



UNIVERSITÀ DEGLI STUDI DI CATANIA  
CORSO DI LAUREA MAGISTRALE IN INFORMATICA

---

RICCARDO RACITI  
ROSARIO CANNAVO'  
MARIO BENISSIMO

PROGETTO MULTIMEDIA

---

RELAZIONE

---

---

Anno Accademico 2022–2023



## Indice

<b>1</b>	<b>Clustering</b>	<b>1</b>
1.1	KMeans . . . . .	1
1.2	PCA . . . . .	2
1.3	HOG . . . . .	3
<b>2</b>	<b>Corner Detection</b>	<b>4</b>
2.1	FAST . . . . .	4
<b>3</b>	<b>Segmentazione</b>	<b>5</b>
3.1	SLIC . . . . .	6
3.2	Felzenszwalb . . . . .	6
3.3	Quickshift . . . . .	6
3.4	Watershed . . . . .	7

# Introduzione

L'obiettivo di questo progetto è quello di creare uno strumento di analisi e visualizzazione di semplice utilizzo basato su una pipeline realtime di image processing robusta e resiliente in grado di applicare degli algoritmi di Machine learning e Computer vision a delle rilevazioni fotografiche effettuate su Marte quotidianamente dai Rover Curiosity, Perseverance e Spirit per la ricerca di informazioni di interesse. L'endpoint implementato è un'interfaccia web messa a disposizione degli utenti che potranno visualizzare le analisi effettuate in modo rapido, avendo la possibilità di confrontare più risultati contemporaneamente.

## 1 Clustering

Una parte del progetto è stata incentrata sulla produzione di cluster in modo da poter classificare le immagini e poterle distinguere e categorizzare per similarità. I primi algoritmi implementati sono stati *KMeans*, *PCA* e clustering tramite *HOG*.

### 1.1 KMeans

Con il termine *K – means* si indica una classe di algoritmi di clustering partizionali basati su assegnamenti di punti. Essi permettono di suddividere un insieme di oggetti in  $k$  gruppi sulla base delle loro feature. Lavorano su spazi euclidei e assumono la conoscenza a-priori del numero di cluster  $k$ , che costituisce un iperparametro per l'algoritmo. Sono tuttavia presenti alcune tecniche per dedurre il miglior valore di  $k$  attraverso una serie di esperimenti.

## 1.2 PCA

*PCA* è l'acronimo di *AnalisiComponentiPrincipali*, è una tecnica finalizzata a ridurre la dimensionalità di un insieme di dati con finalità esplorative, di visualizzazione dei dati o feature extraction, per un eventuale uso in analisi successive. L'analisi delle componenti principali (*PCA*) è una tecnica finalizzata a derivare, a partire da un set di variabili numeriche correlate, un insieme più ridotto di variabili ortogonali “artificiali”. L'insieme ridotto di proiezioni ortogonali lineari (noto come “componenti principali” o “principal components”, “PC”) è ottenuto combinando linearmente in maniera appropriata le variabili originarie.

### 1.3 HOG

*Histogram of Oriented Gradients (HOG)*, è un descrittore di feature che viene spesso utilizzato per estrarre feature dai dati dell'immagine. In generale, è una rappresentazione semplificata dell'immagine che contiene solo le informazioni più importanti sull'immagine. Il descrittore di feature HOG conta le occorrenze dell'orientamento del gradiente nelle porzioni localizzate di un'immagine. È ampiamente utilizzato nelle attività di visione artificiale per il rilevamento di oggetti. Il descrittore HOG si concentra sulla struttura o sulla forma di un oggetto, è importante notare che è diverso dall'estrazione delle caratteristiche dei corner che possiamo estrarre per le immagini perché nel caso delle caratteristiche HOG, vengono estratti sia i corner che la direzione. L'immagine completa è suddivisa in regioni più piccole (porzioni localizzate) e per ogni regione vengono calcolati i gradienti e l'orientamento. Infine l'HOG genera un istogramma per ciascuna di queste regioni separatamente. Gli istogrammi vengono creati utilizzando i gradienti e gli orientamenti dei valori dei pixel, da cui il nome Histogram of Oriented Gradients.

## 2 Corner Detection

### 2.1 FAST

Successivamente si è scelto di implementare un algoritmo di corner detection.

Un corner può essere definito come l'intersezione di due angoli o come un punto per il quale ci sono due direzioni di angoli dominanti e diverse in un area locale del punto.

Un punto di interesse è un punto in un'immagine che ha una posizione ben definita e può essere rilevato in modo robusto. Ciò significa che un punto di interesse può essere un angolo ma può anche essere, ad esempio, un punto isolato di intensità locale massima o minima, terminazioni di linea o un punto su una curva in cui la curvatura è localmente massima.

In pratica, la maggior parte dei cosiddetti metodi di rilevamento degli angoli rileva i punti di interesse in generale e in effetti, il termine "angolo" e "punto di interesse" sono usati in modo più o meno intercambiabile in letteratura. Di conseguenza, se si devono rilevare solo gli angoli, è necessario fare un'analisi locale dei punti di interesse rilevati per determinare quali di questi sono angoli reali. L'algoritmo implementato è FAST (Features from accelerated segment test). Il vantaggio di FAST è la sua efficienza computazionale. Riferendosi al suo nome, è davvero più veloce di molti altri noti metodi di estrazione delle caratteristiche, come ad esempio DoG.

