



Introduzione alla visione artificiale (computer vision)

Alberto Borghese



Parleremo di:

- Introduzione & Applicazioni
- Determinazione del campo di moto dal flusso ottico.
- Structure from Motion da sequenze monoculari (elaborazione flusso ottico).
- Structure from Motion da sequenze monoculari (elaborazione di features).



Computer Vision



Obiettivo: determinazione delle proprietà geometriche, fisiche e dinamiche del mondo che ci circonda mediante elaborazione di immagini o sequenze di immagini.

- *Low level vision (o early vision):* estrazione dalle immagini o sequenze di immagini delle informazioni necessarie al livello superiore.
- *High level vision:* riconoscimento, associazione di un significato semantico all'atto del vedere.



Applicazioni: Controllo industriale

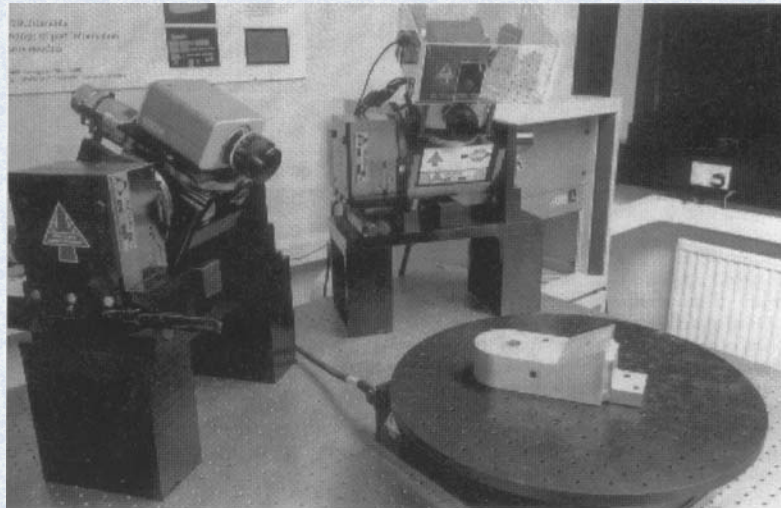


Figure 1.1 A prototype of 3-D inspection cell. The cell includes two types of depth sensors, a laser scanner, and a Moiré fringe system (see Chapter 2), which locate the object in space and perform measurements. Notice the turntable for optimal, automatic object positioning.



Applicazioni: Video-sorveglianza

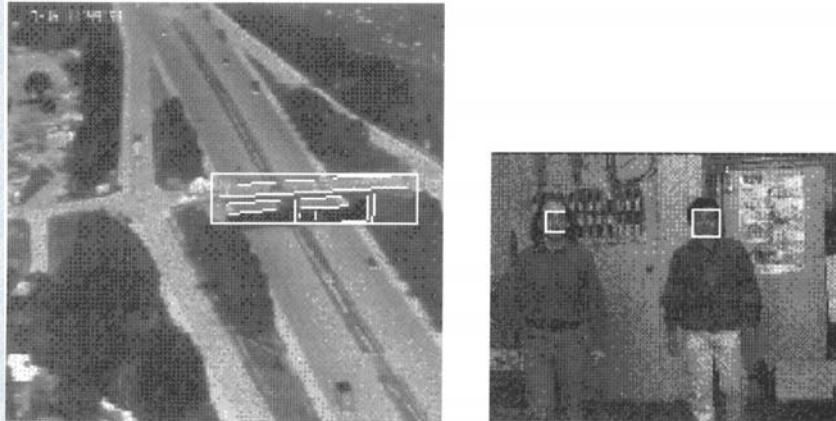


Figure 1.2 Left: automatic recognition of road bridges in aerial infrared images (courtesy of Majid Mirmehdi, University of Surrey; Crown copyright reproduced with the permission of the Controller of Her Majesty's Stationery Office). Right: an example of automatic face detection, particularly important for surveillance and security systems. The face regions selected can be subsequently compared with a database of faces for identification (courtesy of Stephen McKenna, Queen Mary and Westfield College, London).



