



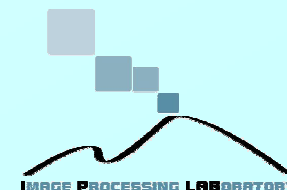
Computer Vision

**Corso di Laurea Specialistica in
Informatica**

A.A. 2008/2009

Sebastiano Battiato

Mercoledì (9.00 -11.00) Venerdì (10.00 – 12.00) – Aula 2





What is Computer Vision?

- La **Visione** è forse il senso più importante che l'uomo possiede. Essa permette di inferire il mondo tridimensionale, di riconoscere e localizzare gli oggetti presenti in una scena, di percepire i rapidi cambiamenti dell'ambiente, ecc.
- La **Computer Vision** è la disciplina che studia come abilitare i computer alla comprensione e alla interpretazione delle informazioni visuali presenti in immagini o video.



Computer Vision

Tra tutte le abilità sensoriali, la **visione** è largamente riconosciuta come quella con le maggiori potenzialità. Le capacità dei sistemi biologici sono formidabili: l'occhio raccoglie una banda di radiazioni elettromagnetiche rimbalzate su diverse superfici e provenienti da fonti luminose diverse ed il cervello elabora questa informazione formando il quadro della scena come noi la percepiamo.

Se volessimo dare una definizione, potremmo dire che la **Visione Computazionale (VC)** o **Computer Vision**, si occupa della **analisi** di immagini numeriche al calcolatore.



Computer Vision

L'analisi è finalizzata a scoprire **cosa** e presente nella scena e **dove**. Non si occupa di:

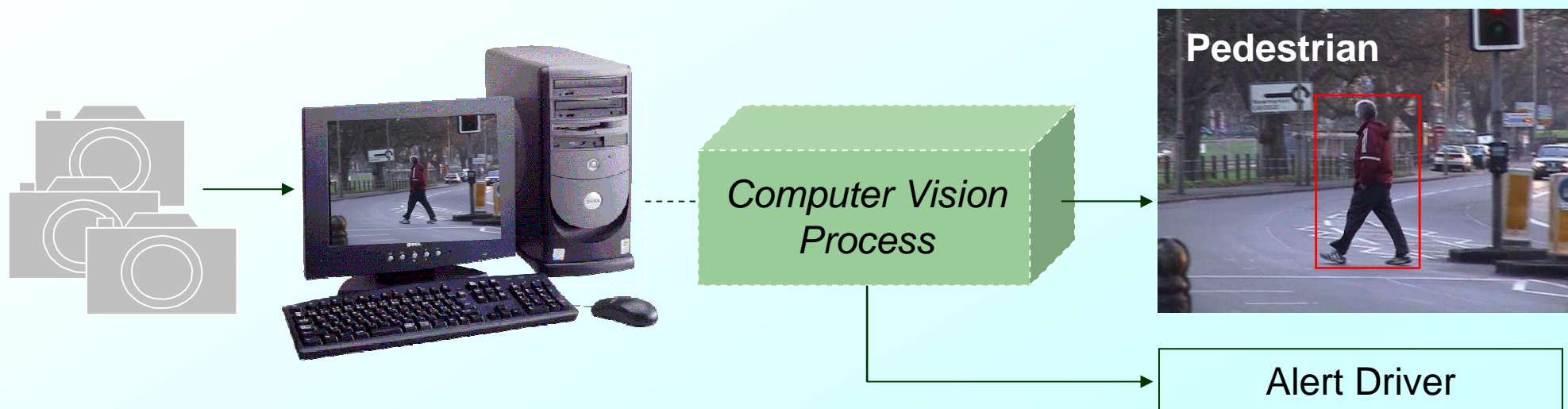
Elaborazione di immagini: miglioramento, restauro e compressione di immagini. Si elabora una immagine per ottenerne un'altra in qualche senso "migliore";

Riconoscimento di pattern: (estrazione), identificazione, classificazione di caratteristiche nelle immagini.