Il rilevatore di corner FAST utilizza un cerchio di 16 pixel (un cerchio Bresenham di raggio 3) per classificare se un punto candidato  $p$  è effettivamente un angolo. Ogni pixel nel cerchio è etichettato con un numero intero da 1 a 16 in senso orario. Se un insieme di  $N$  pixel contigui del cerchio è composto da punti di un'intensità maggiore di una certa soglia, il pixel candidato  $p$  verrà classificato come angolo.

### 3 Segmentazione

Un altro importante strumento dell'Image Computing è la *segmentazione*.

La segmentazione è un processo di partizionamento di un'immagine in regioni disgiunte e omogenee.

Sia  $R$  l'intera regione spaziale occupata dall'immagine. Il processo di segmentazione può essere visto come il partizionamento di  $R$  in  $n$  sottoregioni,  $R_1, R_2, \dots, R_n$  tali che:

1.  $\bigcup_{i=1}^n R_i = R$ .
2.  $R_i$  è un insieme connesso,  $i = 1, 2, \dots, n$ .
3.  $R_i \cap R_j \neq \emptyset$  per tutti i valori di  $i$  e  $j$ , con  $i$  e  $j$  diversi tra loro
4.  $Q(R_i) = \text{True}$  per  $i = 1, 2, \dots, n$ .
5.  $Q(R_i \cap R_j) = \text{False}$  per ogni coppia di regioni adiacenti  $R_i$  intersezione  $R_j$

Ogni pixel deve appartenere ad una regione. I punti appartenenti ad una regione devono essere connessi.

Le regioni devono essere disgiunte e i pixel appartenenti ad una regione devono soddisfare un certo predicato  $Q$ .

Due regioni adiacenti devono essere diverse nel senso del predicato  $Q$ .

Ad esempio il predicato  $Q$  potrebbe essere il seguente:

$Q(R_i) = \text{TRUE}$  se l'intensità media dei pixel di  $R_i$  è inferiore a  $m$  e la loro deviazione standard è minore di  $\sigma$  (con  $m$  e  $\sigma$  parametri costanti).

Per poter suddividere le immagini in gruppi di pixel adiacenti simili sono stati utilizzati degli algoritmi specifici.



### 3.1 SLIC

Questo filtro crea dei superpixel basati sull'algoritmo di clustering k-means. I superpixel sono piccoli cluster di pixels che condividono proprietà simili. I superpixel semplificano le immagini con un gran numero di pixels rendendole più facili da trattare in molti ambiti (computer vision, pattern recognition e intelligenza artificiale). Il colore del superpixel è la media dei colori dei pixel nella regione corrispondente.

### 3.2 Felzenszwalb

Effettua la segmentazione di un'immagine basandosi sulla teoria dei grafi. Sia  $G = (V, E)$  un grafo non diretto con vertici  $v_i \in V$ , l'insieme di elementi da segmentare e angoli  $(v_i, v_j) \in E$  corrispondente a coppie di vertici vicini. Ogni angolo  $(v_i, v_j) \in E$  ha un peso corrispondente  $w((v_i, v_j))$ , che è una misura non negativa della dissomiglianza tra gli elementi vicini  $v_i$  e  $v_j$ . Nel caso della segmentazione dell'immagine, gli elementi in  $V$  sono pixel e il peso di un arco è una misura della dissimilarità tra i due pixel collegati da quell'arco (ad esempio, la differenza di intensità, colore, movimento, posizione o qualche altro attributo locale).

### 3.3 Quickshift

Quick shift è un algoritmo di ricerca rapida che effettua la segmentazione dell'immagine facendo inferenza sui colori, simile al mean shift. L'algoritmo segmenta un'immagine RGB (o qualsiasi immagine con più di un canale) identificando gruppi di pixel nelle dimensioni spaziali e cromatiche congiunte. I segmenti sono locali (superpixel) e possono essere utilizzati come base per ulteriori elaborazioni. Data un'immagine, l'algoritmo calcola una foresta di pixel i cui rami sono etichettati con un valore di distanza. Questo specifica

una segmentazione gerarchica dell'immagine, con segmenti corrispondenti a sottoalberi. I superpixel utili possono essere identificati tagliando i rami la cui etichetta di distanza è superiore a una determinata soglia (la soglia può essere fissata a mano o determinata mediante convalida incrociata).

### 3.4 Watershed

E' una tecnica di morfologia matematica, basata sul principio che ogni immagine in scala di grigio può essere considerata come una superficie topografica (una superficie tridimensionale definita sul piano dell'immagine ed avente l'intensità luminosa come terza dimensione). In tale interpretazione si identificano tre tipi di punti:

1. Punti appartenenti ad una regione di minimo;
2. Punti in cui una goccia d'acqua scorrerebbe verso uno di tali minimi (bacino di raccolta, catchment basin o watershed);
3. Punti in cui l'acqua potrebbe scendere verso più di un minimo con uguale probabilità (linee di separazione, linee di watershed o watershed lines).

L'obiettivo è quello di individuare le linee di watershed sull'immagine. L'idea è semplice: supponiamo che ogni regione di minimo sia perforata e che l'intera topografia sia riempita facendo fluire l'acqua dal basso a velocità costante. Quando la salita dell'acqua sta per unire due distinti bacini adiacenti si costruisce una diga per evitare la fusione. L'insieme finale di tali dighe costituisce le linee di watershed.