Applicazioni: Navigazione autonoma

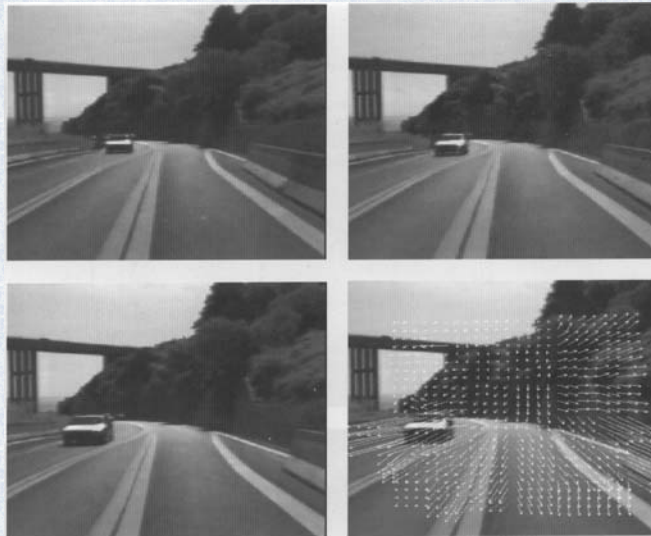


Figure 1.3 Computer vision and autonomous road navigation: some images from a sequence acquired from a moving car, and the estimated motion field (optical flow, discussed in Chapter 8 computed by a motion analysis program, indicating the relative motion of world and camera.



Applicazioni: Esplorazione autonoma

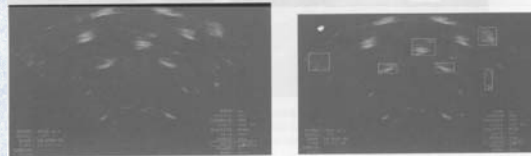


Figure 1.4 Computer vision is becoming increasingly important for remotely operated and autonomous subsea vehicles (ROV/AUVs), like the one shown above, ANGUS, built by the Ocean Systems Laboratory of Heriot-Watt University. As with many ROV/AUVs, ANGUS carries video and sonar sensors (see Chapter 2). Bottom left: an example of underwater sonar image. The white areas are the returns from a diver and the poles of a pier, imaged from above. Bottom right: the result of automatic search for objects of interest (courtesy of Dave Lane, Heriot-Watt University).



Applicazioni: Telepresenza

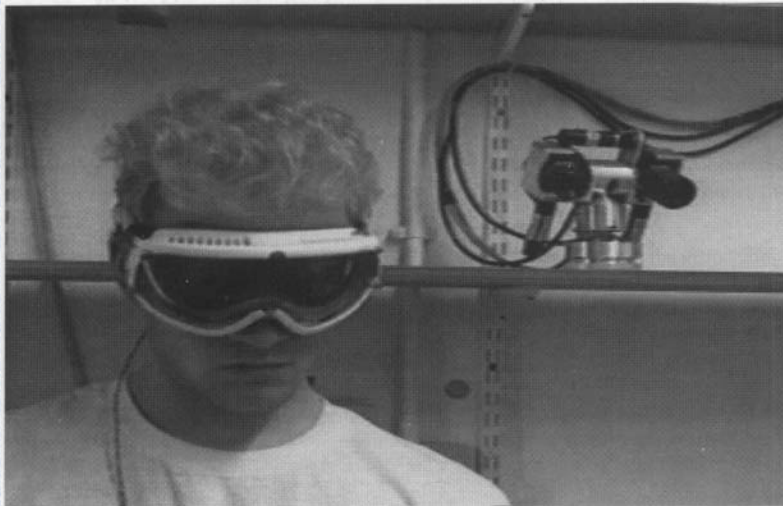


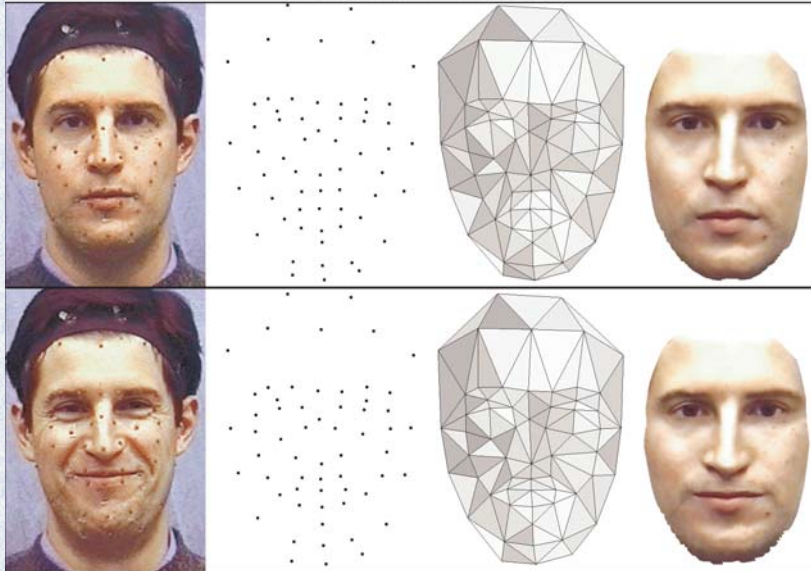
Figure 1.5 Computer vision and virtual telepresence: the movements of the operator's head are tracked by a vision system (not shown) and copied in real time by the head-eye platform (or stereo head) on the right (courtesy of David W. Murray, University of Oxford).