Computer Vision \neq Pattern Recognition
Computer Vision \neq Image Processing

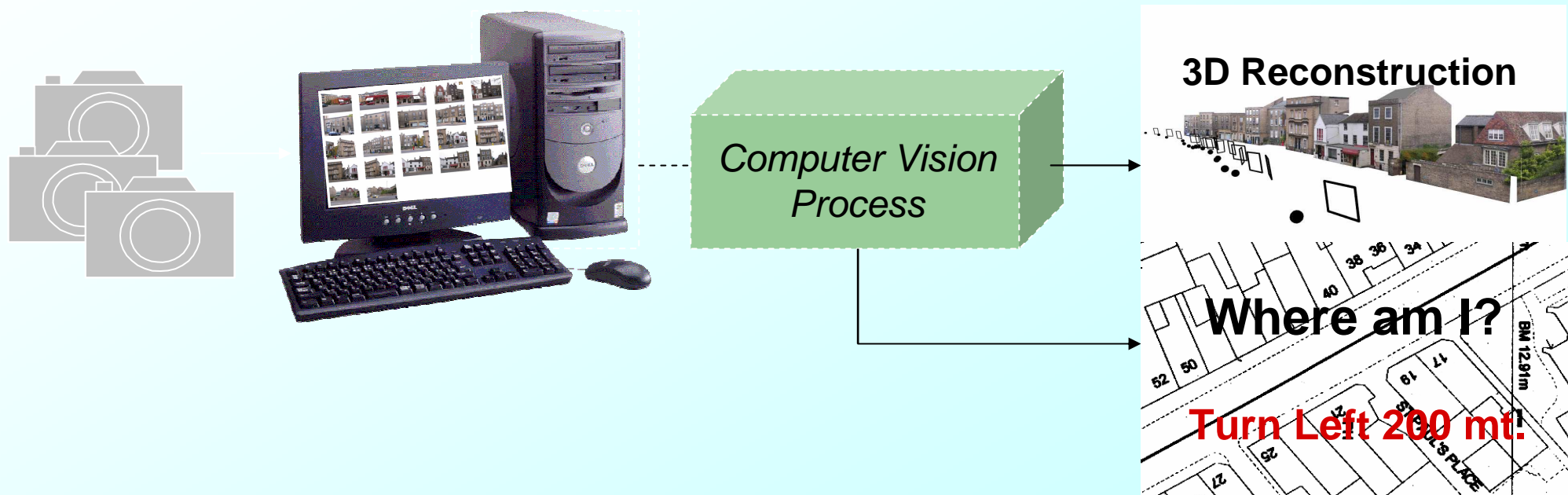
Computer Vision Systems (1)

Un computer elabora immagini di una scena reale, catturata da una o più telecamere, per ottenere informazioni utili (li interpreta) che saranno utilizzate al fine di prendere decisioni in maniera automatica o semiautomatica.



Computer Vision Systems (2)

Un computer elabora immagini di una scena reale, catturata da una o più telecamere, per ottenere una “**interpretazione**” 3D della scena reale. Anche in questo caso sulla base di tale informazione si possono poi prendere “**decisioni**” in maniera automatica e/o semiautomatica.





Computer Vision Goals

- **Costruire sistemi capaci di prendere decisioni a partire da una descrizione della scena estrapolata da immagini/video;**
- **Inferire il mondo 3D a partire da immagini digitali;**
- **Riconoscimento di oggetti, scene, contesto a partire da immagini digitali.**
- **.....**

Applications: The Computer Vision Industry (1)

- **Automobile driver assistance**
 - Systems that warn automobile drivers of danger, provide adaptive cruise control, and give driver assistance.
- **Automobile traffic management**
 - Systems for reading automobile license plates.
- **Film and Television**
 - Systems for tracking objects in video or film action to provide enhanced broadcasts.
- **General purpose vision systems**
 - Vision systems for object recognition and navigation. Applications include mobile robotics, grocery retail, and recognition from cell phone cameras.
- **Image search**
 - Image retrieval based on content.
- **Industrial automation and inspection**
 - **Automotive industry:** Systems for vision-guided robotics in the automotive industry.
 - **Electronics industry:** Electronics inspection systems for component assembly and semiconductor manufacturing.
 - **Food and agriculture:** Vision systems for inspecting and grading fruits and vegetables.
 - **Printing and textiles:** Inspection for the printing and packaging industries.



