

Elaborazione delle Immagini

Sara Angeretti

@Sara1798

Fabio Ferrario

@fefabo

2023/2024

Indice

1	Il corso	3
1.1	Programma del corso	3
1.2	Gli orari	3
1.3	Programma Esteso	4
1.4	Libri di testo	4
1.5	Contatti	5
1.6	L'esame	5
2	Cenni sulla visione	8
2.1	Formazione del segnale immagine	8
2.1.1	Imaging	10
2.1.2	Radiometria e grandezze radiometriche	10
3	Digitalizzazione delle immagini	11
3.0.1	Nei video	11
3.1	Segnale immagine o irradianza	13
3.2	Campionamento vs quantizzazione	13
3.2.1	Campionamento	14
3.2.2	La risoluzione	14
3.2.3	Vari modi per rappresentare un'immagine	15
3.2.4	Quantizzazione	15
3.3	Digitalizzazione immagini a colori	16
3.3.1	Lo spazio colore RGB	16
3.3.2	HUI Hetc ho perso	17
3.3.3	Memorizzare i colori	17
3.4	Compressione delle immagini	17

Capitolo 1

Il corso

Il corso di Elaborazione delle Immagini 2023-2024 è erogato dai professori **Schettini Raimondo** e **Ciocca Gianluigi**.

1.1 Programma del corso

Durante il corso lo studente acquisirà competenze specifiche che lo porranno in grado di comprendere la catena di elaborazione, analisi e classificazione di immagini e video. Lo studente acquisirà inoltre le competenze necessarie per progettare, sviluppare ed integrare specifici moduli in sistemi applicativi complessi.

Competenze per cui questo corso prepara: esempi

Image Recognition in the Consumer Goods and Services Industry (supported by Accenture). Per esempio, una certa marca di prodotto paga un costo aggiuntivo al proprio valore per essere venduto ad altezza occhi per 10-12 metri di scaffale, mentre il prodotto sull'ultimo scaffale ha meno costi aggiuntivi. Parole sue: ***meglio essere quello che automatizza di quello che viene automatizzato.***

1.2 Gli orari

Martedì U24 C1 13:30-16:30 Lezione

Giovedì U14 LabA1 11:30-13:30 Lezione

Venerdì U14 LabA1 08:30-11:30 Lezione fino a metà Ottobre, poi Laboratorio

Le aule potrebbero cambiare di settimana in settimana.
Sono previsti seminari, in orario di lezione e non, e sono parte integrante del corso.

1.3 Programma Esteso

1. Cenni sulla percezione visiva, la visione umana e artificiale, il colore. Acquisizione e digitalizzazione di immagini.
2. Miglioramento delle immagini con operatori puntuali.
3. Filtraggio spaziale lineare e non-lineare.
4. Spazi colore. Elaborazione delle immagini a colori.
5. Segmentazione di immagini per regioni e per contorni.
6. Analisi tessiturale; Morfologia Matematica.
7. Descrizione e rappresentazione di immagini (regioni, contorni, approssimazione poligonale).
8. Riconoscimento, classificazione supervisionata e non supervisionata.
9. Introduzione alle reti neurali convoluzionali profonde.

1.4 Libri di testo

- R. Gonzalez, R. Woods, Digital Image Processing, Second Edition, 2002. Prentice Hall.
- R. Gonzalez, R. Woods, Digital Image Processing, Third edition, 2008 Prentice Hall. <http://www.imageprocessingplace.com/>
- R. Gonzalez, R. Woods, elaborazione digitale delle immagini, terza edizione (utile per chi ha difficoltà con l'inglese) non ha tutti i capitoli.
- R. Gonzalez, R. Woods, Digital Image Processing, forth edition, 2018 Prentice Hall.

Credo sia quest'ultimo perché parlava dell'ultimo del 2018. Comunque fa che costicchia, tipo tra i 100 e i 150€, ma se lo compri sull'Amazon indiano costa tipo 25\$ wth. Comunque giovedì 28 lo porta in aula. Che mentiroso,

non l'ha fatto.

