

Testi del Syllabus

Resp. Did. **COLOMBO CARLO** **Matricola: 097762**

Docente **COLOMBO CARLO, 6 CFU**

Anno offerta: **2024/2025**

Insegnamento: **B024316 - COMPUTATIONAL VISION**

Corso di studio: **B241 - INTELLIGENZA ARTIFICIALE**

Anno regolamento: **2024**

CFU: **6**

Settore: **ING-INF/05**

Tipo Attività: **B - Caratterizzante**

Anno corso: **1**

Periodo: **Secondo Semestre**



Testi in italiano

Lingua insegnamento	ITALIANO
Obiettivi formativi	<p>Il corso intende dotare lo studente di strumenti teorici e pratici per l'analisi computazionale (a partire da modelli espliciti basati sulla geometria e sulla fisica della visione) di immagini singole, collezioni di immagini e sequenze video.</p> <p>Tali conoscenze sono finalizzate allo sviluppo di sistemi e algoritmi di computer vision</p> <p>2D e 3D, con applicazioni nei più svariati campi, quali la robotica autonoma e la guida automatica, la fruizione e preservazione dei beni culturali, gli ausili per disabili, l'automazione industriale, l'interazione avanzata uomo-macchina basata su movimenti e gesti, la grafica 2D/3D interattiva ed adattativa, l'informatica forense.</p>
Contenuti (Dipl.Sup.)	<ol style="list-style-type: none">1. LA VISIONE NELL'UOMO E NELLA MACCHINA: ASPETTI COMPUTAZIONALI2. FORMAZIONE DELL'IMMAGINE3. VISIONE MONOCULARE: VISTE DI SINGOLI PIANI4. REGISTRAZIONE DI IMMAGINI5. TELECAMERE: MODELLI E CALIBRAZIONE6. Ricostruzione 3D densa da sequenze video.7. ALGORITMI DI STEREOPSI, Stereo fotometrico.8. RICOSTRUZIONE 3D DA VISTE SINGOLE E MULTIPLE9. APPLICAZIONI: Beni culturali, Realta' aumentata, INTERAZIONE UOMO-MACCHINA, ROBOTICA avanzata, MULTIMEDIA, Informatica forense, etc.
Prerequisiti	Conoscenze di base sulla rappresentazione ed elaborazione delle immagini. Nozioni di algebra e di geometria.
Metodi didattici	Lezioni in aula con lavagna e videoproiettore. Sperimentazioni in aula con computer portatili e rete wireless.

Modalità di verifica dell'apprendimento

Esame orale. Il colloquio prende avvio dalla teoria di uno degli argomenti in programma, che si chiede di illustrare in dettaglio. Seguono domande di ordine pratico-realizzativo, ed esercizi volti a verificare la comprensione della teoria come strumento per risolvere problemi reali. Elaborato o ricerca bibliografica (solo studenti 9 CFU). L'elaborato verte sulla scrittura di un programma (in MATLAB, C++ o Python) per la soluzione di un semplice problema attraverso le tecniche studiate nel corso o altre tecniche che il candidato desideri approfondire. Al programma deve essere allegata una relazione che descriva i dettagli matematici e di implementazione degli algoritmi impiegati. La ricerca bibliografica (anch'essa proponibile dallo stesso studente) consiste in una review di una particolare tecnica di visione (es. algoritmi per il calcolo della disparità stereo) alla luce della letteratura recente, ovvero nell'approfondimento di un ambito applicativo della visione (es. algoritmi di visione per l'industria cinematografica).

Programma esteso

1. LA VISIONE NELL'UOMO E NELLA MACCHINA: ASPETTI COMPUTAZIONALI

Introduzione al corso. Ambiguità in visione. Illusioni visive. Il ruolo della semantica nella percezione. Indizi 3D in un'immagine.

2. FORMAZIONE DELL'IMMAGINE

L'immagine come sintesi di luce, materiale e geometria da parte di un osservatore. Tipi di superfici. BRDF. Albedo. Componenti diffusa e speculare. Dispositivi di acquisizione delle immagini. Ottiche.

3. VISIONE MONOCULARE: VISTE DI SINGOLI PIANI. Omografie e loro anatomia. Rettificazione di immagini basata sui punti circolari.

4. REGISTRAZIONE DI IMMAGINI

Stima robusta di omografie: RANSAC etc. Mosaici. Mosaicing in presenza di parallasse. Image-based rendering. Tecniche super-risoluzione.

5. TELECAMERE: MODELLI E CALIBRAZIONE

Camera a foro stenopeico, natural camera, affine camera. Distorsione radiale. Calibrazione fotogrammetrica. Autocalibrazione.

6. ANALISI DI SEQUENZE VIDEO

Optical flow vs motion field. Structure from motion (caso continuo). Tempo all'impatto. SLAM (simultaneous localization and mapping).

7. ALGORITMI DI STEREOPSI

Geometria di due viste. Matrice fondamentale ed essenziale. Parallasse. Rettificazione di una coppia stereo. Ricostruzione proiettiva e metrica. Disparità. Algoritmi per lo stereo denso. Stereo fotometrico.