Applications: The Computer Vision Industry (2)

- **Medical and biomedical**
 - Uses real-time stereo vision to detect and track the pose of markers for surgical applications.
- **Pedestrian tracking**
 - Systems for counting and tracking pedestrians using overhead cameras.
- **Safety monitoring**
 - System monitors swimming pools to warn of accidents and drowning victims.
- **Security**
 - Vision systems for video surveillance, including tracking, object monitoring, and behavior analysis.
- **Biometric**
 - Systems for Fingerprint recognition and biometric face recognition
- **Three-dimensional modeling**
 - Creation of texture-mapped 3-D models from a small number of photographs.
- **Video Games**
 - Interactive advertising for projected displays that tracks human gestures.



Videos Examples

- Object Classification
- Automatic Object Detection and Recognition
- Pedestrian Detection
- Pedestrian Detection in Crowds
- Face Tracking
- Body Tracking
- People Counting in store
- In/out People counting
- Detection of scene in video
- Detection of Actions in Video
- Detection of independent motion in Crowds
- 3D city modelling from photos
- 3D bone classification and Reconstruction
- 3D from single photo
- 3D Object Modelling from images

Demos

- Visualizzazione di fotografie in un ambiente 3D “virtuale”
 - <http://labs.live.com/photosynth/collectionHome.htm>





Links in Rete

- The Computer Vision Home Page
 - <http://www.cs.cmu.edu/~cil/vision.html>
- Computer Vision Education
 - <http://www.cved.org/>
- The Computer Vision Industry
 - <http://www.cs.ubc.ca/spider/lowe/vision.html>
- CVOnline
 - <http://homepages.inf.ed.ac.uk/rbf/CVonline/>

Programma di massima del Corso (1/2)

Il corso si propone di approfondire teorie e tecniche specificatamente rivolte alla visione artificiale con una serie di applicazioni.

La prima parte del corso verterà su:

- Modelli di Formazione dell'Immagine: Camera Calibration
- Filters e Features
 - Edge, Linee, Trasformata di Hough
 - Piramidi Laplaciane
 - Corner Detection (Harris, ...)
 - SIFT: Teoria e Applicazioni
- Tecniche di segmentazione:
 - Thresholding
 - Seeded Region Growing
 - Statistical Region Merging, ...





Programma di massima del Corso (2/2)

La seconda parte verterà su:

- Modelli probabilistici applicati alla Visione

Con alcuni casi di studio e applicazioni

CBIR Retrieval

Video Stabilization

L'ultima parte del corso è dedicata (come ogni anno) ad un tema "specialistico" d'approfondimento.



Books

- D. A. Forsyth, J. Ponce, *“Computer Vision a Modern Approach”*, Prentice Hall PTR, 2002
- E. Trucco, A. Verri, *“Introductory Techniques for 3-D Computer Vision”*, Prentice Hall, 1998
- Gonzalez and Woods, *“Elaborazione delle Immagini Digitali”*, PBM, Terza Edizione, 2008
- L. Shapiro, G. Stockman, *“Computer Vision”*, Prentice Hall PTR, 2001
- Mubarak Shah, *“Fundamentals of Computer Vision”* (On-Line), 1997
- S. Battiato, F. Stanco, *“Fondamenti di Image Processing”*, EdiArgo 2006
- R. Hartley and A. Zisserman, *“Multiple View Geometry in Computer Vision”*, 2004
- R. O. Duda, P. E. Hart, and D. G. Stork, *“Pattern Classification”*, Wiley Interscience, 2001
- C. M. Bishop, *“Pattern Recognition and Machine Learning”*, 2006