Ocio non quello in italiano perché mancano i capitoli finali. <http://www.imageprocessingplace.com/>

Dispense a cura del docente (formato elettronico e/o cartaceo)

Alcuni libri on-line su argomenti affini al corso possono essere trovati al seguente indirizzo: <http://homepages.inf.ed.ac.uk/rbf/CVonline/books.htm>

1.5 Contatti

Comunicazioni docente-studenti

- Messaggi di posta elettronica raimondo.schettini@unimib.it
- Nel subject usare la parola chiave [EI]
- Firmare i messaggi con nome e cognome e mandarli solo quando essenziale
- Per esigenze di tipo organizzativo (prenotazioni, esami, ecc.) rivolgersi alla segreteria del dipartimento, non al docente.

Ricevimento studenti: Raimondo Schettini

1. U14, 2° piano, stanza 2059
2. Per appuntamento via mail (specificare il motivo della richiesta).
3. Dopo le lezioni.

Possibilità di stage/tesi:

1. Consultare il sito del laboratorio per avere una idea delle tematiche possibili per stage interni www.ivl.disco.unimib.it.
2. Inviare CV e piano di studi (con i voti degli esami sostenuti) a raimondo.schettini@unimib.it
3. Verrete convocati per un colloquio in cui vi verranno proposti stage interni od esterni.

1.6 L'esame

L'esame è composto di due prove:

Prova Scritta

- Sostenuta il giorno dell'appello d'esame come da SIFA
- Valutata in 30-esimi
- Consiste in un insieme di domande con risposte a scelta multipla e libera
- Due step
 1. 5 domande "base" di sbarramento (5 punti) max 15 minuti
Chi NON risponde correttamente a 4 su 5 domande di sbarramento non ha superato l'esame
 2. N domande con punteggio variabile (25 punti in totale) ha fatto vedere un esempio, un esercizio sulle chiavi
- Se il voto è < 18 l'esame non è superato.
- Se il voto è ≥ 18 si passa al progetto.

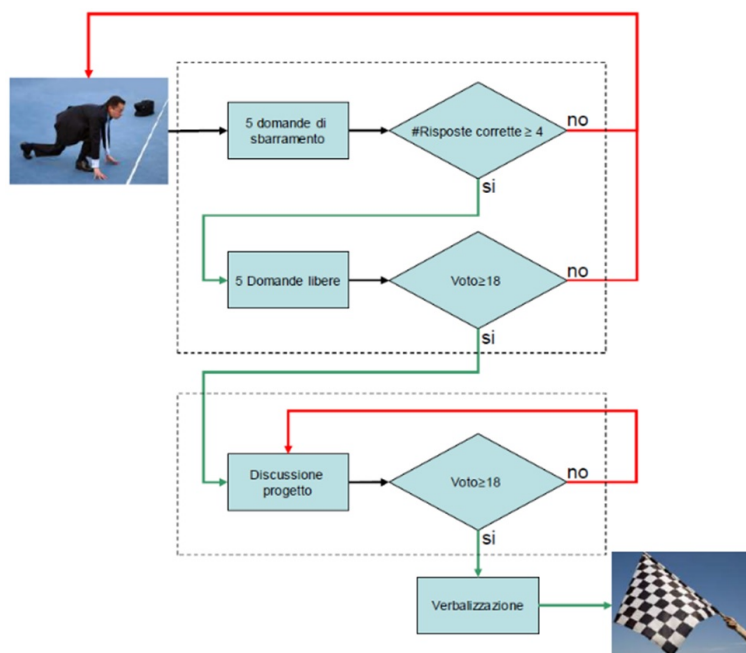
Prova Pratica

- La prova pratica consiste nella valutazione e discussione dello svolgimento di un progetto
 - Possibile svolgere il progetto in gruppi di **minimo 2 persone** massimo 3 persone
- Verranno valutate
 - Funzionalità del codice del progetto (demo live)
 - Slides che illustrino la struttura del progetto, le tecniche e algoritmi utilizzati, risultati ottenuti
- Se il voto è < 18 l'esame non è superato.
- Se il voto è ≥ 18 si passa alla verbalizzazione.
- Il progetto può essere svolto in qualunque linguaggio di programmazione
- Le slides devono descrivere schematicamente il funzionamento dell'applicazione sviluppata:

