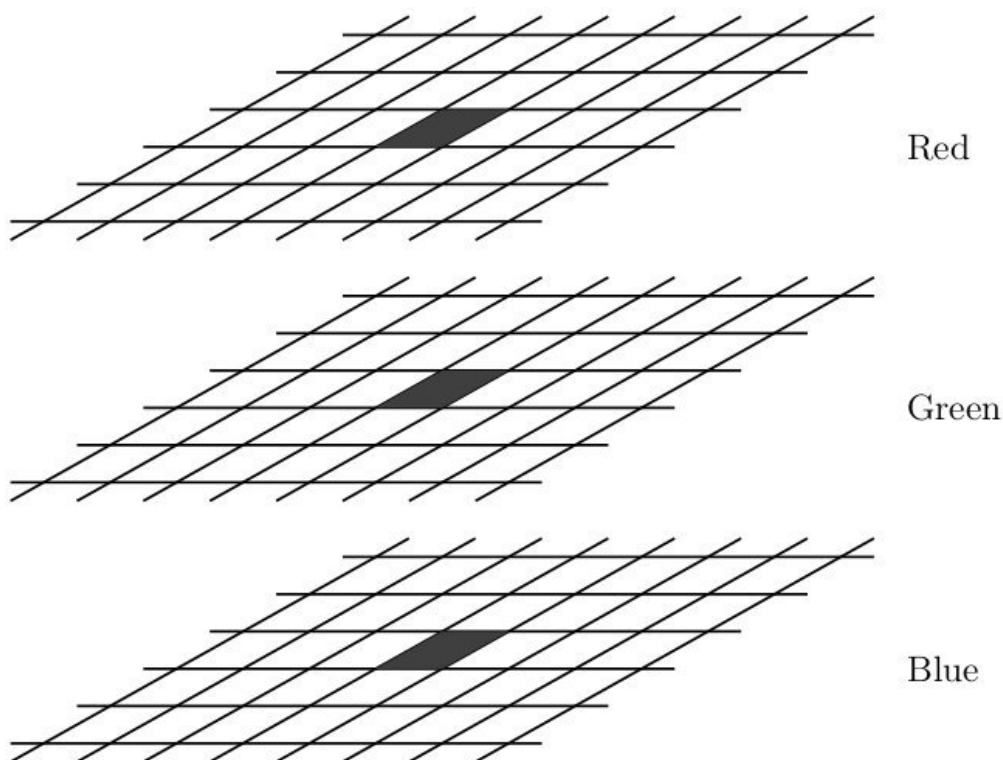
Images de couleur

- Un modèle de couleur est une méthode permettant de spécifier les couleurs d'une façon standard. Ce modèle se compose d'un système de coordonnées 3D et d'un

sous-espace de ce système, dans lequel chaque couleur est représentée par un seul point.

- RVB: Dans ce modèle, chaque couleur est représentée par 3 valeurs R, V et B, indiquant la quantité de rouge, vert et bleu qui forment la couleur.
- TSV:
- Teinte: L'attribut "true color" (rouge, vert, bleu, orange, jaune, et ainsi de suite).
- Saturation: Le montant par lequel la couleur comme a été dilué avec du blanc. Plus il y a de blanc dans la couleur plus basse est la saturation.
- Valeur: Le degré de luminosité: une couleur bien éclairé a une haute intensité, une couleur foncée a une faible intensité

Image colorée



Conversion de couleur

Function	Use	Format
ind2gray	Indexed to Greyscale	<code>y=ind2gray(x,map);</code>
gray2ind	Greyscale to indexed	<code>[y,map]=gray2ind(x);</code>
rgb2gray	RGB to greyscale	<code>y=rgb2gray(x);</code>
gray2rgb	Greyscale to RGB	<code>y=gray2rgb(x);</code>
rgb2ind	RGB to indexed	<code>[y,map]=rgb2ind;</code>
ind2rgb	Indexed to RGB	<code>y=ind2rgb(x,map);</code>

Chapitre 3

Méthodes / Application

Ce chapitre examine les algorithmes développés dans le cadre de l'application de reconnaissance faciale. Le but de cette étude est de mesurer les progrès globaux dans la reconnaissance du visage, de déterminer la maturité des algorithmes de reconnaissance de visage, et disposer d'un moyen indépendant de comparaison des algorithmes. Les tests mesurent la capacité des algorithmes pour gérer de grandes bases de données, des changements dans l'apparence des individus au fil du temps, les variations de l'éclairage, l'échelle et forme et les changements dans le fond.

Aussi décrit la mission principale du programme de technologie de reconnaissance faciale (*Face Recognition Technology program*) dans le développement des capacités de reconnaissance automatique des visages qui peuvent être employées pour aider la sécurité, l'intelligence ...

Pour atteindre ses objectifs, le programme effectue plusieurs tâches et s'est concentré sur trois points:

- Le développement de la base de la technologie nécessaire pour un système de reconnaissance faciale.
- La deuxième tâche importante commençant au début du programme
- La troisième est la collecte d'une grande base de données d'images faciales. Cette base de données des images de visage est une partie essentielle de projet global.

3-1 Méthode de Viola et Jones

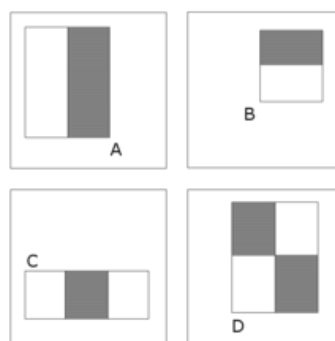
Cadre de détection d'objet Viola Jones

Le cadre de détection d'objet Viola Jones est le premier cadre de la détection d'objet dans la fourniture des taux de détection d'objets compétitifs en temps réel proposées en 2001 par Paul Viola et Michael Jones. Bien qu'il puisse être formé pour détecter une variété de classes d'objets, il a été principalement motivé par la problématique

de la détection des visages. Cet algorithme est implémenté dans OpenCV comme `cvHaarDetectObjects()`.

Types d'entités et d'évaluation

Les caractéristiques utilisées par le cadre de détection impliquent universellement les sommes de pixels de l'image au sein des zones rectangulaires. Comme tels, ils ressemblent quelque peu aux fonctions de base de Haar, qui ont été utilisés précédemment dans le domaine de la détection d'objet à base d'images. Toutefois, depuis les fonctions utilisées par Viola et Jones reposent toutes sur plus d'une zone rectangulaire, elles sont devenues généralement plus complexe.



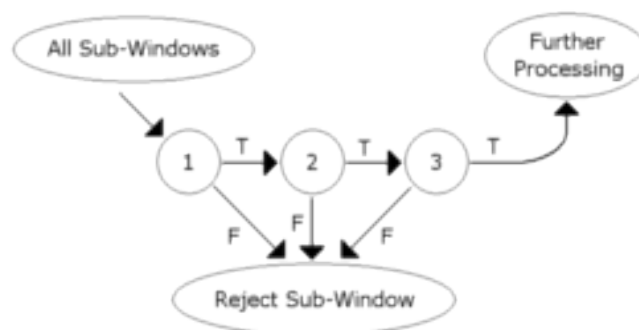
La figure illustre les quatre types de fonctions utilisées dans le cadre. La valeur d'un élément donné est toujours simplement la somme des pixels à l'intérieur des rectangles clairs soustraits de la somme des pixels à l'intérieur des rectangles ombragés. Comme on pouvait s'y attendre, les caractéristiques rectangulaires de ce genre sont assez primitives par rapport à d'autres solutions telles que les filtres orientables. Bien qu'ils soient sensibles aux traits verticaux et horizontaux, leur réaction est beaucoup plus grossière. Cependant, avec l'utilisation d'une représentation d'image appelé image intégrale, les caractéristiques rectangulaires peuvent être évaluées en temps constant, ce qui leur donne un avantage de vitesse considérable sur leurs ancêtres plus sophistiqués. Parce que chaque zone rectangulaire dans une fonction est toujours adjacente à au moins un autre rectangle, il s'ensuit que toute fonction de deux rectangles peut être calculée en six références de tableau, une caractéristique de trois rectangle en huit ans, et une fonction de quatre rectangle en seulement neuf.

Algorithme d'apprentissage

La rapidité avec laquelle les caractéristiques peuvent être évaluées ne compense pas adéquatement leur nombre. Cependant, par exemple, dans une sous fenêtre de pixels 24x24 standard, il ya un total de 162 336 les fonctions possibles, et il serait trop coûteux de les évaluer tous. Ainsi, le cadre de la détection d'objets utilise une variante de l'algorithme d'apprentissage AdaBoost qui permet à la fois de sélectionner les meilleurs éléments et de former les classificateurs qui les utilisent.