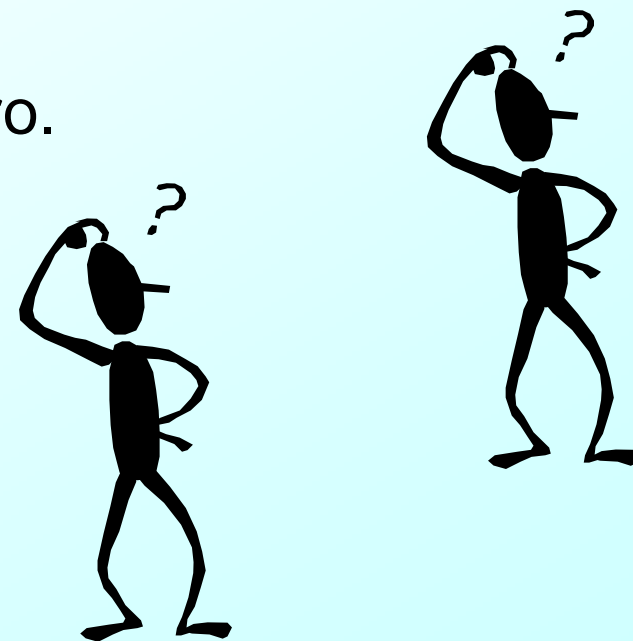


Modalità d'esame

Progetto SW personale (in JAVA ?) da concordare con il docente.

Prove in Itinere (almeno una) con esonero.

Colloquio Orale
comprendente la Demo del progetto



Utility

Slides e Materiale Vario:

www.dmi.unict.it/~battiato/cvision0809/cvision.htm

Forum su Uniweb

E-mail:

battiato@dmf.unict.it

Ricevimento:

(Consultare il web)





Outline del corso

- Introduzione
- Camera Calibration
- Low Level Computer Vision
 - Edge, Linee, Texture, Corner
 - SIFT: Teoria ed Applicazioni
- Tecniche di Segmentazione applicate alle immagini digitali

- Modelli probabilistici applicati alla Visione
- Applicazioni
 - Video Stabilization
 - CBIR Retrieval
 - ...
- Tema “Specialistico” d’approfondimento

Possibili Seminari da parte di esperti esterni



Computer Vision

Distinguiamo la CV di **basso livello** e di **alto livello**.

La prima si occupa di estrarre determinate proprietà fisiche dell'ambiente visibile, come profondità, forma tridimensionale, contorni degli oggetti.

I processi di visione di **basso livello** sono tipicamente paralleli, spazialmente uniformi e relativamente indipendenti dal problema e dalla conoscenza a priori associata a particolari oggetti.



Computer Vision

Viceversa, la visione di **alto livello** si occupa della *estrazione* delle proprietà delle forme e di relazioni spaziali, di *riconoscimento* e *classificazione* di oggetti. I processi di alto livello sono di solito applicati ad una porzione dell'immagine, dipendono dall'obiettivo della computazione e dalla conoscenza a priori associata agli oggetti.



- **Geometria della scena**
 - forma e posizione delle superfici
- **Fotometria della scena**
 - tipo, intensità e direzione dell'illuminazione
 - proprietà di riflettanza delle superfici visibili
- **Modello del sensore**
 - ottico/geometrico
 - fotometrico
 - elettronico
- ...

Grafica
→

Visione
←

Immagine/i



Problemi tipici (1)

- *Condizioni di illuminazione* che producono una variazione nella distribuzione dell'intensità luminosa della scena.
- *Trasformazioni geometriche rigide* dell'oggetto (in ordine di difficoltà crescente):
 - *roto-traslazioni e variazioni di scala in 2D (e in 3D).*
- *Rumore.*
- *Gap*: tipo particolare di rumore consistente nella mancanza di elementi nell'immagine.
- *Occlusione.*



Problemi tipici (2)

- *Segmentazione*: partizionamento dei dati di input in entità semantiche distinte (linee, regioni, oggetti).
- *Indexing*: effettuare una ricerca efficiente in un catalogo di modelli.
- *Identificazione*: riconoscere l'istanza di un oggetto in un'immagine.
- *Oggetti non rigidi* (forbici, volti umani, ...). Il loro riconoscimento è complicato dalla possibilità che ha la loro forma di variare.
- *Classificazione*: riconoscere l'appartenenza ad una data classe di un oggetto in un'immagine.



Calendario di massima

Camera Calibration (con il dott. Eugenio Rustico, circa 6 lezioni a partire dalla prossima)

Low Level Vision

Mid Level Vision: Tecniche di Segmentazione

Applicazioni

