

SUPSI

Tecniche base di image processing

Fondamenti di Multimedia Processing

Tiziano Leidi

09.11.2017

Digital image processing

Esistono una moltitudine di tecniche di digital image processing.

Le più comuni si possono suddividere in:

- Trattamento dei colori dell'immagine
- Ricostruzione e ristrutturazione dell'immagine
- Segmentazione
- Operazioni di morfologia matematica
- Riconoscimento di forme e oggetti

Equalizzazione dell'istogramma

L'equalizzazione dell'istogramma è un metodo dell'immagine processing per l'adeguamento del contrasto.

L'obiettivo è di ridistribuire le intensità nell'istogramma, fornendo, come conseguenza, maggiore contrasto a zone che hanno basso contrasto locale.

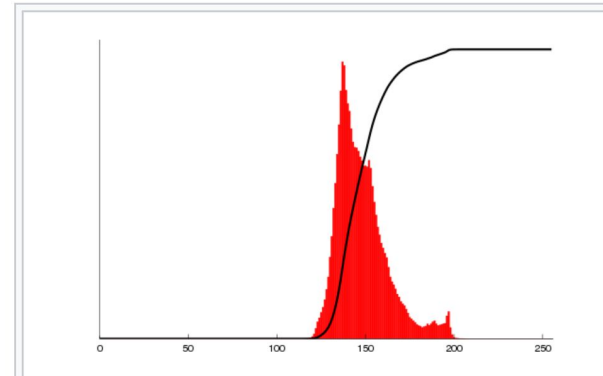
Può essere applicato sia a immagini in scale di grigio, sia a immagini a colori.

L'equalizzazione dell'istogramma può produrre risultati poco realistici nella fotografia. È molto efficace nel caso di immagini scientifiche.

Equalizzazione dell'istogramma



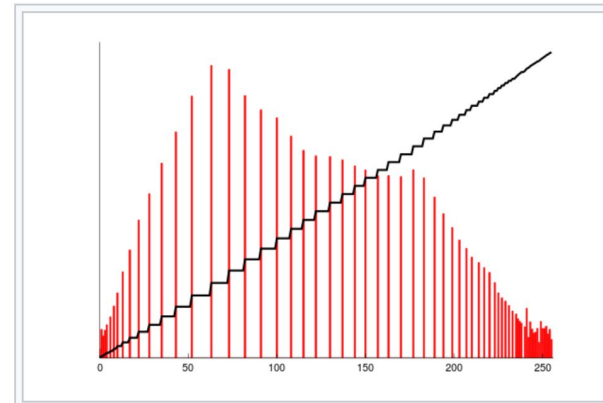
Una immagine non equalizzata



Istogramma corrispondente (rosso) e istogramma cumulativo(nero)



La stessa immagine dopo l'equalizzazione dell'istogramma



Istogramma corrispondente (rosso) e istogramma cumulativo (nero)

Normalizzazione dell'immagine

La normalizzazione è un processo affine all'equalizzazione dell'istogramma, in cui vengono modificate le intensità dei pixels (estendendone il range) per ottenere una migliore occupazione del range dinamico disponibile.

Viene anche chiamato “contrast stretching” o “histogram stretching”, più in generale anche “dynamic range expansion”.

Filtri digitali tramite convoluzione o DFT

Tecniche di convoluzione o di DFT possono essere utilizzate per eseguire varie operazioni sulle immagini, in particolare operazioni di sfocatura (blur) e messa a fuoco (sharpening).

Il filtraggio può essere eseguito nel dominio spaziale tramite convoluzione, applicando dei kernels, o nel dominio delle frequenze, tramite la trasformata di Fourier (DFT), modificando specifiche regioni di frequenze.

Gaussian blur

Nell'immagine processing una sfocatura gaussiana (anche detta Gaussian smoothing) si ottiene applicando all'immagine una funzione gaussiana tramite convoluzione.