Architecture cascade

L'évaluation des classificateurs forts générés par le processus d'apprentissage peut être faite rapidement, mais ce n'est pas assez rapide pour exécuter en temps réel. Pour cette raison, les classificateurs forts sont disposés en cascade par ordre de complexité, où chaque classificateur successif est formé uniquement sur les échantillons sélectionnés qui traversent les classificateurs précédents. Si, à n'importe quel stade de la cascade un classificateur rejette la sous fenêtre sous contrôle, aucun autre traitement n'est effectué et il continue sur la recherche de la prochaine sous fenêtre. La cascade a donc la forme d'un arbre dégénéré. Dans le cas des visages, le premier classificateur en cascade - appelé l'opérateur attentionnel - utilise seulement deux traits pour atteindre un taux de faux négatifs de l'ordre de 0% et un taux de faux positifs de 40%. L'effet de ce seul classificateur est de réduire environ de moitié le nombre de fois où l'ensemble de cascade est évalué [7].



L'architecture de la cascade a des implications intéressantes pour la performance des classificateurs individuels. Parce que l'activation de chaque classificateur

dépend entièrement du comportement de son prédécesseur, le taux de faux positifs pour une cascade entière est:

$$F = \prod_{i=1}^K f_i.$$

De même, le taux de détection est:

$$D = \prod_{i=1}^K d_i.$$

Ainsi, pour faire correspondre les taux de faux positifs généralement obtenus par d'autres détecteurs, chaque classificateur peut s'en tirer étonnamment en ayant de faibles performances. Par exemple, pour une cascade de 32 étapes pour atteindre un taux de faux positifs, chaque classificateur a besoin seulement d'atteindre un taux de faux positifs d'environ 65%. Dans le même temps, cependant, chaque classificateur doit être exceptionnellement capable s'il veut atteindre le taux de détection adéquats. Par exemple, pour atteindre un taux de détection de 90%, chaque classificateur dans la cascade susmentionné a besoin pour atteindre un taux de détection de l'ordre de 99,7%.

Éléments de la méthode

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. Une première innovation de la méthode est l'introduction des images intégrales, qui permettent le calcul rapide de ces caractéristiques. Une deuxième innovation importante est la sélection de ces caractéristiques par boosting, en interprétant les caractéristiques comme des classifiées. Enfin, la méthode propose une architecture pour combiner les classifiées boostés en un processus en cascade, ce qui apporte un net gain en temps de détection.

Limitations et extensions

De très nombreuses améliorations ont été proposées par la suite, visant à améliorer le paramétrage de la méthode, ou à en combler un certain nombre de limitations.

L'une des premières améliorations est apportée par Lienhart et Maydt en 2002. Ils proposent d'étendre l'ensemble de caractéristiques pseudo-Haar utilisé de 4 à 14 caractéristiques. De même, ils introduisent des caractéristiques « de biais » (tournées de 45°), ainsi qu'une méthode pour les calculer basée sur une extension des images intégrales.

D'autres types de caractéristiques ont également été utilisés en remplacement des caractéristiques de Haar : les histogrammes de gradients orientés, les motifs binaires locaux ou la covariance de région. Les chercheurs ont également proposé d'utiliser des variantes de l'algorithme de boosting, notamment RealBoost, qui produit un indice de confiance à valeurs réelles, en plus de la classification. Plusieurs travaux ont ainsi montré la supériorité de RealBoost sur AdaBoost dans le cadre de l'algorithme de Viola et Jones.

Viola et Jones étendent en 2003 leur système à la détection de piétons dans des vidéos, en incluant une information de mouvement en plus de l'information d'apparence.

Une des limitations de la méthode est son manque de robustesse à la rotation, et sa difficulté à apprendre plusieurs vues d'un même objet. En particulier, il est difficile d'obtenir un classifieur capable de détecter à la fois des visages de face et de profil. Viola et Jones ont proposé une amélioration qui permet de corriger ce défaut, qui consiste à apprendre une cascade dédiée à chaque orientation ou vue, et à utiliser lors de la détection un arbre de décision pour sélectionner la bonne cascade à appliquer. Plusieurs autres améliorations ont été proposées par la suite pour apporter une solution à ce problème.

Une autre limitation importante de la méthode de Viola et Jones concerne le temps d'apprentissage de la cascade, compris généralement entre plusieurs jours et plusieurs semaines de calcul, ce qui limite sévèrement les possibilités de tests et de choix des paramètres.

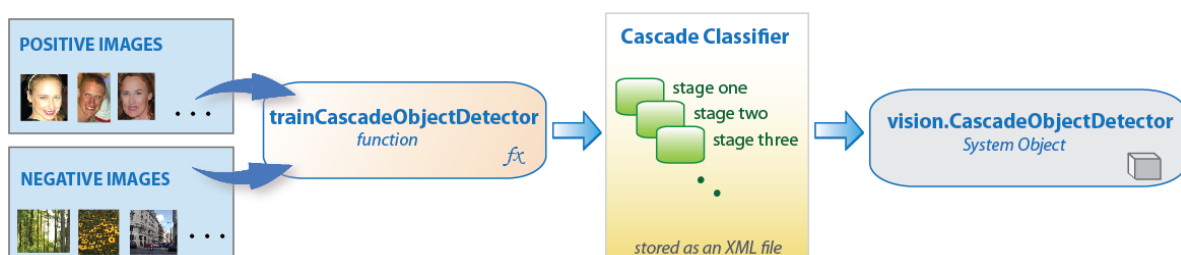
Un des problèmes majeurs de la méthode proposée par Viola et Jones est qu'il n'existe pas de méthode optimale pour choisir les différents paramètres régissant l'algorithme : le nombre d'étages, leur ordre ou les taux de détection et de fausses alarmes pour chaque étage doivent être choisis par essais et erreurs. Plusieurs méthodes sont proposées pour déterminer certains de ces seuils de manière automatique.

Un reproche également fait à la méthode concerne la perte d'information subie au passage d'un étage à l'autre de la cascade, perte due à l'effet couperet des décisions d'acceptation ou de rejet prises à chaque étage. Certains chercheurs proposent la solution de garder l'information contenue dans la somme pondérée des classifiées faibles, par exemple le «boosting Chain» de Xiao. Une modification plus radicale de structure est proposée par Bourdev et sa notion de cascade souple, qui consiste essentiellement à supprimer le concept d'étages, en formant un seul classifieur fort, donc une seule somme, et en permettant de prendre une décision à chaque évaluation de classifieur faible et de s'affranchir de la contrainte des taux de détection et de fausses alarmes cibles.

3-2 Cascade Object Detector

Pourquoi former un détecteur?

L'objet du système `vision.CascadeObjectDetector` est livré avec plusieurs classificateurs pré entraînés pour détecter des faces frontales, des visages de profil, le nez, l'ensemble du buste et les yeux. Cependant, ces classificateurs peuvent ne pas être suffisants pour une application particulière. Computer Vision System Toolbox fournit la fonction `trainCascadeObjectDetector` qui sert à former un classificateur personnalisé.



Quel genre d'objets peut être détecté?

Le détecteur d'objet en cascade Computer Vision System Toolbox peut détecter des catégories d'objets dont le rapport d'aspect ne varie pas de façon significative. Les objets dont le rapport reste un quasi aspect fixe comprennent des visages, des panneaux d'arrêt, ou des voitures vues de côté.

L'objet du système `vision.CascadeObjectDetector` détecte des objets dans des images en faisant glisser une fenêtre sur l'image. Le détecteur utilise ensuite un

classificateur en cascade pour décider si la fenêtre contient l'objet d'intérêt. La taille de la fenêtre varie afin de détecter des objets à différentes échelles, mais son ratio d'aspect reste fixe. Le détecteur est très sensible à la rotation hors plan, parce que le rapport d'aspect change pour la plupart des objets 3-D. Ainsi, il faut former un détecteur pour chaque orientation de l'objet. La formation d'un détecteur unique pour gérer toutes les orientations ne fonctionnera pas.

Comment fonctionne le classificateur en cascade?

Le classificateur en cascade se compose d'étapes, où chaque étape est un ensemble de faibles apprentis. Les faibles apprentis sont de simples classificateurs appelés «ouches de décision». Chaque étape est formée à l'aide d'une technique appelée «rappel». La fonction «Stimuler» offre la possibilité de former un classificateur très précis en prenant une moyenne pondérée des décisions prises par les faibles apprentis.