Analisi della comunicazione facciale



Riconoscimento di marker



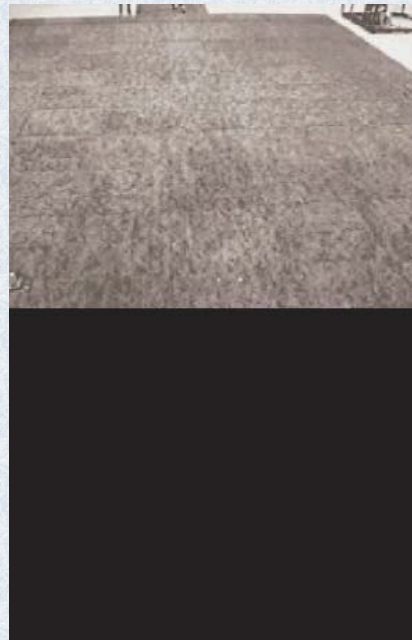
Copyright N.A. Borghese Università di Milano 19/03/2003

<http://homes.dsi.unimi.it/~borghese>

9/20



Analisi del movimento



Copyright N.A. Borghese Università di Milano 19/03/2003

10/20



Computer Vision



Obiettivo: determinazione delle proprietà geometriche, fisiche e dinamiche del mondo che ci circonda mediante elaborazione di immagini o sequenze di immagini.

- *High level vision:* riconoscimento, associazione di un significato semantico all'atto del vedere.
- *Low level vision (o early vision):* estrazione dalle immagini o sequenze di immagini delle informazioni necessarie al livello superiore (features).



I 3 stadi della visione



- Estrazione di "feature".
 - **Calcolo del Flusso ottico.**
 - Calcolo delle sorgenti di illuminazione e stima dell'albedo.
 - Riconoscimento di tessiture.
 - **Stereo-matching.**
 - Contorni (edge)
 - Forme dai contorni.
 - Forme da tessitura.
 - Forme da ombreggiatura.
 - **Determinazione della struttura 3D e del movimento 3D di oggetti da immagini monoculari.**
 - **Ricostruzione 3D da stereo di oggetti della scena.**
 - Costruzione di superfici.
 - Calcolo del colore di una superficie.
 - Parametri geometrici del sistema di visione x
- Interpretazione (la visione artificiale deriva storicamente dall'Intelligenza Artificiale).



Cosa sono le features?



1) *Località.*

2) *Significatività.*

3) *Riconoscibilità.*

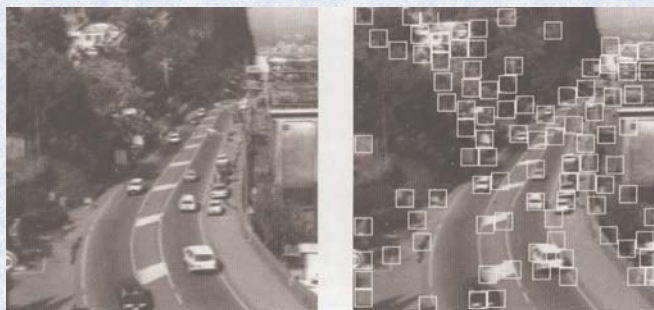
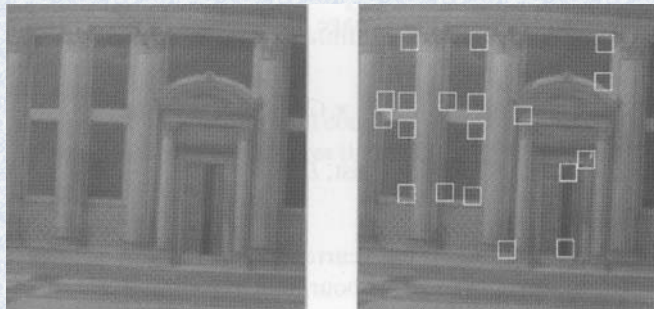
Feature attraverso la deteazione dei bordi



Copyright N.A. Borghese Università di Milano 19/03/2003



Riconoscimento di spigoli



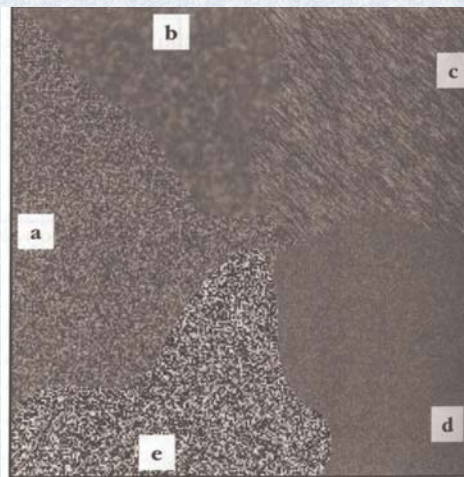
Copyright N.A. Borghese Università di Milano 19/03/2003