Viene spesso utilizzata per ridurre il rumore e la presenza di dettagli nelle immagini.



box blur

Un box blur (anche conosciuto come box linear filter) è un filtro lineare nel dominio spaziale, in cui ogni pixel dell'immagine risultante ha un valore equivalente alla media dei suoi vicini nell'immagine di partenza.

Si tratta di un filtro low-pass, che può essere scritto come matrice determinante 3x3 moltiplicata per 1/9. I box blurs vengono spesso utilizzati per approssimare i Gaussian blur.

$$\frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Sharpening

Un'immagine è a fuoco se ha buona risoluzione e un buon contrasto sui contorni (edges). Più l'immagine ha risoluzione, meglio può essere messa a fuoco. Per quanto concerne il contrasto dei bordi, si tratta di un fattore soggettivo. Può essere migliorato applicando filtri che migliorano il contrasto dei contorni.

Sharpening



Convoluzione

Nel trattamento d'immagine, la convoluzione è il processo di aggiungere ogni elemento ai propri vicini, moltiplicato per i valori specificati in un kernel di dimensione 3×3 .

Tramite questa semplice operazione, si implementa una forma di convoluzione matematica.

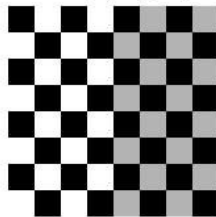
Va sottolineato che la moltiplicazione da eseguire è un array multiplication, non una moltiplicazione matriciale tradizionale.

Kernel

I kernel usati nella convoluzione vengono anche detti matrici di convoluzione o maschere. Possono essere utilizzati per implementare una moltitudine di filtri, incluse alcune soluzioni per la detezione dei contorni.

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

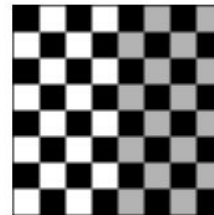
Identità



Originale

$$\frac{1}{9} \times \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

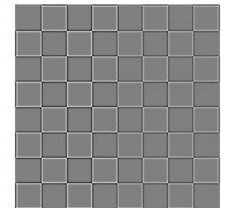
Lowpass



3x3 mean blur

$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

Highpass

Laplacian
edge detection

Normalizzazione del kernel

La normalizzazione dei valori di un kernel è definita come la divisione di ogni elemento del kernel per la somma di tutti gli elementi del kernel, in modo che la somma di tutti i valori all'interno del kernel dia 1.

Questa soluzione garantisce che i pixels siano in media luminosi nell'immagine risultante quanto lo sono nell'immagine originale.

Detezione dei contorni

Il riconoscimento dei contorni (edge detection) ha lo scopo di marcare i punti in cui l'intensità luminosa cambia bruscamente. L'operazione genera immagini contenenti molte meno informazioni rispetto a quelle originali, eliminando la maggior parte dei dettagli non rilevanti alla descrizione della forma e delle caratteristiche strutturali e geometriche degli oggetti.



Trasformata di Fourier 2D

Visto che le immagini contengono dati nel dominio dello spazio (si tratta di uno spazio bi-dimensionale) invece che in quello del tempo, per l'immagine processing viene solitamente utilizzata la trasformata di Fourier di tipo 2D.



Trasformata di Fourier 2D

Come nell'audio, anche la trasformata di Fourier 2D sfrutta funzioni armoniche per eseguire la decomposizione spettrale:

$$F(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-j2\pi(ux+vy)} dx dy,$$
$$f(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u, v) e^{j2\pi(ux+vy)} du dv$$

F è il risultato (lo spettro) della trasformata di Fourier dell'immagine f .

L'operazione inversa permette di ricostruire l'immagine a partire dallo spettro.

Trasformata di Fourier 2D

I parametri u e v , vengono chiamate frequenze spaziali.

$F(u,v)$ è complesso: $F(u,v) = F_R(u,v) + jF_I(u,v)$.