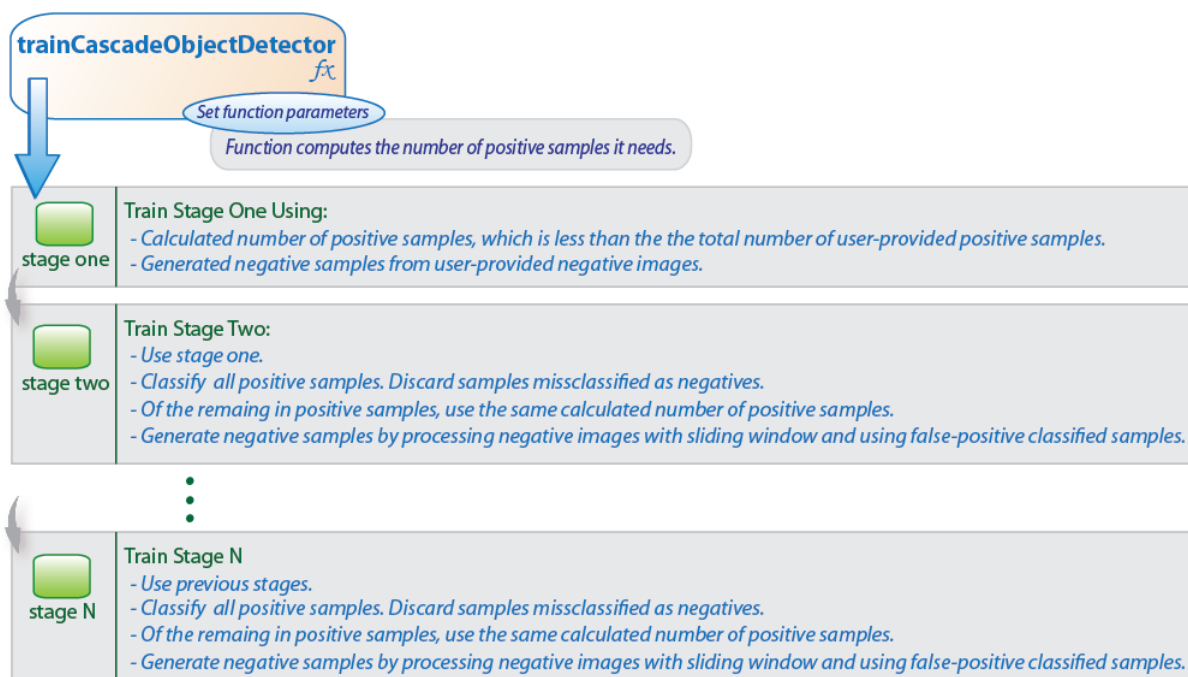
Chaque étape de l'étiquette du classificateur, la région définie la position actuelle de la fenêtre glissante comme positive ou négative. Positif indique qu'un objet a été trouvé et négatif indique qu'aucun objet n'est présent. Si le marqueur est négatif, la classification de cette région est terminée, et le détecteur glisse la fenêtre à l'emplacement suivant. Si le marqueur est positif, le classificateur passe la région à l'étape suivante. Le détecteur signale un objet situé à l'emplacement de la fenêtre lorsque la dernière étape région classifie la région comme positive.

Les étapes sont conçues pour rejeter les échantillons négatifs aussi vite que possible. L'hypothèse est que la grande majorité des fenêtres ne contiennent pas l'objet d'intérêt. Inversement, les vrais positifs sont rares, et il faut prendre le temps de vérifier. Un vrai positif survient lorsqu'un échantillon positif est correctement classé. Un faux positif survient lorsqu'un échantillon est négatif et est considéré à tort comme positif. Un faux négatif se produit quand un échantillon positif est à tort considéré comme négatif. Pour bien fonctionner, chaque étape de la cascade doit avoir un faible taux de faux négatifs. Si une étape qualifie à tort un objet comme négatif, le classement s'arrête, et il n'existe aucun moyen de corriger l'erreur. Cependant, chaque étape peut avoir un taux élevé de faux positifs. Même si elle qualifie à tort un non objet comme positif, l'erreur peut être corrigée par des étapes ultérieures.

Le taux de faux positifs globale du classificateur en cascade est f_s , où f est le taux de faux positifs par étape dans l'intervalle (0, 1), et s est le nombre d'étapes. De même, le taux de vrais positifs global est t_s , où t est le taux de vrais positifs par étape dans l'intervalle (0 1]. Ainsi, il est possible de voir que l'ajout de plus d'étapes réduit le taux de faux positifs dans l'ensemble, mais il réduit également le taux général de vrais positifs.

Comment utiliser la fonction `trainCascadeObjectDetector` pour créer un classificateur en cascade

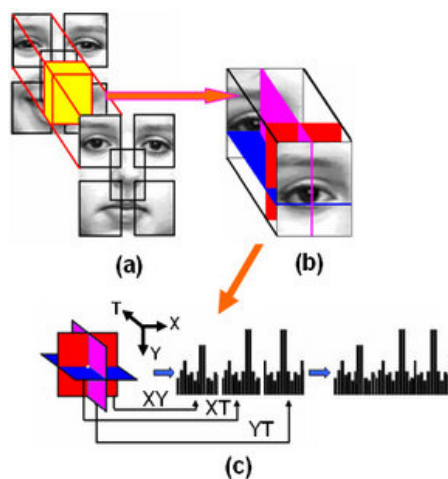
La formation du classificateur en cascade nécessite un ensemble d'échantillons positifs et d'un ensemble d'images négatives. Il faut fournir un ensemble d'images positives avec des régions d'intérêt prévues pour être utilisés comme échantillons positifs. Il faut également fournir un ensemble d'images négatives à partir duquel la fonction génère automatiquement les échantillons négatifs. Il faut également définir le nombre d'étapes, type de fonction, et d'autres paramètres de la fonction pour obtenir une précision de détection acceptable.



3-3 Local Binary Pattern (LBP)

LBP est un opérateur de texture simple mais très efficace qui marque les pixels d'une image en souillant le voisinage de chaque pixel et en considérant le résultat

comme un nombre binaire. Grâce à son pouvoir discriminant et la simplicité de calcul, l'opérateur de texture LBP est devenu une approche populaire dans diverses applications. Il peut être considéré comme une démarche fédératrice pour les modèles statistiques et structurels traditionnellement divergentes de l'analyse de texture. La caractéristique la plus importante de l'opérateur LBP dans les applications du monde réel est sa robustesse aux variations de gris monotones causées, par exemple, par des variations d'éclairage. Une autre propriété importante est sa simplicité de calcul, ce qui permet d'analyser les images en contestant les paramètres en temps réel [8].



LBP dans le domaine spatial

L'idée de base pour le développement de l'opérateur LBP était que les textures de surface à deux dimensions peuvent être décrites par deux mesures complémentaires: structures spatiales locales et le contraste de gris. L'opérateur d'origine LBP [9] forme des étiquettes pour les pixels d'image par seuillage du voisinage 3 x 3 de chaque pixel avec la valeur centrale et en considérant le résultat sous forme de nombre binaire. L'histogramme de ces $2^8 = 256$ étiquettes différentes peut alors être utilisé comme un descripteur de texture. Cet opérateur utilisé conjointement avec une mesure de contraste local simple fournit de très bonnes performances dans la segmentation de texture non supervisée [10]. Après cela, de nombreuses approches connexes ont été développées pour la texture et la segmentation de texture de couleur.

L'opérateur LBP a été étendu pour utiliser des voisinages de différentes tailles [12]. L'utilisation d'un voisinage circulaire et de valeurs d'interpolation bilinéaire à coordonnées non entières de pixels autorise n'importe quel rayon et n'importe quel

nombre de pixels dans le voisinage. La variance de l'échelle de gris du voisinage local peut être utilisée comme mesure de contraste complémentaire. Par la suite, la notation (P, R) sera utilisé pour les quartiers de pixels qui signifie points de prélèvement P sur un cercle de rayon R. voir fig. 2 pour un exemple de calcul LPB.

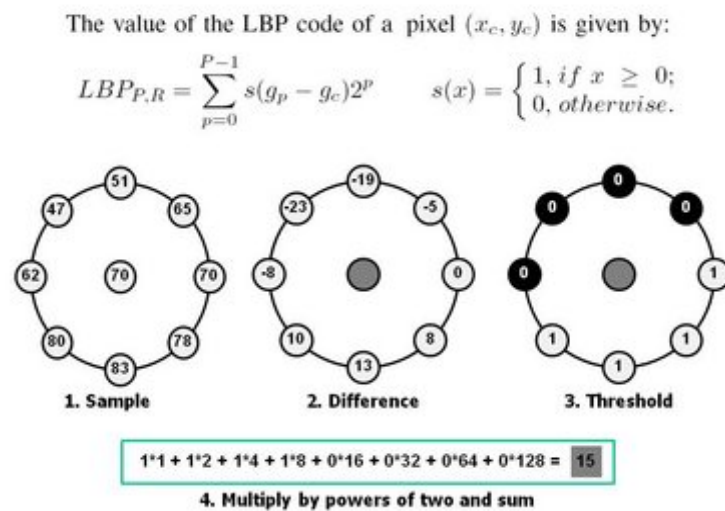


Figure 2 : exemple de calcul de LBP

Une autre extension à l'opérateur d'origine est la définition de modèles uniformes, qui peuvent être utilisés pour réduire la longueur du vecteur de caractéristiques et la mise en œuvre d'une simple rotation descripteur invariant. Cette extension a été inspirée par le fait que certains motifs binaires sont plus fréquents dans les images de texture que les autres. Un motif binaire locale est appelée uniforme si la configuration binaire contient au plus deux transitions de bits de 0 à 1 ou inversement lorsque la configuration de bits est traversé circulairement. Par exemple, les modèles 00000000 (0 transitions), 01110000 (2 transitions) et 11001111 (2 transitions) sont uniformes, tandis que les modèles 11001001 (4 transitions) et 01010010 (6 transitions) ne le sont pas. Dans le calcul de l'étiquette LPB, des motifs uniformes sont utilisés de telle sorte qu'il y ait une étiquette séparée pour chaque motif uniforme et tous les motifs non uniformes sont étiquetés avec une étiquette unique. Par exemple, lorsque vous utilisez le voisinage (8, R), il y a un total de 256 motifs, dont 58 sont uniformes, ce qui donne en 59 étiquettes différentes. Ojala et al. (2002) ont remarqué dans leurs expériences avec des images de texture qui comptent des modèles uniformes pour un peu moins de 90% de tous les modèles lorsque vous utilisez le voisinage (8,1) et environ 70% dans le quartier

(16,2). Chaque bac (code LBP) peut être considéré comme un micro-texton. Les primitives locales qui sont codifiées par ces bacs comprennent différents types de bords incurvés, des taches, des zones plates, etc.