- Mediante diagrammi di flusso
- Descrizione di alto livello (macro-moduli)
- Descrizione dei singoli moduli

Informazioni finali

- La prova pratica può essere sostenuta in un appello diverso dalla prova scritta (purché superata)
 - Es. Prova scritta sostenuta all'appello di Febbraio, superata e prova pratica sostenuta all'appello di Luglio.
- Il voto della prova scritta rimane valido fino all'ultimo appello dell'anno corrente
 - Indicativamente Febbraio 2021 - Settembre/Novembre 2022
 - Dal primo appello del nuovo anno (2022), il voto viene, di norma, annullato
- Il progetto è unico per tutti gli appelli dell'anno corrente
- La prova pratica deve essere presentata da tutti i membri del gruppo
 - Almeno un membro del gruppo deve aver superato lo scritto
- Il voto della prova pratica è individuale



Capitolo 2

Cenni sulla visione

2.1 Formazione del segnale immagine



Le immagini vengono generate da radiazioni percepite da un sensore, siccome esistono diversi tipi di radiazioni (e di sensori), esistono anche diverse tipologie di immagini. La banda Elettro Magnetica che il nostro sistema visivo è in grado di percepire è molto ridotta sullo spettro elettromagnetico. Questa banda che noi possiamo vedere é chiamato **Spettro di Luce Visibile**.

Ci sono sempre nuovi sensori che permettono di rivelare altri tipi di radiazioni.

Es.: immagini della Luna, o lastre a raggi X fatte in ospedale, o angiogrammi, o altre immagini mediche.

Mentre diversi tipi di illuminazione sono: raggi gamma, raggi X, UV, luce visibile, infrarosso, onde radio. . .

Una stessa scena acquisita con diversi rivelatori genera immagini anche molto diverse fra loro. Potremmo avere:

- immagini di *riflessione* (e.g. immagini radar, sonar, slide stampata...)
- immagini di *emissione* (e.g. termografie, questa slide a monitor...)
- immagini di *trasmissione* (e.g. raggi X, questa slide sul proiettore...)

Riprendo da qua con interazione fra radiazione e materia, slide 9.

La lunghezza d'onda è proporzionale all'energia: lunghezze più basse (e pericolose) sono con energia più alta e viceversa. Ma non ha senso, ho capito male?

Ha fatto vedere immagini di luna e mediche. Sono tutte grigie. Come mai? Ad ogni immagine è associato uno scalare che danno informazioni per la mappazione.

Come mai un oggetto (es la felpa del compagno) è di un certo colore es blu? La luce è bianca, ma i range vengono quasi tutti assorbiti tranne il range che effettivamente vediamo.

L'immagine è un array dove per ogni elemento c'è un'informazione. Es a livello di grigio o sul colore.

Vector vs Raster Che differenza c'è fra un'immagine vector e un'immagine Raster?

La vector salva l'immagine in forme geometriche, es di un cerchio salva centro e raggio e ogni volta che la usi o sposti o copi viene ridisegnata, si salvano i vettori dei poligoni e il colore e vengono ridisegnate invece di memorizzare tutta l'immagine. Il vantaggio è la riutilizzabilità, nel senso che posso zoommare quanto voglio e l'immagine rimarrà nitida. La Raster invece è formata da pixel quindi ha una sua risoluzione e se vado ad ingrandirla diventa sgranata e illeggibile.

I pixel Il concetto di pixel è molto relativo: le immagini tipo da satellite avranno dei pixel giganteschi, mentre un'immagine piccina del pc o della mia fotocamera li avrà molto piccoli.

Problema delle scale: una delle prime difficoltà che troviamo quindi è che in una stessa immagine trovo oggetti che sono su scale diverse. Ma non tratteremo questo aspetto: ci sono algoritmi in grado di stimare la distanza fra oggetti raffigurati ma se ne parla alla magistrale (mi sono persa il nome tecnico di questo studio). Il problema è che è la grandissima maggioranza delle applicazioni che troviamo.

2.1.1 Imaging

Fabio, qua mi manca un pezzettino.