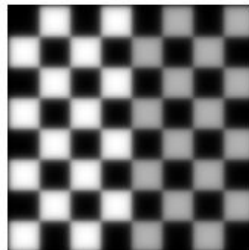
Il valore assoluto è lo spettro di magnitude: $|F(u,v)|$

In maniera analoga si può definire la DFT 2D da applicare alle immagini raster nell'ambito del digital image processing.

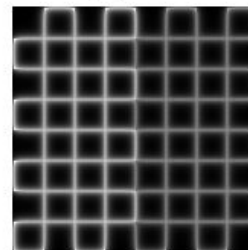
Trasformata di Fourier 2D

Il fatto che la modellazione eseguita dalla DFT sfrutti segnali periodici può presentare delle complicazioni nel caso delle immagini, visto che non presentano necessariamente periodicità.

In maniera simile alla convoluzione, La DFT 2D viene utilizzata per applicare varie tipologie di filtri.



Lowpass



Highpass

Denoise

Immagini acquisite con camere digitali o convertite da film fotografico convenzionale, contengono rumore introdotto da una moltitudine di fattori, come ad esempio i circuiti elettrici, rispettivamente i sensori in caso di illuminazione insufficiente o eccessi di temperatura.

Molto frequentemente il processing di queste immagini richiede la rimozione (almeno parziale) di questo rumore, sia per questioni estetiche, ma anche per questioni prettamente pratiche, come nel caso della computer vision.

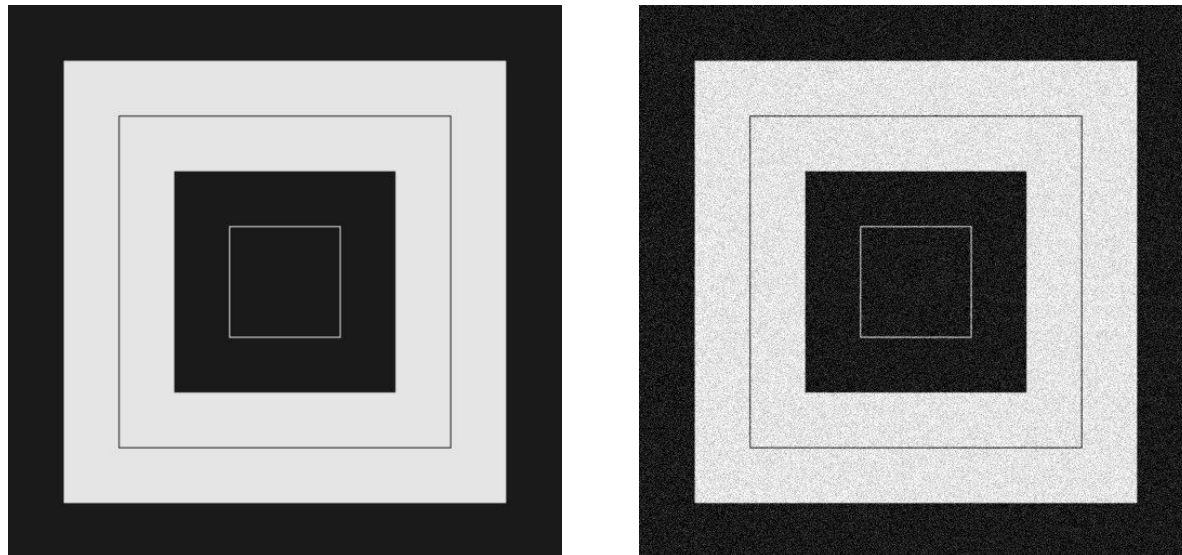
Rumore salt-and-pepper

Il rumore di tipo salt-and-pepper, anche conosciuto come rumore impulsivo, causa disturbi improvvisi e netti all'interno dell'immagine. Si presenta di norma come occorrenze sparse di pixels neri o bianchi. Un metodo efficace per la sua riduzione è l'utilizzo di un filtro mediano.