La notation suivante est utilisée pour l'opérateur LBP: LBPP, Ru2. L'indice représente à l'aide de l'opérateur dans une zone (P, R). L'exposant u2 signifie en utilisant uniquement des modèles uniformes et l'étiquetage de tous les modèles restants avec une seule étiquette. Après la LBP l'image $f_l(x, y)$ a été obtenue, l'histogramme LPB peut être définie comme

$$H_i = \sum_{x,y} I\{f_l(x,y)=i\}, i=0, \dots, n-1 ;$$

dans laquelle n est le nombre d'étiquettes différentes produites par l'opérateur LBP, et $I\{A\}$ est égal à 1 si A est vrai et 0 si A est faux.

Lorsque les taches d'image dont les histogrammes doivent être comparés sont de tailles différentes, les histogrammes doivent être normalisés pour obtenir une description cohérente.

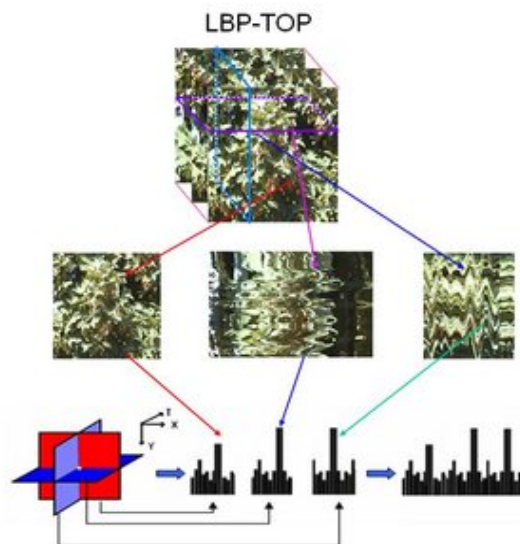
Le LPB spatio-temporelle

L'opérateur LPB originale a été défini pour traiter uniquement l'information spatiale. Plus tard, il a été étendu à une représentation spatio-temporelle pour l'analyse de texture dynamique. A cet effet, le volume que l'on appelle modèle d'opérateur binaire local (VLBP) a été proposé [13]. L'idée derrière VLBP consiste à examiner la texture dynamique comme un ensemble de volumes dans l'espace (X, Y, T) où X et Y représentent les coordonnées spatiales et T désigne l'indice de trame (temps). Le voisinage de chaque pixel est ainsi défini dans un espace tridimensionnel. Puis, comme pour le LPB dans le domaine spatial, les textures de volume peuvent être définies et extraites dans les histogrammes. Par conséquent, VLBP combine le mouvement et l'apparence de concert pour décrire les textures dynamiques.

Pour faire le VLBP simplement calculable et facile à étendre, un opérateur basé sur les cooccurrences de motifs binaires locaux sur trois plans orthogonaux (LBP- TOP) a également été introduit. Le LBP -TOP considère trois plans orthogonaux : XY, XT et YT, et concatène la cooccurrences locales statistiques de motifs binaires dans ces trois directions comme illustré à la figure 1. Les quartiers circulaires sont généralisés à l'échantillonnage elliptique pour s'adapter aux statistiques spatio-temporelles.

La figure 3 montre des exemples d'images de trois plans. Le plan XY représente les informations de l'apparence, alors que le plan XT donne une impression visuelle d'une rangée évolutive dans le temps et YT décrit le mouvement d'une colonne dans l'espace temporel. Les codes LBP sont extraits pour tous les pixels de la XY, XT et avions YT, noté XY- LBP, XT- LBP et YT- LBP, et les histogrammes de ces plans sont calculés et concaténer en un seul histogramme. Dans une telle représentation, une texture dynamique est codée par une apparition (XY- LBP) et de deux statistiques de co- occurrence spatio-temporels (XT- LBP et YT- LBP).

Régler le rayon dans l'axe de temps pour être égal au rayon dans l'axe de l'espace n'est pas raisonnable pour les textures dynamiques. Nous avons donc différents paramètres de rayon dans l'espace et le temps à définir. Dans les plans XT et YT, de rayons différents peuvent être affectés pour échantillonner des points voisins dans l'espace et le temps. Plus généralement, les rayons dans les axes X, Y et T, et le nombre de points voisins dans les plans XY, XT et YT peuvent aussi être différents.



Description du visage à l'aide du LPB

Dans l'approche du LPB pour la classification de la texture, les occurrences du code LPB dans une image sont collectées dans un histogramme. Le classement est ensuite effectué en calculant les simples similitudes de l'histogramme. Toutefois, compte tenu d'une approche similaire pour l'image faciale résultats de représentation à une perte d'information spatiale et donc on devrait coder les informations de texture tout en conservant également leurs emplacements. Une façon d'atteindre cet

objectif est d'utiliser les descripteurs de texture LPB pour construire plusieurs descriptions locales du visage et les combiner en une description globale. Ces descriptions locales ont gagné en intérêt ces derniers temps ce qui est compréhensible étant donné les limites des représentations holistiques. Ces méthodes basées sur les caractéristiques locales sont plus robustes contre les variations de pose ou d'éclairage que les méthodes holistiques. La méthodologie de base pour la description du visage est proposée par [14] et se présente comme suit: L'image du visage est divisé en régions locales et les descripteurs de texture LPB sont extraits de chaque région de façon indépendante. Les descripteurs sont ensuite enchaînés pour former une description globale du visage, comme le montre la figure. 4

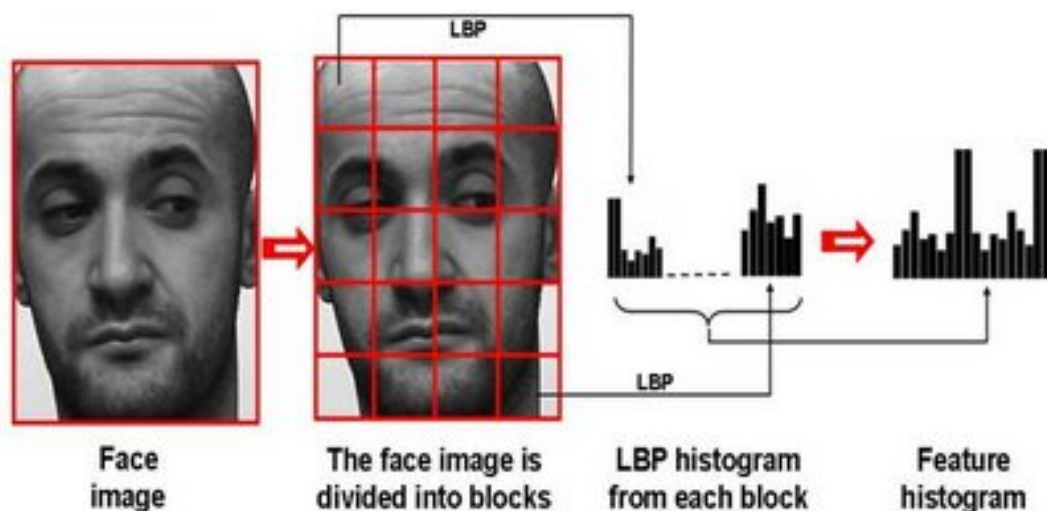


Figure 4: Description de visage avec des motifs binaires locaux

Cet histogramme a effectivement une description de la face sur trois niveaux différents de localité: les étiquettes LPB pour l'histogramme contient des informations sur les motifs sur un niveau de pixel, les étiquettes sont résumées sur une petite région pour produire de l'information au niveau régional et les histogrammes régionaux sont enchaînés pour construire une description globale du visage. Il convient de noter que lorsque vous utilisez les méthodes basées sur l'histogramme les régions n'ont pas besoin d'être rectangulaire. Pas plus qu'ils ne doivent être de la même taille ou la forme, et ils ne doivent pas nécessairement couvrir l'ensemble de l'image. Il est également possible d'avoir recouvrement partiel des régions.

