

*République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université des Sciences et de la Technologie Houari Boumediene
Faculté d'Electronique et d' Informatique
Département d'Informatique*

Traitement et Analyse d'Images et de la Vidéo

Master Informatique Visuelle

Introduction à la Vision par Ordinateur

Cours de Traitement et Analyse d'images et de la vidéo
Master MIV, TAI- par N.BAHA

Introduction

- L'univers dans lequel nous évoluons est un espace tridimensionnel (3D). Une personne à qui on demanderait soit de décrire ce qu'elle voit, ou de se déplacer dans ce monde n'aurait aucune difficulté à le faire. Pourtant l'information qui est réellement disponible sur la rétine de ses yeux n'est, ni plus ni moins, une collection de points tel que chaque point est la projection d'un fragment de la scène vue par l'oeil portant aussi une information sur la quantité de lumière (quantité d'énergie) et la couleur qui proviennent de l'espace environnant.
- On pourrait alors faire correspondre cette donnée à une représentation.
- De cette différence entre données collectées et interprétation fournie par la faculté sensorielle de vision, on est amené rapidement à tirer une déduction logique. **Le processus visuel permet une interprétation non ambiguë** (reconnaissance) des objets et des scènes 3D à partir d'information sensorielle (la lumière et la couleur).

Introduction

- La vision a été et restera toujours un grand mystère pour les philosophes et biologistes, malgré toutes les expériences et études menées pour comprendre l'anatomie et le fonctionnement du système visuel en particulier: le cerveau.
- La tâche s'est avérée très complexe et illusoire.
 - Complexe, car la création divine surpasse toute science et
 - illusoire car ses limites ne sont pas connues.

David Hubel a merveilleusement exprimé cette conquête [Hubel79]:

le cerveau peut-il comprendre le cerveau ?

Introduction

- De nombreuses recherches ont été menées dans le domaine de la vision, afin de comprendre le fonctionnement de ce mécanisme de perception complexe, et d'essayer de le reproduire sur des machines.
- Cela a donné naissance à de nombreuses études qui ne permettent malheureusement pas encore la reconnaissance d'un environnement complet sans restriction ou connaissance à priori.
- La vision par ordinateur a donc pour but d'essayer de créer un système artificiel de vision, à un niveau de complexité moindre, mais le plus autonome possible.

Historique

- Télévision : 1935 analogique, TNT 2005
- Imagerie médicale : 1960 premier échographe; échographe numérique au milieu des années 90
- Imagerie satellitaire : pellicules puis premiers satellites numériques (espions) en 1976
- Photos : photo argentique (ordinaires) 1858, brevet appareil numérique 1972, premiers modèles 91
- 1960 : reconnaissance de caractères dactylographiés
- 1970 : imagerie médicale, imagerie satellitaire, télévision

Historique

- David Marr [Mar 82] s'est inspiré des résultats de Neurophysiologie pour proposer un modèle calculatoire pour la Vision par Ordinateur. Ce modèle s'appuie sur le fait que la perception du monde par la vision peut être considérée comme des processus de traitement d'informations symboliques.
- Années 1990: reconnaissance de visage;
- Années 2000: de la reconnaissance plus large; grands ensembles de données annotées disponibles; traitement vidéo commence à faire ses débuts
- Soulèvement des robots 2030!!!!
- À l'heure actuelle : des images numériques partout...

C'est quoi la vision par ordinateur



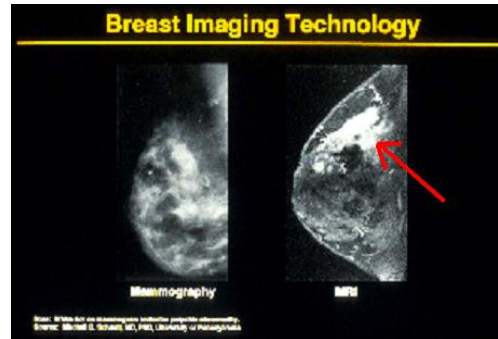
Réponse:

Rendre les ordinateurs capables de comprendre les images et les vidéos.

Pourquoi la VO est importante?