Rumore gaussiano

Il rumore gaussiano ha una distribuzione statistica di tipo normale. Un caso speciale è quello del rumore bianco, in cui i valori sono tutti indipendenti e distribuiti in maniera identica. Una delle tecniche convenzionali per la rimozione del rumore gaussiano è quella della media tramite convoluzione.



Restoration: Wiener filter

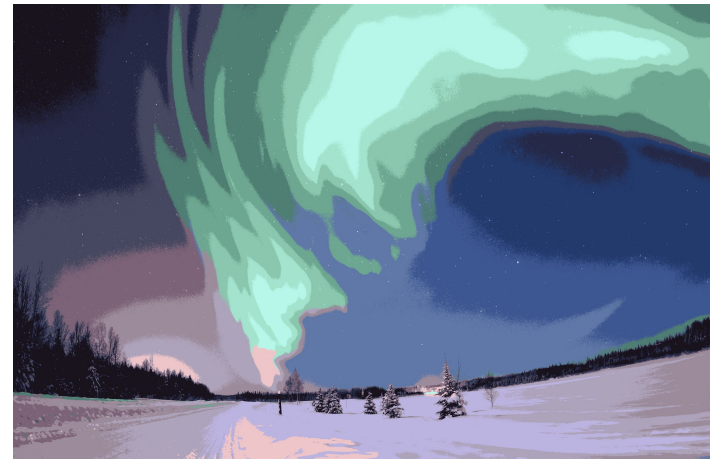
I filtri più tradizionali eseguono la separazione tra segnale e rumore a condizione che questi occupino diverse bande di frequenza. Il filtro di Wiener supera questa limitazione con un approccio statistico.

Si assume di avere conoscenza delle caratteristiche spettrali del segnale originale e del rumore, e si cerca il filtro LTI il cui risultato sia "il più vicina possibile" al segnale originale (minimizzando una prefissata misura dell'errore).

Questo filtro è spesso utilizzato nel processo di deconvoluzione.

Segmentazione

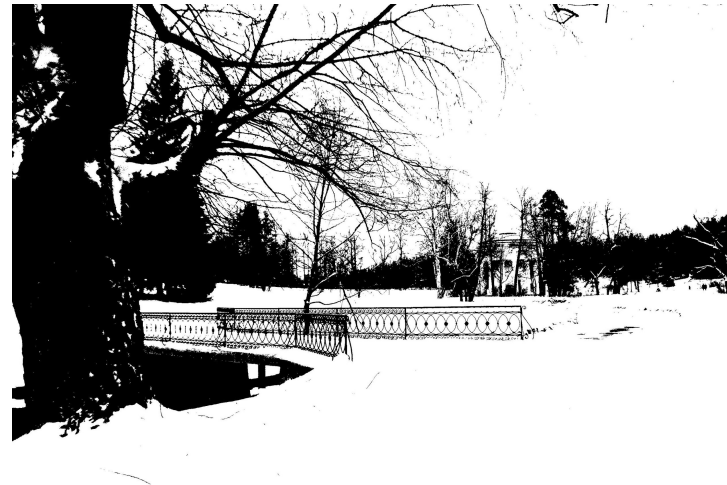
La segmentazione è il processo in cui si suddivide un'immagine digitale in più segmenti (sets di pixel). L'obiettivo è quello di semplificare/modificare la rappresentazione dell'immagine in qualcosa che sia più facile e sensato da analizzare. La segmentazione viene normalmente utilizzata per identificare forme e contorni.



Thresholding

Il metodo più semplice per eseguire una segmentazione è chiamato metodo di sogliatura (thresholding).

Nella sua versione più essenziale, un'immagine in scale di grigi viene convertita in un'immagine con pixels solo bianchi o neri. Il valore usato come discriminante è la soglia (threshold).



Thresholding

Esistono molte tecniche utilizzate per il thresholding, come ad esempio il thresholding bilanciato tramite istogramma, il maximum entropy method, l'Otsu's method (maximum variance) e il k-means clustering.