La méthode de description visage en deux dimensions a été étendue en domaine spatio-temporel [13]. La figure 1 représente la description de l'expression du visage à l'aide du LBP-TOP. L'excellente performance de reconnaissance des expressions faciales a été obtenue avec cette approche.

Depuis la publication de la description du visage à base LBP, la méthodologie a déjà atteint une position établie face à la recherche et des applications d'analyse. Un exemple notable est le système d'illumination invariant système de reconnaissance faciale proposée dans [15], en combinant l'imagerie NIR avec des fonctionnalités LBP et l'apprentissage Adaboost. Zhang et al. [16] ont proposé l'extraction de caractéristiques LBP à partir d'images obtenues par filtrage d'une image faciale avec 40 filtres de Gabor à différentes échelles et orientations, en obtenant des résultats remarquables. Hadid et Pietikäinen [17] ont utilisé le LBP spatio-temporelles pour la reconnaissance du visage et du sexe à l'aide de séquences vidéo, tandis que Zhao et al. [18] a adopté l'approche LBP-TOP de reconnaissance vocale visuelle en atteignant des performances de pointe sans erreurs segmentation de mouvement.

Extensions et applications

La méthodologie LBP a conduit à des progrès significatifs dans l'analyse de texture. Il est largement utilisé partout dans le monde à la fois dans la recherche et les applications. Grâce à son pouvoir discriminant et la simplicité de calcul, la méthode a très bien réussi dans beaucoup de ces problèmes de vision par ordinateur qui n'existait pas auparavant, même considérés comme des problèmes de texture, telles que l'analyse du visage et l'analyse du mouvement [19].

Pour augmenter l'applicabilité de LBP, diverses extensions et modifications de celui-ci ont été proposées. Par exemple, Liao et al. [20] ont proposé des modèles binaires locaux dominants qui font appel à des modèles plus fréquemment survenus de LBP pour améliorer la précision de la reconnaissance. L'utilisation de descripteurs de région d'intérêt (comme les EIPD) à divers problèmes de vision par ordinateur a été d'un grand intérêt récemment. A cet effet, un roman descripteur combinant les forces de EIPD et LBP a été proposé [21], dans lequel les motifs binaires locaux centre symétriques (CS- LBP) ont été utilisés pour remplacer l'opérateur gradient utilisé par l'opérateur EIPD. Mäenpää et Pietikäinen [22] ont proposé un adversaire de couleur LBP, et étudiés l'utilisation conjointe et séparée de la couleur et de la texture dans la

classification. La combinaison des caractéristiques PSL et Gabor a été étudiée [23 et 24].

La première méthode basée sur la texture pour soustraction de fond a été proposée dans [25]. Chaque pixel est modélisé comme un ensemble d'histogrammes de motifs locaux adaptatifs binaires qui sont calculées sur une zone circulaire autour du pixel. Ceci a montré être tolérants aux variations d'éclairage, la multi modalité de l'arrière-plan, et l'introduction ou la suppression d'objets de fond. En outre, le procédé est capable de traiter en temps réel. Un algorithme de prétraitement sur la base des LPB a été mis au point pour traiter les variations d'éclairage dans un système d'authentification de visage [26]. L'utilisation des LPB dans la reconnaissance des actions a été examinée par [27]. L'utilisation des LPB dans la classification d'âge du visage a été étudiée [24]. D'autres approches basées sur LBP liés à ces problèmes ont été proposées récemment.

En plus de la reconnaissance faciale et l'expression du visage, la LBP a également été utilisé dans de nombreuses autres applications de la biométrie, y compris la localisation de l'œil, reconnaissance de l'iris, la reconnaissance d'empreintes digitales, reconnaissance de démarche et classification par âge du visage. Les références à la plupart de ces œuvres peuvent être trouvées dans la bibliographie LBP.

Due to its discriminative power and computational simplicity, LBP texture operator has become a popular approach in various applications. It can be seen as a unifying approach to the traditionally divergent statistical and structural models of texture analysis. Perhaps the most important property of the LBP operator in real-world applications is its **robustness to monotonic gray-scale changes caused, for example, by illumination variations**. Another important property is its computational simplicity, which makes it possible to analyze images in challenging real-time settings

3-4 Suivi de points dans une vidéo à l'aide de l'algorithme de Kanade-Lucas-Tomasi (KLT)

Description [28]

L'objet de traqueur du point suit un ensemble de points à l'aide de l'algorithme fonction de suivi Kanade-Lucas-Tomasi (KLT). Vous pouvez utiliser le traqueur de

point pour la stabilisation vidéo, l'estimation du mouvement de la caméra, et le suivi d'objets. Il fonctionne particulièrement bien pour le suivi d'objets qui ne changent pas la forme et pour ceux qui présentent une texture visuelle. Le point traqueur est souvent utilisé pour le suivi à court terme dans le cadre d'un cadre de suivi plus large.

Comme l'algorithme de suivi de point de progresse dans le temps, des points peuvent être perdues en raison d'une variation de l'éclairage, de la rotation du plan, ou le mouvement articulé. Pour suivre un objet sur une longue période de temps, vous pouvez être amené à acquérir périodiquement points.

Construction

`pointTracker = vision.PointTracker` renvoi à un objet système, `pointTracker`, permet de suivre le traqueur de point dans une video.

`pointTracker = vision.PointTracker(Name,Value)` configure les propriétés de l'objet Tracker, définie comme un ou plusieurs de paires d'arguments de nom-valeur. Les propriétés non spécifiées ont des valeurs par défaut

Initialiser le processus de suivi

Pour initialiser le processus de suivi, il faut utiliser la méthode d'initialisation pour indiquer les emplacements initiaux des points et l'image vidéo initiale. `initialize(pointTracker,points,I)` initialise les points à suivre et fixe le cadre de la vidéo initiale. Les emplacements initiaux de POINTS, doit être un tableau de Mx2 intégrant des coordonnées [x,y]. L'image vidéo initiale, I, doit être un 2-D en niveaux de gris ou une image RVB et doivent avoir la même taille et le type de données que les images vidéo transmises à la méthode « step ».

Les fonctions `detectFASTFeatures`, `detectSURFFeatures`, `detectHarrisFeatures`, et `detectMinEigenFeatures`, ainsi que l'objet système `vision.CornerDetector` sont quelques-unes des nombreuses façons d'obtenir les points de départ pour le suivi.

Pour suivre un ensemble de points:

1. Définir et mettre en place votre objet traqueur de point à l'aide du constructeur
2. Appelez la méthode `step` avec l'image d'entrée, I, et l'objet de traqueur de points, `pointTracker`. Voir la syntaxe suivante pour utiliser la méthode `step`.

Après l'initialisation du processus de suivi, utilisez la méthode `step` pour suivre les points dans des images vidéo ultérieures. Vous pouvez également réinitialiser les points à tout moment en utilisant la méthode `setPoints`.

`[points,point_validity] = step(pointTracker,I)` suit des points dans la trame d'entrée, `I` utilisant le point traqueur objet, `pointTracker`. Les données de sorties comprennent un tableau de `Mx2` données avec des coordonnées `[x,y]` ceci correspondent aux nouveaux emplacements des points dans la trame d'entrée, `I`. La sortie `point_validity` fournit un tableau logique de `Mx1` données, ce qui indique si oui ou non chaque point a été suivi de façon fiable.

Un point peut être invalide pour plusieurs raisons. Le point peut devenir non valide s'il se retrouve à l'extérieur de l'image. En outre, il peut devenir invalide si la matrice de gradient spatial calculée dans son voisinage est singulière. Si l'erreur bidirectionnelle est supérieure au seuil `MaxBidirectionalError`, cette condition peut aussi rendre le point invalide.

`[points,point_validity,scores] = step(pointTracker,I)` renvoie en outre le score de confiance pour chaque point. La matrice de données de sortie `Mx1`, `scores`, contient des valeurs comprises entre 0 et 1. Ces valeurs correspondent au degré de similitude entre la zone autour de l'emplacement précédent et le nouvel emplacement de chaque point. Ces valeurs sont calculées en fonction de la somme des différences au carré entre les voisinages précédents et les nouveaux. La plus grande confiance de suivi correspond à un score parfait de 1.

`setPoints (pointTracker, points)` définit les points de suivi. La méthode définit le tableau `Mx2` de `[xy]` coordonnées avec les points à suivre. Vous pouvez utiliser cette méthode si les points doivent être redétection parce que trop d'entre eux ont été perdus pendant le suivi.

`setPoints(pointTracker,points,point_validity)` permet en plus de marquer de points comme valide ou non valide. Le vecteur d'entrée logique `point_validity` de longueur `M`, contient la valeur vraie ou fausse correspondant à la validité du point à suivre. La longueur `M` correspond au nombre de points. La valeur fausse indique un point invalide qui ne devrait pas être suivi.