<http://homes.dsi.unimi.it/~borghese>

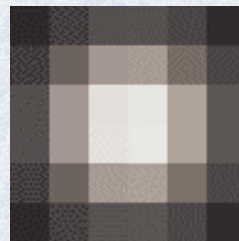
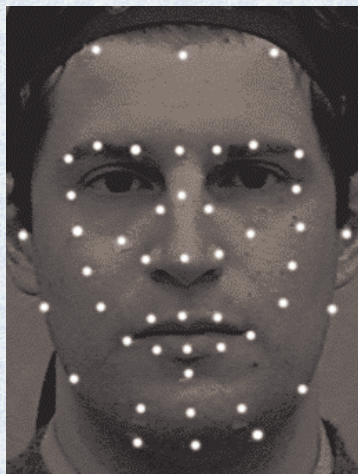
14/20



Estrazione di regioni



Marker detection



-8	-8	-8	-8	-8	-8
-8	0	1	1	0	-8
-8	1	7	7	1	-8
-8	1	7	7	1	-8
-8	0	1	1	0	-8
-8	-8	-8	-8	-8	-8

Cross-correlazione



Il flusso ottico

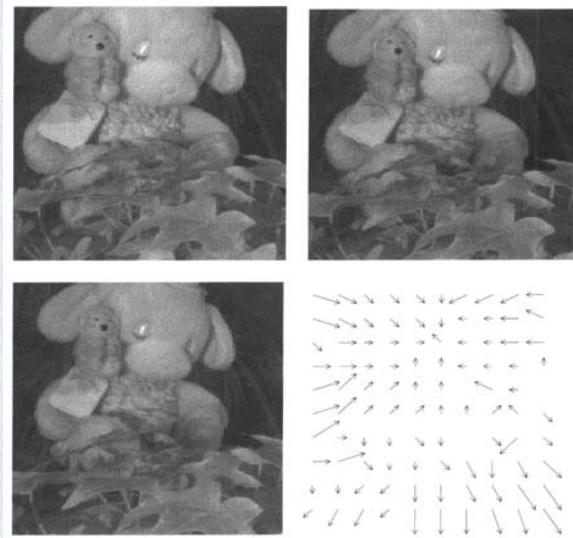


Figure 8.3 Three frames from a long image sequence (left to right and top to bottom) and the optical flow computed from the sequence, showing that the plant in the foreground is moving towards the camera, and the soft toys away from it.



I problemi di visione sono mal posti



- Perché non è facile costruire un sistema di visione?
- Risoluzione spaziale limitata.
- Gli oggetti reali non sono mai uniformemente illuminati.
- I contorni non sono netti.
- Le superfici non hanno albedo costante.
- L'illuminazione genera campi di irradianza "difficili".

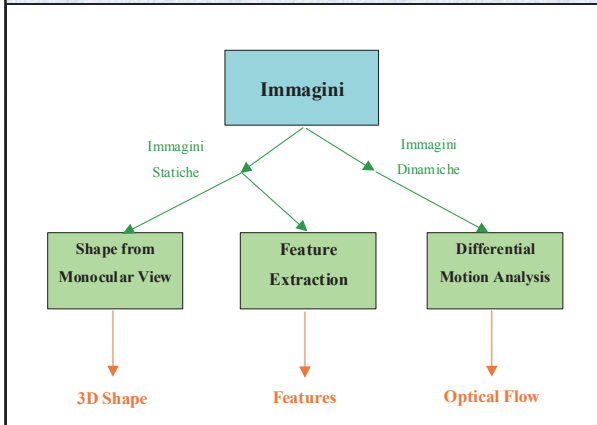


Visione 3D ed Elaborazione di immagini



- **Visione 3D:** Immagine/i \Rightarrow Ricostruzione 3D della scena statica o dinamica ed interpretazione.
- **Grafica 3D:** Modello 3D della scena, statico o dinamico \Rightarrow Visualizzazione.

Si incontrano sul terreno della visualizzazione 3D.



L'elaborazione delle immagini costituisce il primo livello di un sistema di visione. Fornisce le features di base.



Parleremo di:



- Flusso ottico.
- Riconoscimento di feature (corners) mediante algoritmi di analisi locale dell'immagine.
- Ricostruzione 3D di feature della scena da immagini stereoscopiche.
- Ricostruzione 3D di feature della scena ed il movimento 3D da sequenze monoculari (Structure from Motion).