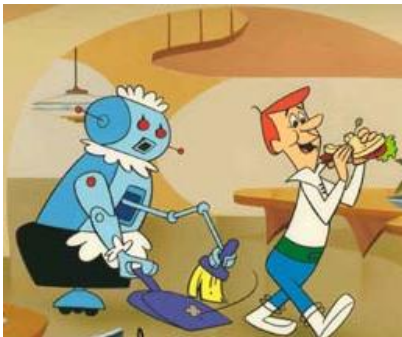
Sécurité



Médical



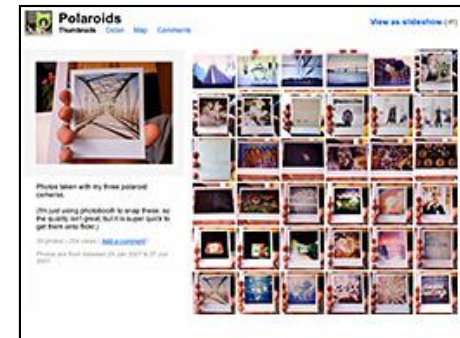
sûreté



Confort



S'amuser

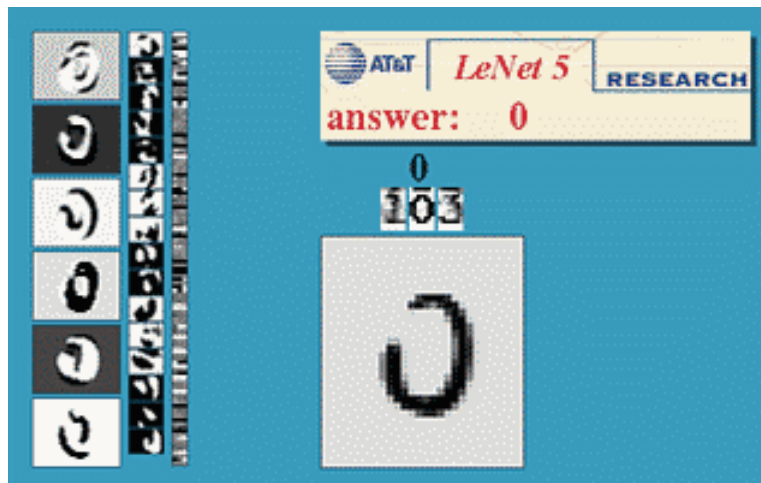


Accès

Comment est utilisée la VO actuellement?

Exemples de l'état de l'art:

- Reconnaissance des caractères



Digit recognition, AT&T labs
<http://www.research.att.com/~yann/>

Comment est utilisée la VO actuellement?

- **Face detection:** actuellement, la plupart des appareils photo détectent les visages: Canon, Sony, Fuji,

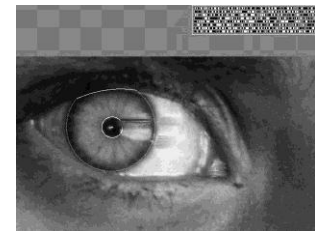


Comment est utilisée la VO actuellement?

- Reconnaissance d'objets dans les super marchés



- **Biometrie:** reconnaissance de l'Iris,
Reconnaissance des empreintes



Comment est utilisée la VO actuellement?

Capture du mouvement :
motion capture



Google Car: Oct 9, 2010. ["Google Cars Drive Themselves, in Traffic"](#). [The New York Times](#). John Markoff



Comment est utilisée la VO actuellement?

Robotique pour l'industrie



Comment est utilisée la VO actuellement?