Exemple

- Créer des objets du système de lecture et d'affichage de la vidéo et de l'élaboration

```
videoFileReader = vision.VideoFileReader('visionface.avi');  
videoPlayer = vision.VideoPlayer('Position', [100, 100, 680, 520]);
```

- Lire la première trame vidéo, qui contient l'objet, définir la région.

```
objectFrame = step(videoFileReader);  
objectRegion = [264, 122, 93, 93];
```

Comme alternative, vous pouvez utiliser les commandes suivantes pour sélectionner l'objet de la région en utilisant une souris. L'objet doit occuper la majeure partie de la région.

```
Figure ; imshow(objectframe) ;  
objectRegion=round(getposition(imrect))
```

- Afficher trame initiale avec une boîte rouge.

```
objectImage = insertShape(objectFrame, 'Rectangle',  
objectRegion,'Color', 'red');  
figure; imshow(objectImage); title('Yellow box shows object region');
```

- Détecter les points d'intérêt dans la région de l'objet.

```
points = detectMinEigenFeatures(rgb2gray(objectFrame), 'ROI',  
objectRegion);
```

- Afficher les points détectés.

```
pointImage = insertMarker(objectFrame, points.Location, '+', 'Color',  
'white');  
figure, imshow(pointImage), title('Detected interest points');
```

- Créer un objet de suivi.

```
markerInsertervision.MarkerInserter('Shape','Plus','
```

```
BorderColor','White');
```

- Créer un objet de suivi.

```
tracker= vision.PointTracker('MaxBidirectionalError', 1);
```

- Initialiser le tracker.

```
initialize(tracker, points.Location, objectFrame);
```

- Lire, suivre, afficher des points, et les résultats de chaque image vidéo.

```
while ~isDone(videoFileReader)
```

```
    frame = step(videoFileReader);
```

```
    [points, validity] = step(tracker, frame);
```

```
    out = insertMarker(frame, points(validity, :), '+');
```

```
    step(videoPlayer, out);
```

```
end
```

- libérer le lecteur vidéo

```
release(videoPlayer);
```

```
release(videoFileRea
```

3-5 Application

Système de reconnaissance de visage

La reconnaissance faciale (**Face Recognition**) est un thème général qui comprend à la fois la reconnaissance des visages et la détection de visage (également appelée authentification). D'une part, la détection de visage est concernée par la validation d'une identité revendiquée sur la base de l'image d'un visage, et l'acceptation ou au rejet de la demande d'identité (aussi appelé authentification).

D'autre part, l'objectif de la reconnaissance de visage est d'identifier une personne sur la base de l'image d'un visage. Cette image du visage doit être comparée avec toutes les personnes inscrites (correspondance point à point).

La principale approche à la reconnaissance faciale considère à la fois la forme et les informations de texture pour représenter des images de visage. La zone du visage est d'abord divisé en petites régions à partir de modèle histogrammes binaires locaux (LBP) sont extraites et enchaînés en une seule caractéristique spatiale améliorée histogramme représentant efficacement l'image du visage. La reconnaissance est effectuée en utilisant un classificateur de voisin le plus proche dans l'espace de caractéristiques calculées avec le carré de Chi en tant que mesure de différence.

Le système mis au point est capable de reconnaître un visage inconnu donné une base de données de visages connus. Le code doit reconnaître même si le visage d'entrée n'est pas présent dans la base.

À partir d'une image d'entrée, nous devons détecter tous les visages. Chaque visage détecté est comparé à la base de données. Lorsqu'une correspondance est trouvée, toutes les informations relatives à l'identité correspondante sont visualisées.

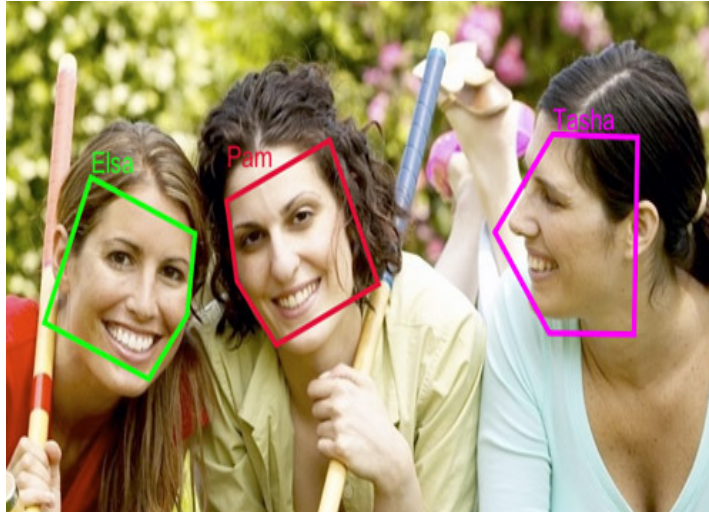
La base de données, pour chaque visage, possède les valeurs suivantes :

Nom et prénom

Adresse

Date et lieu de naissance

Voix enregistrée



Après la détection de visage, nous extrayons le vecteur de caractéristiques en utilisant des motifs binaires locaux.

Le motif binaire Local (LBP) est un opérateur de texture simple mais très efficace qui marque les pixels d'une image par seuillage le voisinage de chaque pixel et considère le résultat comme un nombre binaire.

L'application que nous développons doit tenir compte des besoins entrés. Dans ce qui suit, nous allons identifier les différentes exigences fonctionnelles et non fonctionnelles auxquels notre application doit répondre.

Les avantages du système de reconnaissance du visage

- Le système de reconnaissance faciale est une application informatique ou un dispositif qui permet d'identifier les individus en fonction de leurs caractéristiques uniques du visage. Contrairement à d'autres méthodes d'identification, ils n'ont pas besoin d'entrer en contact direct avec une personne afin de vérifier son identité. Ceci peut être avantageux dans des environnements propres, pour la surveillance ou le suivi, et dans les systèmes d'automation. Au lieu d'exiger d'une personne ses empreintes digitales ou le scan de son iris ou de sa rétine, les caractéristiques faciales peuvent être filmées ou capturés facilement sans nécessiter aucun contact.
- Les taux élevés de reconnaissance et les temps de traitement court font de ces systèmes comme une option efficace

Algorithmes implémentés

Voilà Jones Algorithme

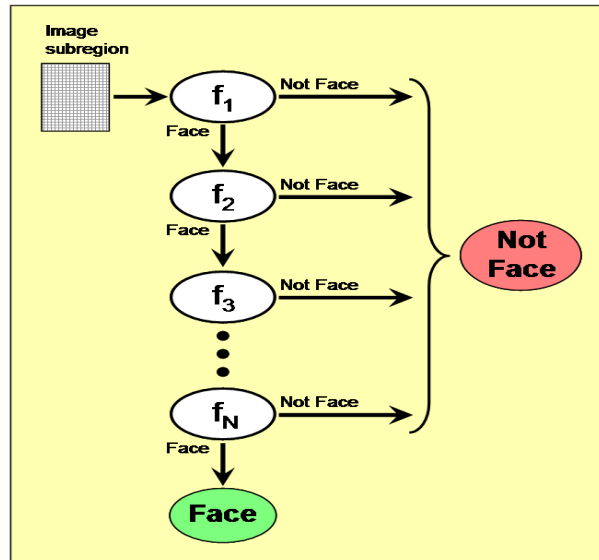
Détection de visage:

La détection de visage est réalisée en utilisant la boîte d'outils de vision de l'ordinateur. L'algorithme qui a été implémenté est l'algorithme de Viola Jones. Les caractéristiques utilisées par le cadre de la détection impliquent universellement les sommes de pixels de l'image au sein des zones rectangulaires. Ces caractéristiques ressemblent quelque peu aux fonctions de base de Haar, qui ont été utilisés précédemment dans le domaine de la détection d'objet basé sur l'image.



AdaBoost Algorithme

- Classificateur fort pour la détection de visage:
- Le cadre de détection d'objet utilise une variante de l'algorithme d'apprentissage d' AdaBoost pour à la fois sélectionner les meilleurs éléments et pour former des classificateurs qui les utilisent.
- L'évaluation des classificateurs forts générés par le processus d'apprentissage peut être faite rapidement, mais ce n'est pas assez rapide pour l'exécuter en temps réel. Pour cette raison, les classificateurs forts sont disposés en cascade par ordre de complexité, où chaque classificateur successif est formé uniquement à partir des échantillons sélectionnés des classificateurs précédents.



LBP Algorithmme

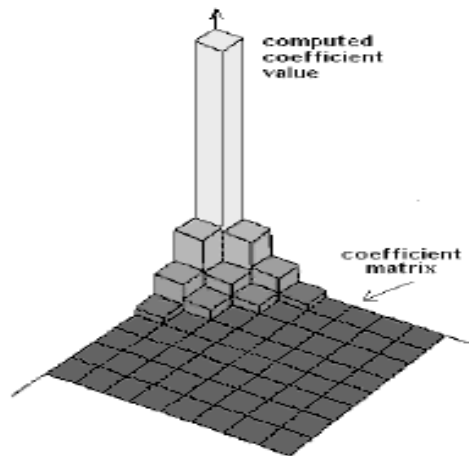
Après la détection de visage, nous extrayons le vecteur de caractéristiques en utilisant des **motifs binaires locaux**.