2.1.2 Radiometria e grandezze radiometriche

Studio del trasferimento di energia, energia emessa da un certo oggetto, etc.
Immagine.

Capitolo 3

Digitalizzazione delle immagini

Slide delle fasi dei Digital Image Fundamentals, siamo nella prima, digital world.

Immagini binarie, $x \in [0, 1]$, appartiene ad una classe o all'altra.

Abstract? Immagini scala di grigio, la trasformo in array. chiedo a fabio

$l(x, y) \in (a, b)$

la digitalizzazione di immagini a colori è più complessa, la rappresento con 3 valori, una per ogni canale.

I canali non sono colori, sono visualizzati a livello di grigio (0-255) e in fase di visualizzazione esce il colore.

Altri tipi di immagini, satellitari ad es., che sono a 7 bande. Perché 7? La visione umana è a 3 bande, quindi 3 canali.

Oltre al visibile, potremmo essere interessati ad altre informazioni es. certi range per le onde elettromagnetiche (es. se l'uva è matura, se è inquinata, etc., cose non visibili a occhio).

3.0.1 Nei video

Abbiamo una dimensione in più. partiamo dai diversi frame che prendono ad es. le basse luci, le luci medie e le luci alte che vengono poi combinate per creare le immagini Hcos?

Perché utile? Perché nelle camere più moderne brevi sequenze video vengono combinate per creare immagini migliori.

Slide differenze immagini analogiche e digitali. La prima ha parametri spaziali che replicano le dimensioni reali, le seconde sono rappresentate da livelli di grigio.

Acquisizione di immagini

Due dispositivi:

- dispositivi analogici
- dispositivi digitali

La qualità del colore per le immagini analogiche è migliore per quanto riguarda il colore.

quando dico che un'immagine è digitale, è digitalizzata nel dispositivo con cui la acquisisco.

Caratterizzazione dell'immagine

il processo di formazione dell'immagine implica il mapping di una scena 3D in 2D, viene proiettata. Con questa proiezione rischio di perdere il senso di distanza (o lo perdo proprio sempre?) (slide digitalizzazione di immagini).

a meno che non si metta nell'immagine un riferimento assoluto che conosco per certo, non riesco a capire le reali dimensioni dell'oggetto. È quello che diceva nella prima lezione quando parlava delle iridi come campione di misura.

Digitalizzazione di immagini

Le camere fotografiche digitali moderne sfruttano lo stesso concetto alla base della camera oscura.

È anche l'idea alla base della Polaroid. Perché quando togliamo la pellicola l'immagine diventa tutta grigia? Perché come ha detto Fabio non c'è la lente, non c'è la camera, perciò esce tutta grigia.

Ha nominato un algoritmo "Gray cosa?" che useremo per bilanciare le immagini.

La camera pinhole crea un'immagine riflessa dell'oggetto. Due slide su camera pinhole.

Potremmo fare un buco più grande per fare entrare più luce, ma non avendo la lente l'immagine esce molto sfocata.

Si usano quindi le **lenti**, che hanno una teoria molto complicata ma focalizza l'immagine.

Seguono slide utili per completezza. Una è su parametri geometrici, ottici, fotometrici. . . più ovviamente i parametri dei sensori.

I sensori

Slide con la matrice grigia.

Ho un filtro, un vettore e un array.

Quando uso il filtro? Per togliere infrarossi e ultravioletti, riduce tutto ciò che non è luce visibile.

Quando uso il vettore (la linea)? Oggi si usa su ali dell'aereo o droni per fare foto aeree.

Quando uso la matrice? Boh.

Slide figura 2.13

La risoluzione più alta ce l'ho a sinistra (che si muove) o quella a destra su cui ho messo i sensori?

Decisamente la prima perché la controllo io, la seconda è la dimensione del sensore. Lo scanner migliore (ha fatto l'esempio dello scanner a casa) è quello che ha la risoluzione più alta. Quindi sinistra.

3.1 Segnale immagine o irradianza

Questo è troppo semplificato sul libro. Slide con l'immagine della donna.

È il prodotto dell'illuminazione e della riflettanza.

Guarda la slide manca il secondo pezzo.