8. RICOSTRUZIONE DA VISTE SINGOLE E MULTIPLE

Pipeline di ricostruzione da viste multiple. Bundle adjustment. Ricostruzione da viste singole: vincoli sulla scena (piani, superfici di rivoluzione, etc.). Metrologia da una vista singola.

9. APPLICAZIONI: Interfacce naturali uomo-macchina, Esterocepsi per robot, Post-produzione video, videoproiettori intelligenti, fotografia computazionale, televisione 3D, informatica forense, etc.

Testi di riferimento

1. Hartley and Zisserman, MULTIPLE VIEW GEOMETRY IN COMPUTER VISION.

Cambridge University Press, 2003 (2nd edition).

2. Torralba, Isola and Freeman, FOUNDATIONS OF COMPUTER VISION. MIT Press, 2024.

Su diversi argomenti saranno messi a disposizione degli studenti sia appunti di lezione che articoli in lingua inglese tratti dalla

letteratura recente.

Altre informazioni

Sito web del corso:
http://cvg.dsi.unifi.it/colombo_now/viscomp.html

Obiettivi per lo sviluppo sostenibile

Codice	Descrizione
--------	-------------



Testi in inglese

Language	Italian
	The course aims at providing students with theoretical and practical tools for the visual analysis of images and videos, based on explicit geometric and physical models. Such knowledge will be used to develop 2D and 3D computer vision algorithms and systems, with applications in several fields, such as autonomous robotics and AGVs, fruition and valorization of cultural heritage, aids for disabled people, industrial automation, human-computer interaction based on movements and gestures, interactive 2D/3D computer graphics, information forensics.
	<ol style="list-style-type: none">1. Vision in Man and Machine: Computational aspects.2. Image formation.3. Monocular vision: views of planar surfaces.4. Image registration.5. Cameras: models and calibration.6. 3D Structure from dense 2D motion: video sequences.7. Stereo vision. Photometric stereo.8. 3D reconstruction from single images and image collections.9. Applications: Cultural heritage, Augmented reality, Human-machine interaction, Advanced robotics, multimedia, Information forensics, etc.
	Basic knowledge on image representation and processing, algebra and geometry.
	Classroom teaching with blackboard and videoprojector. Classroom experiments with laptops and wi-fi.
	Oral test. The student will be asked to outline in some detail the theory of a selected topic, from which the discussion will move on to practical-implementation aspects, and exercises aimed at assessing the student's ability to use theory as a tool for solving concrete, real problems. Homework assignment (only for 9 CFU students), which will consist in the writing of either a computer program (in MATLAB, C++ o Python) or a bibliographical essay. The topic will be either suggested by the teacher, or proposed by the student. The computer program will be accompanied by a document detailing the theoretical and implementation aspects of the work. The bibliographical essay will either review a specific vision topic (e.g. computation of stereo disparity) as addressed by the recent literature, or discuss in detail a selected application scenario of computer vision (e.g., vision algorithms for the movie industry).
	<ol style="list-style-type: none">1. VISION IN HUMAN AND MACHINES: COMPUTATIONAL ASPECTS. General introduction. Ambiguity of vision. Visual illusions. The role of semantics in perception. Visual attention. Basic 3D cues.2. IMAGE FORMATION. The image as the combination of light, material,

geometry and the observer. Tipi di superfici. BRDF. Albedo. Diffuse and specular components. Image acquisition devices. Lenses.

3. MONOCULAR VISION: PLANAR SURFACES. Homographies and their anatomy.
Image rectification based on circular points.

4. IMAGE REGISTRATION

Robust homography estimation: RANSAC etc. Mosaicing.
Mosaicing in the presence of parallax. Image-based rendering. Super-resolution.

5. CAMERA MODELS AND CALIBRATION

Camera pinhole, natural and affine camera. Radial distortion.
Photogrammetric calibration. Self-calibration.

6. VIDEO ANALYSIS

Optical flow vs motion field. Structure from motion (continuous case).
Time to collision. SLAM (simultaneous localization and mapping).

7. STEREOPSIS

Two-view geometry. Fundamental and Essential matrices. Parallax.
Stereo rectification. Projective and metric reconstruction. Disparity.
Dense stereo algorithms. Photometric stereo.

8. 3D RECONSTRUCTION FROM SINGLE VIEWS AND IMAGE COLLECTIONS

3D reconstruction pipeline. Bundle adjustment. Single-view reconstruction: scene constraints (planes, surfaces of revolution, etc.).
Single view metrology.

9. APPLICATIONS

Natural human-machine interfaces, advanced robot exteroception, video post-production, smart videoprojectors, computational photography, 3D television, information forensics, etc.

1. Hartley and Zisserman, MULTIPLE VIEW GEOMETRY IN COMPUTER VISION.

Cambridge University Press, 2003 (2nd edition).

2. Torralba, Isola and Freeman, FOUNDATIONS OF COMPUTER VISION. MIT Press, 2024.

Teaching notes and recent papers on selected topics will be made available to students.

Course webpage:

http://cvg.dsi.unifi.it/colombo_now/viscomp.html

Obiettivi per lo sviluppo sostenibile

Codice

Descrizione