Le Motif binaire locale (LBP) est un opérateur de texture simple mais très efficace qui marque les pixels d'une image par seuillage du voisinage de chaque pixel et considère le résultat comme un nombre binaire.

Grâce à son pouvoir discriminant et sa simplicité de calcul, l'opérateur de texture LBP est devenu une approche populaire dans diverses applications. Il peut être considéré comme une démarche fédératrice pour les modèles statistiques et structurels traditionnellement divergentes de l'analyse de texture. La caractéristique la plus importante de l'opérateur LBP dans les applications du monde réel est **sa robustesse aux variations de gris monotones causées, par exemple, par des variations d'éclairage**. Une autre propriété importante est sa simplicité de calcul, ce qui permet d'analyser les images en contestant les paramètres en temps réel.

DCT Algorithmme

- Une fois la matrice LBP calculée, on obtient une matrice de la même taille que l'image d'entrée. A cette matrice 2D nous appliquons l'opérateur DCT.
- La transformée bidimensionnelle en cosinus discrète est capable de réduire la dimension du problème puisque une grande partie de l'énergie du signal se situe dans les basses fréquences; ceux-ci apparaissent dans le coin supérieur gauche du DCT



- Les coefficients DCT sont disposés dans un ordre en zigzag, tout comme dans l'algorithme de compression à base JPEG et seulement les coefficients N du début sont pris en compte, les restants étant négligés
- Les autres coefficients sont associés aux hautes fréquences. Les hautes fréquences ont tendance à être principalement bruit.

Normalisation du vecteur caractéristique

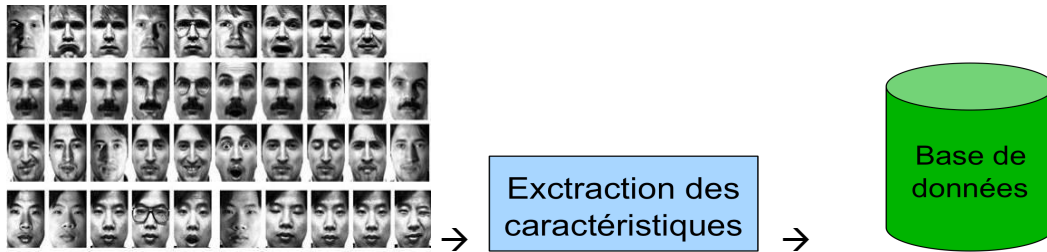
- Après que toutes les caractéristiques des vecteurs ont été extraites des entrées des images faciales, tous les vecteurs 1D sont normalisés à des valeurs minimales et maximales.
- Cette normalisation est nécessaire pour décaler la plage de chaque valeur de coefficients dans l'intervalle [0, 1]
- La mise à l'échelle de caractéristique est utilisée pour normaliser la gamme de variables indépendantes ou de caractéristiques des données. Alors que la plage de valeurs de données brutes est très variable, après une telle normalisation chaque caractéristique contribue approximativement proportionnellement à la distance finale
- Les coefficients de normalisation min et max sont enregistrés pour la reconnaissance

Processus de reconnaissance

- Quand une image d'entrée inconnue doit être identifiée, le vecteur caractéristique est extrait à partir du patch du visage
- Le vecteur est ensuite normalisé en utilisant les coefficients min et max.

- Le vecteur normalisé est comparé avec tous les vecteurs présents dans la base. La mesure qui a été adoptée est la distance de Manhattan
- Si la distance est inférieure à une valeur seuil, l'image faciale est reconnue, sinon elle est rejetée.

Génération de base de données

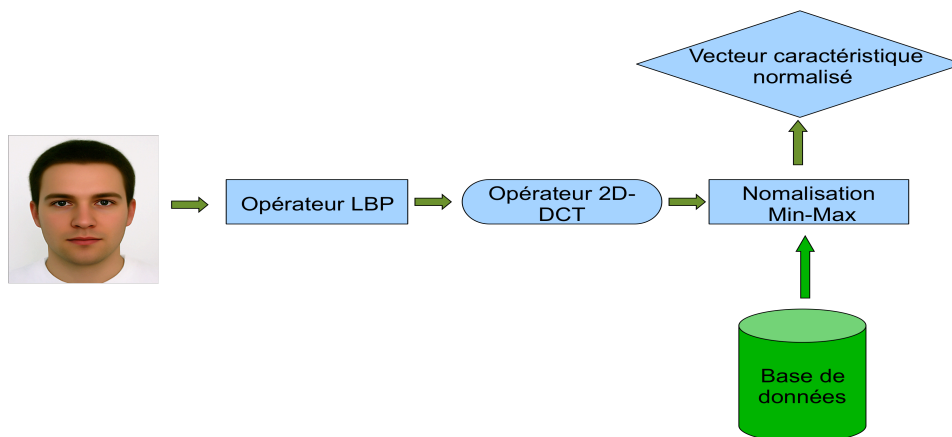


Pour chaque valeur min et max de caractéristique sont calculés les normalisations min et max.

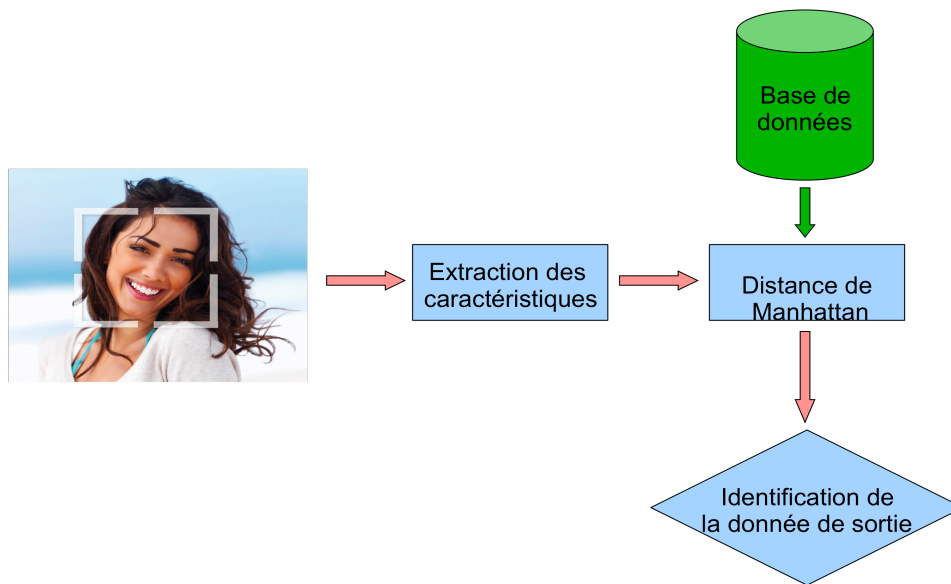
Vecteur caractéristique :

- nom,
- prénom,
- date et lieu de naissance,
- voix enregistrée

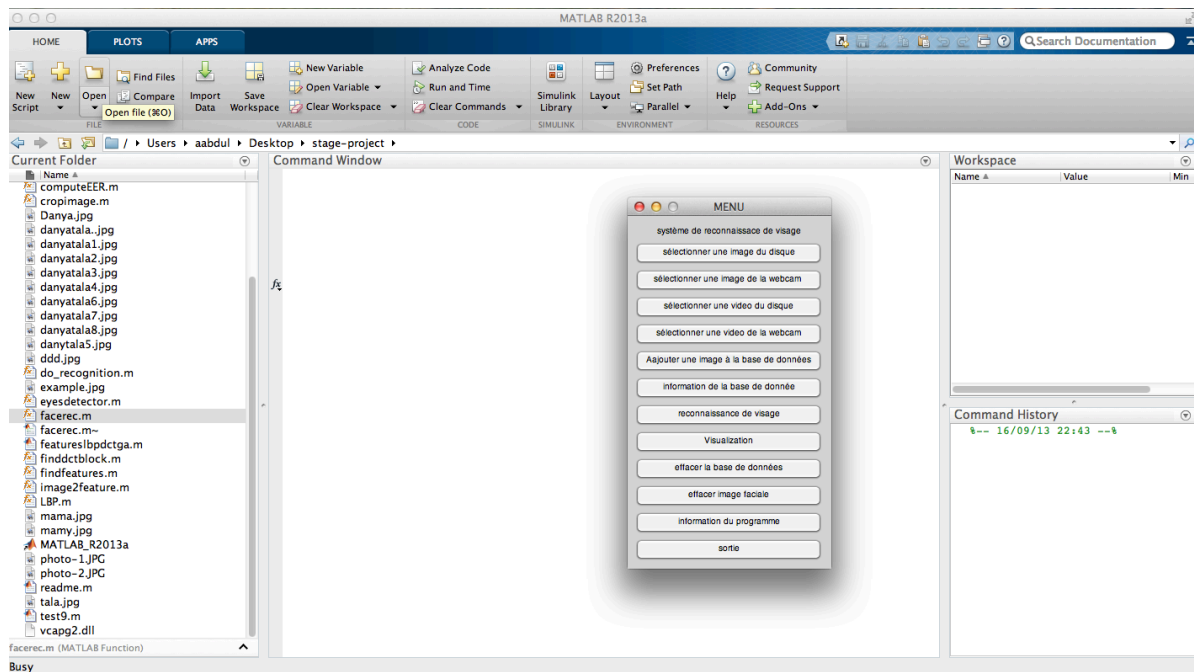
Processus d'extraction de caractéristiques