La riflettanza va da 0 a 1, l'illuminazione dipende dalla sorgente. Ma è *sempre* fra 0 e 1? Ha fatto l'esempio della discoteca e dei raggi UV. Alcune sostanze, alcuni sbiancanti, assorbono, rielaborano e rimettono la luce dei raggi UV.

3.2 Campionamento vs quantizzazione

Slide presa dal Gonzales, c'è una matrice. Il prof mette l'origine nell'elemento in basso a sinistra, il libro in alto a sinistra. La slide successiva mostra i grafici differenze fra campionamento e quantizzazione, dove è possibile vedere il *rumore*.

La slide successiva ne mostra la digitalizzazione. un pixel rappresenta quindi una regione dell'immagine e ha associato un solo valore di grigio. Ma è digitalizzato bene? Dipende dagli scopi della mia indagine.

3.2.1 Campionamento

Il sensore realizza un campionamento spaziale dell'immagine acquisita. La cella del sensore del telefono deve diventare più piccola perché devo tenere insieme i pixel, l'area sensibile diventa più piccola. Cosa vuol dire? Che perdo luce, prende più dettagli ma perde segnale. Quindi il mio sensore deve essere molto ma molto più sensibile.

Poco segnale = poco rumore.

Non sempre la risoluzione spaziale è un parametro che va massimizzato.

3.2.2 La risoluzione

È un concetto relativo.

Tre tipi:

- r. dell'apparecchiatura di acquisizione
- r. dell'apparecchiatura di resa
- r. di ripresa di una scena

Aliasing e campionamento

Quando digitalizzo un'immagine, il campionamento in alto non è necessario. Due slide, una è un buon esempio di campionamento, la seconda no. Si va a creare un'immagine che prima non avevo. È il fenomeno di Aliasing. Slide successiva con definizione. Poi c'è una slide con l'esempio di immagine corrotta per il fenomeno di Aliasing, esempio della persiana.

Ci sono teoremi che legano cosa al mio campionamento.

Il nostro campionamento deve essere almeno la metà della più piccola cosa che vogliamo sia visibile. Parliamo del teorema di Shannon. Qualsiasi segnale deve essere la somma di diverse curve, per essere sicuri di prendere tutti i più piccoli campionamenti.

Slide orologio da tasca: immagine divisa in 1250 punti.

Campionamento e tecnologia dei sensori

Slide risoluzione ottimale e tecnologia dei sensori.

Non va bene un sensore? Se ne cerca un altro.

3.2.3 Vari modi per rappresentare un'immagine

Slide con matrici.

Ci sono dimensioni più o meno standard.

Genericamente il più importante è il livello di quantizzazione. È il livello di grigio. È quasi sempre una potenza di due.

Es. slide della rosa: i pixel hanno sempre la stessa dimensione, ma cambia il numero di pixel. una 1024x1024 sarà grande e nitida, una 32x32 avrà grosse porzioni di immagine coperte da un solo pixel. Ovviamente se vado a prendere una piccola con pochi pixel es 32x32 e vado a ingrandirla, perdo dettagli e l'immagine uscirà molto molto sgranata. È la tecnica dietro all'oscuramento di volti alla tv. Ci sono algoritmi per simulare il recupero di definizione (visti alla magistrale), però sono una stima, una volta che l'informazione è persa è persa. È la slide di es. di elaborazione di basso livello, il lavoro che hanno fatto per la Canon, CNN con immagini di scimmia, farfalla, boh e bambino.

3.2.4 Quantizzazione

Quanti livelli di grigio servono?

La risoluzione di livelli di grigio è la più piccola differenza del livello di grigio è la più piccola che posso distinguere.

Slide con il tizio, i numeri sono il numero di livelli di grigio. Con la quantizzazione prendiamo un numero finito di valori nel range $f(x, y)$.

Esempio della stampante: fa quantizzazione, partendo da 4 colori ne seleziona i livelli per andare ad ottenere milioni di colori. La quantizzazione entra in campo quando diminuisco la dimensione dei dot nella visualizzazione puntinata dell'immagine.

Quantizzazione non uniforme

Slide. L'altra volta con la legge di Weber avevamo accennato alla quantizzazione, la discretizzazione dei livelli di grigio che non sempre opera su modelli lineari.