Imagerie Médicale

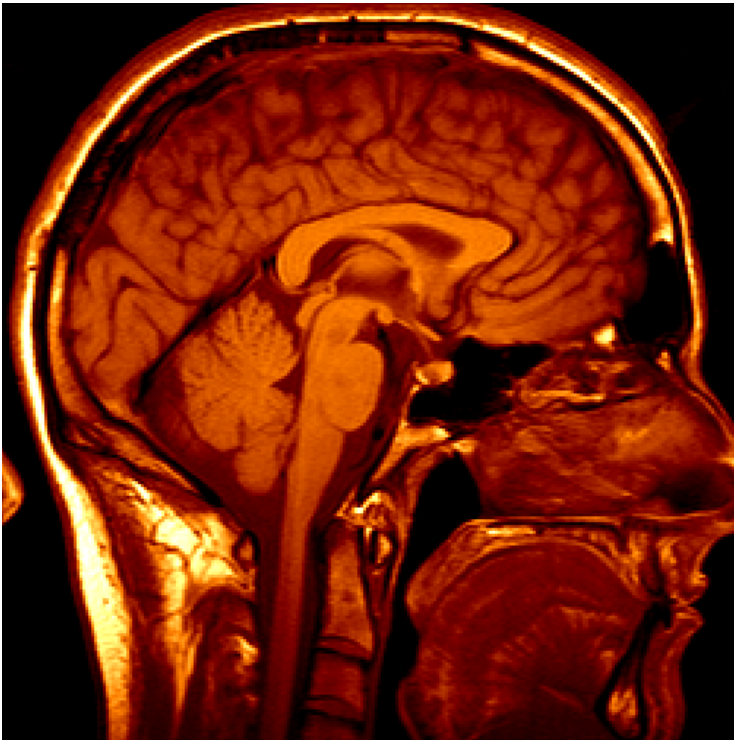
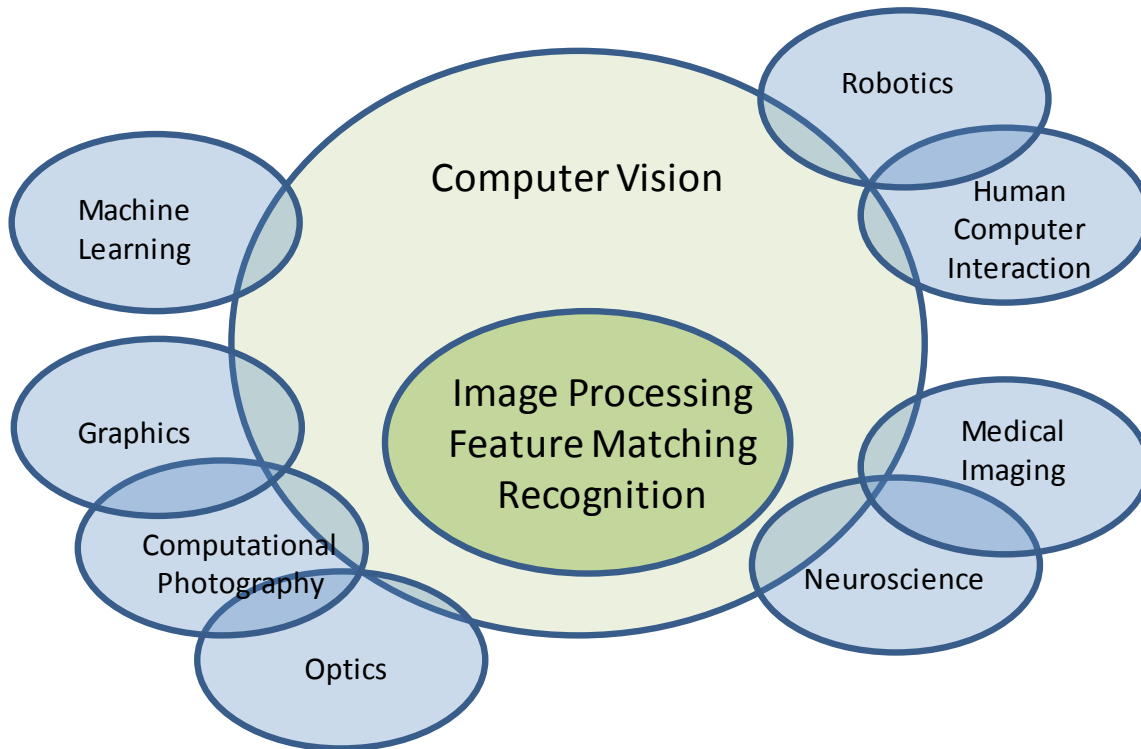


Image guided surgery
[Grimson et al., MIT](#)

Exemples d'états de l'art utilisant la VO



Vision par ordinateur

- Le capteur visuel de l'être humain constitue une façon quasi complète de communication avec le monde extérieur.
- La capacité de ce capteur a fasciné l'homme et l'a poussé à tenter de le reproduire .



- **La vision artificielle?**

Employer des moyens informatiques et mathématiques pour aboutir à des résultats similaires à ceux parvenus par la vision de l'être humain.

En entrée, le système dispose:

- D'images de la scène prises par un ensemble de capteurs,
- D'une base de connaissances acquise et enrichie dans le temps.

En sortie, le système doit être capable de reconnaître et situer relativement les objets composant la scène, en temps réel.

Reconnaissance d'objets



Pourquoi la VO est difficile?

- Challenges :

- Point de vues différent



Pourquoi la VO est difficile?

➤ Illumination



Pourquoi la VO est difficile?

➤ Echelle: Scaling

scale



slide credit: Fei-Fei, Fergus & Torralba

Pourquoi la VO est difficile?

Arrière plan complexe:
Complex Background



Pourquoi la VO est difficile?

➤ Variation des objets de la même classe(intra-classe)



Pourquoi la VO est difficile?