Processus d'indentification du visage



Interface d'utilisateur graphique pour le système de reconnaissance facial :



Acquisition d'image ou de vidéo et reconnaissance faciale



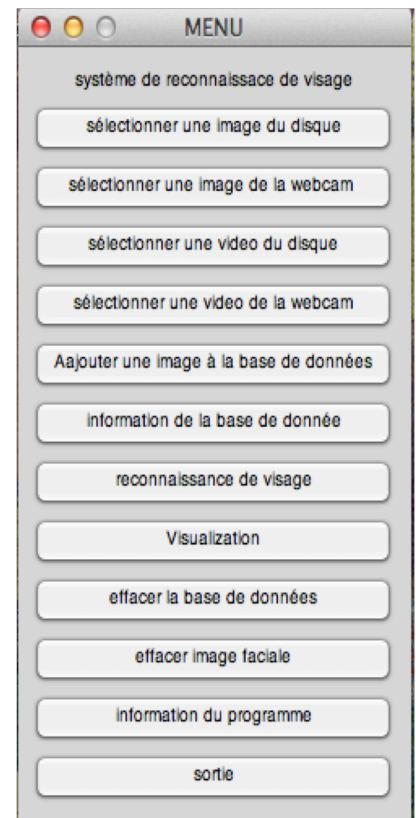
enregistrer le visage dans la base de données après extraction du vecteur caractéristique et afficher les informations de la base de données



reconnaissance faciale



les options de la base de données et les informations GUI



Résultats

Les images ont été prises à des moments différents, allant de différents éclairages, à des expressions du visage différentes (yeux ouverts / fermés, souriant / pas de sourire, le clignotement des yeux,...)

Les résultats sont :

- Une à une : pas d'erreur
- Selon le nombre extrait d'images dans chaque classificateur, les résultats montrent une progression. Les classificateurs de 4 extractions dans lequel apparaissent un taux de 45% d'erreurs de plus que 19 extractions qui donnent un résultat d'erreur inférieure à 4%

Perspectives :

Améliorations possibles :

- Amélioration de l'algorithme de détection des visages pour une meilleure localisation des yeux

La combinaison de classificateur de visage avec un classificateur vocal pour un système de reconnaissance multi biométriques

- L'application a été développée sur la base du comportement de conception, où l'objectif futur sera d'augmenter les performances et l'efficacité pour élargir l'application, y compris aux utilisateurs ayant des besoins spéciaux et précisément aux personnes aveugles.

Conclusions

Les systèmes de reconnaissance du visage utilisés aujourd'hui fonctionnent mais dans des conditions limitées, bien que tous les systèmes fonctionnent beaucoup mieux avec des images de face et de profil et un éclairage constant. Tous les algorithmes actuels de reconnaissance de visage ne parviennent pas à identifier des visages dans les conditions largement variables en vertu de laquelle les humains sont capables d'identifier d'autres personnes. Les systèmes de reconnaissance de personne de la prochaine génération devront reconnaître les gens en temps réel et dans des situations beaucoup moins contraintes. Nous croyons que les systèmes d'identification qui sont robustes dans les milieux naturels, en présence de changements de bruit et d'éclairage, ne pourront pas compter sur une seule modalité, ce qui signifie que la fusion avec d'autres modalités sera essentielle. Compte tenu de toutes les exigences, les systèmes d'identification qui utilisent la reconnaissance faciale et l'identification vocale nous semblent avoir le plus de potentiel pour des applications très répandues.

Références Bibliographiques

- [1] [INNOROBO 2012 : Innovation, handicap et accessibilité: http://www.clubic.com/robotique/actualite-482172-innorobo-2012-robotique-innovante-handicap.html#ixzz2ebeV0EZ1](http://www.clubic.com/robotique/actualite-482172-innorobo-2012-robotique-innovante-handicap.html#ixzz2ebeV0EZ1)
- [2] <http://www.news-medical.net/health/What-is-visual-impairment.aspx>
- [3] <http://cimss.ssec.wisc.edu/wxwise/class/aos340/spr00/whatismatlab.htm>
- [4] <http://cimss.ssec.wisc.edu/wxwise/class/aos340/spr00/whatismatlab.htm>
- [5] http://www.mathworks.fr/fr/help/matlab/creating_plots/figures-plots-and-graphs.html
- [6] http://www.igp.ethz.ch/photogrammetry/education/lehrveranstaltungen/photogrammetry/matlab_imageprocessing.pdf
- [7] http://en.wikipedia.org/wiki/Viola-Jones_object_detection_framework
- [8] http://www.scholarpedia.org/article/Local_Binary_Patterns
- [9] Ojala, T., Pietikäinen, M. and Harwood, D. (1996), A Comparative Study of Texture Measures with Classification Based on Feature Distributions. *Pattern Recognition* 19(3):51-59.
- [10] Ojala, T. and Pietikäinen, M. (1999), Unsupervised Texture Segmentation Using Feature Distributions. *Pattern Recognition* 32:477-486.
- [11] Ojala, T., Pietikäinen, M. and Mäenpää, T. (2002), Multiresolution Gray-scale and Rotation Invariant Texture Classification with Local Binary Patterns. *IEEE Trans. Pattern Analysis and Machine Intelligence* 24(7): 971-987.
- [13] Zhao, G. and Pietikäinen, M. (2007), Dynamic Texture Recognition Using Local Binary Patterns with an Application to Facial Expressions. *IEEE Trans. Pattern Analysis and Machine Intelligence* 29(6):915-928.
- [14] Ahonen, T., Hadid, A. and Pietikäinen, M. (2006), Face Description with Local Binary Patterns: Application to Face Recognition. *IEEE Trans. Pattern Analysis and Machine Intelligence* 28(12):2037-2041.
- [15] Li, S.Z., Chu, R.F., Liao, S.C. and Zhang, L.(2007), Illumination Invariant Face Recognition Using Near-Infrared Images. *IEEE Trans. Pattern Analysis and Machine Intelligence* 29(4):627-639.
- [16] Zhang, W., Shan, S., Gao, W., Chen, X. and Zhang H. (2005), Local Gabor Binary Pattern Histogram Sequence (LGBPHS): A Novel Non-statistical Model for Face

- Representation and Recognition. In Proc. Tenth IEEE International Conference on Computer Vision, 1:786-791.
- [17] Hadid, A. and Pietikäinen, M. (2009), Combining Appearance and Motion for Face and Gender Recognition from Videos. *Pattern Recognition* 42(11):2818-2827.
- [18] Zhao, G., Barnard, M. and Pietikäinen, M. (2009), Lipreading with Local Spatiotemporal Descriptors. *IEEE Trans. Multimedia* 11(7):1254-1265.
- [19] Pietikäinen, M., Hadid, A., Zhao, G. and Ahonen, T. (2011), *Computer Vision Using Local Binary Patterns*, Springer, 207
- [20] Liao, S., Law M.W.K. and Chung A.C.S. (2009), Dominant Local Binary Patterns for Texture Classification. *IEEE Trans. Image Processing* 18(5):1107-1118.
- [21] Heikkilä, M., Pietikäinen, M. and Schmid, C. (2009), Description of Interest Regions with Local Binary Patterns. *Pattern Recognition* 42(3):425-436.
- [22] Mäenpää, T. and Pietikäinen, M. (2004), Classification with Color and Texture: Jointly or Separately? *Pattern Recognition* 37(8):1629-1640.
- [23] Tan X. and Triggs B. (2007), Fusing Gabor and LBP feature Sets for Kernel-based Face Recognition, *Proceedings of the IEEE International Workshop on Analysis and Modeling of Face and Gesture*, pp.235-249.
- [24] Wang J.-G., Yau W.-Y. and Wang H. L. (2009), Age Categorization via ECOC with Fused Gabor and LBP Features. *Proceedings of the IEEE Workshop on Applications of Computer Vision (WACV)*, pp.313-318
- [25] Heikkilä, M. and Pietikäinen, M. (2006), A Texture-Based Method for Modeling the Background and Detecting Moving Objects. *IEEE Trans. Pattern Analysis and Machine Intelligence* 28(4):657-662.
- [26] Heusch G., Rodriguez Y. and Marcel S. (2006), Local Binary Patterns as an Image Processing for Face Authentication, *Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition*.
- [27] Kellokumpu, V., Zhao, G. and Pietikäinen, M. (2010), Recognition of Human Actions Using Texture Descriptors. *Machine Vision and Applications*.
- [28] <http://www.mathworks.fr/fr/help/vision/ref/vision.pointtrackerclass.html>