Tempo di acquisizione

Quando lo cambio cambio la quantità di luce che arriva al sensore. quando tengo un tempo troppo corto su immagine statica, l'immagine esce scura perché non entra abbastanza luce. Al contrario, quando è troppo alta va in saturazione e va a contaminare i pixel vicini (un po' quello che succedeva con

la Polaroid).

In caso dinamico ovviamente il tempo di esposizione deve essere diverso. Per esempio per un treno in corsa il tempo di acquisizione deve essere molto basso, ha bisogno di un sensore più sensibile così che bastino meno fotoni per generare il segnale.

Range dinamico

Molto diverso dal numero di pixel (numero di livelli, che fa che è tipo un salame, ovvero non importa quante fette faccio ma quanto è lungo il salame), o dal numero di bit.

C'è la slide con la definizione che mi sono persa.

In base a quello che voglio prendere (esempio finestra in una stanza buia) cambia il range dinamico che devo prendere.

Come si risolve? con l'HDR dice Fabio? Mi serve che il sensore veda più luce nel buio e meno luce nella zona illuminata.

Cosa succede ? Negli scuri si vede più rumore, nelle zone troppo illuminate si satura.

3.3 Digitalizzazione immagini a colori

Per ogni pixel devo avere le 3 informazioni, per RGB. Per acquisire tutti e tre i canali potrei usare prismi. ma è complicato per molti motivi, tra cui un costo eccessivo per 3 sensori, oltre ad un'ottica troppo complicata. Come fa ad esempio un cellulare ad acquisire i colori per ogni singolo pixel? Ogni pixel può vedere un solo canale, poi vengono combinati. Riprendiamo il concetto di curva luminosa di efficienza. Il sensore acquisisce quello che si chiama un Bayer pattern, che viene interpolato.

3.3.1 Lo spazio colore RGB

Lo possiamo vedere come un cubo con i 3 colori sugli assi cartesiani, Dove saranno i grigi? Sulla diagonale che va dal punto nero (0, 0, 0) al punto bianco (255, 255, 255). Allontanandomi dalla diagonale dei grigi, vado verso le superfici del cubo che hanno i colori più saturi più carichi.

Come faccio ad editare i colori? Ruoto il cubo? Chiedi a Fabio.

Slide con l'immagine colorata da scomporre in bande, seguono le versioni grigie e nere per ogni canale. Quella che rappresenta il canale verde è grigia e non bianca dove c'era il verde semplicemente perché non è il verde massimo.

Le immagini digitali possono esser visualizzate o riprodotte su appositi dispositivi quali monitor, display, proiettori e stampanti.

Differenza fra visione additiva e sottrattiva. La visione umana è additiva, ma i colori primari che usiamo sono Red, Green e Blue, non rosso giallo e blu come ci hanno insegnato alle elementari. Il sistema RGB però è sottrattivo.

— **Inizio Piccola Digressione** —

Nella slide delle fasi dei Digital Image Fundamentals, l'editing rientra nella fase finale, mental world.

Parliamo di related e unrelated colours. Un esempio del secondo è che in una situazione di tutto buio il marrone si vede giallo, in luce si vede marrone, praticamente giallo scuro. Praticamente è il colore che vediamo in relazione ad un altro colore.

— **Fine Piccola Digressione** —

3.3.2 HUI Hetc ho perso

Ho perso la slide, quella con la definizione di HUI.

3.3.3 Memorizzare i colori

Slide con la matrice.

3.4 Compressione delle immagini

Standard più usato JPEG. Prima di fare una compressione, devo fare una trasformazione del colore.

Le singole bande vengono compresse per ridurre la ridondanza del segnale. Poi l'immagine viene decompressa e restituita. Il JPEG ha un fattore di image quality. Se andiamo ad aumentare il fattore di compressione, diventa più piccola l'immagine in fase di compressione ma in fase di decompressione sui bordi vanno a crearsi degli artefatti.

Ovviamente il JPEG introduce gli artefatti ma è possibile con metodi e tecniche precise rimuovere questi artefatti. Parliamo di restauro immagini compresse.