➤ Problème de l'ombre (shadow)



Vision par ordinateur

Ces systèmes peuvent être divisés en deux grandes classes:

- **Les systèmes de vision bidimensionnelle (2D)**

- Les systèmes de vision bidimensionnelle ne prennent en considération que deux dimensions, la notion de profondeur n'a donc pas d'intérêt.
- Ils sont utilisés dans divers domaines tels que:
 - l'usage biomédical,
 - spatial,
 - la télédétection,
 - la robotique industrielle ou encore dans la reconnaissance de caractères.

Vision par ordinateur

- **Les systèmes de vision tridimensionnelle (3D)**
 - Les systèmes de vision tridimensionnelle intègrent la troisième dimension. La profondeur est donc retrouvée grâce à la prise de plusieurs vues et permet d'obtenir une description 3D de la scène.
 - Ces systèmes sont utilisés dans plusieurs applications telle que:
 - la robotique mobile (la saisie d'objets, l'évitement d'obstacles...).
 - La reconstruction 3D.

Vision par Ordinateur

Nous pouvons distinguer deux classes de systèmes tridimensionnels :

- Les systèmes de vision actifs.
- Les systèmes de vision passifs.

Vision par Ordinateur

➤ Les systèmes de vision actifs

Un système de vision actif est basé sur l'utilisation de systèmes dotés d'un émetteur et un récepteur. L'émetteur envoie vers la scène un faisceau d'ondes radio (radar), d'ondes sonores (sonar) ou de lumière (lidar), dont les éventuelles rétrodiffusions vers le récepteur permettent de localiser les objets présents dans la scène. En mesurant le temps d'aller-retour du faisceau émis, on peut alors connaître la distance des objets par rapport au capteur.

- Utilisés surtout pour les systèmes de détection d'obstacles.
- Les performances de ces systèmes peuvent être réduites par certains problèmes liés:
 - au milieu de propagation qui absorbe et/ ou disperse les faisceaux aller et retour,
 - à la surface réfléchissante de la cible qui peut absorber le faisceau, le diffuser ou le réfléchir dans d'autres directions.

Les systèmes actifs

Comment retrouver la profondeur?

La triangulation : le principe se fonde sur la connaissance de la distance entre le récepteur et l'émetteur, ainsi on peut déduire la profondeur du point qui réfracte l'onde de la direction du signal émis et réfracté.

Le temps de vol : on peut calculer la profondeur du point par la mesure du temps mis par le signal entre l'émetteur et le récepteur.

Les systèmes de vision passif

- Un système de vision passif est un système qui utilise uniquement les images de la scène prises sous des angles différents pour la reconstitution de l'information 3D. Ce sont des caméras numériques qui jouent le rôle de capteurs.
- Le calcul de la profondeur des objets est réalisée grâce à des caractéristiques de l'image telles que les discontinuités photométriques qui permettent d'extraire des points, des contours et des régions.

Modèle De David Marr

➤ Premier niveau : **vision**

Acquiert une image grâce à une chaîne intégrant des éléments optiques et un capteur.

➤ Deuxième niveau : **traitement de l'image**

Modifie le contenu de l'image afin de mettre en évidence des éléments d'intérêt (objets, contours)

➤ Troisième niveau : **reconnaissance**

Utilise des techniques d'intelligence artificielle pour identifier des formes connues dans l'image.

Points → contours → objects 3D

- **Perception 3D = Reconstruction 3D**

- Cette approche est encore implicite dans de nombreux travaux

Objectifs du cours

- Ce cours a pour but de vous introduire les notions fondamentales de la vision par ordinateur. Plus précisément, son objectif est de vous aider à acquérir une certaine intuition du:
 - traitement d'images ,
 - flux vidéo,

Objectifs du cours

- Amélioration d'images
- Codage d'images
- Analyse d'images
- Analyse de la vidéo

- Trois categories de transformations:

➤ **Le traitement des images:** image \longrightarrow image
«**image processing**»

➤ **L'analyse d'images:** image \longrightarrow mesures
«**image analysis**»

➤ **L'analyse sémantique:** image \longrightarrow description de
haut niveau «**image understanding**»

- **L'analyse d'images:** utiliser un ordinateur pour interpréter le monde extérieur en utilisant des images.
- **Synthèse d'images (Computer Graphics) :** utiliser un ordinateur pour générer des images.

- Différentes phases de l'analyse d'une image :
 1. Acquisition
 2. Traitement de bas niveaux : filtrage et extraction d'indices.
 3. Traitement de haut niveaux : reconnaissance des formes, reconstruction, calculs de positions, calculs de mouvements.
- Éléments impliqués :
 1. Les primitives images : pixels, points d'intérêts, segments, contours.
 2. Les caractéristiques photométriques : niveaux de gris, de couleurs.
 3. Les caractéristiques géométriques : caméras, mouvements.