Thresholding basato sull'istogramma

I metodi di thresholding che sfruttano l'istogramma, sono normalmente efficaci rispetto agli altri metodi, perché necessitano di un solo passaggio di processing sui pixels.

Prima di cominciare il processing, è necessario acquisire un istogramma dell'immagine. Picchi e valli dell'istogramma vengono utilizzati per identificare i clusters all'interno dell'immagine. Può funzionare sia sulla base delle intensità di grigio, sia sulla base dei colori.

K-means

L'algoritmo di K-means sfrutta una tecnica iterativa per partizionare l'immagine in K clusters.

L'algoritmo di base è il seguente:

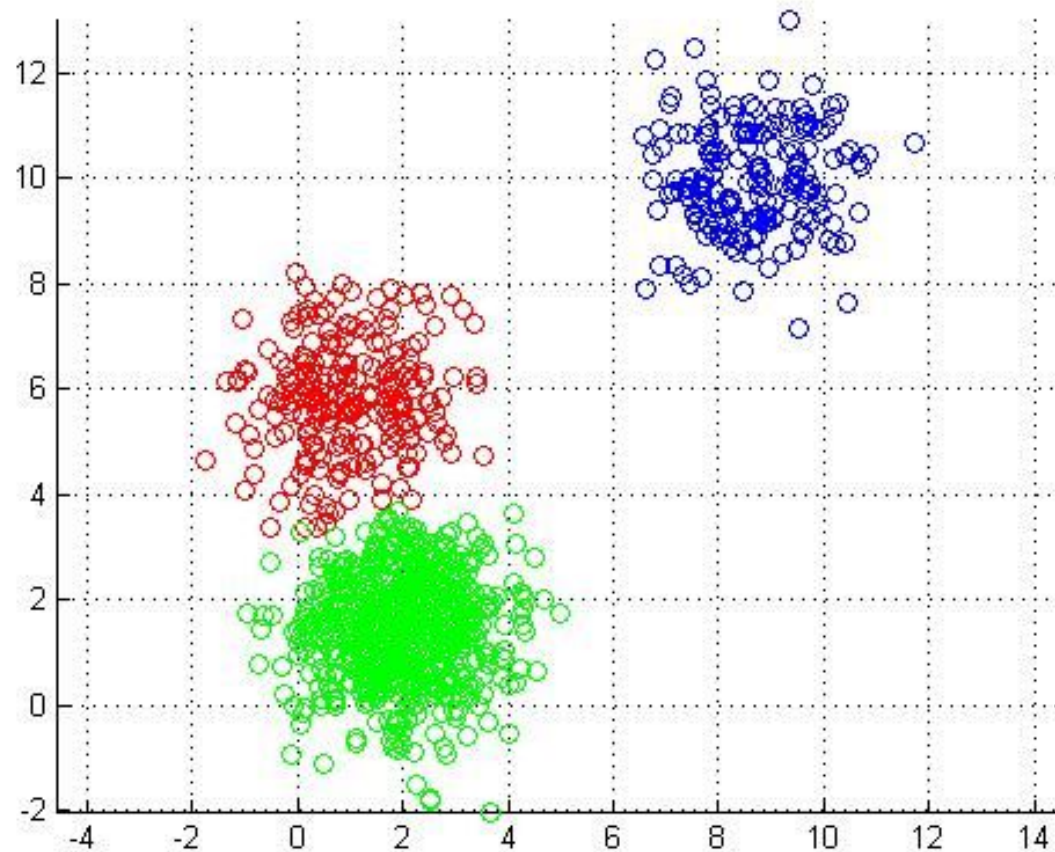
- Scegli K centri dei clusters (a caso o in base ad un'euristica).
- Assegna ogni pixel dell'immagine al cluster che minimizza la distanza fra il pixel e il centro del cluster.
- Ricalcola il centro del cluster posizionandolo nel punto medio di tutti i pixels del cluster.
- Ripeti il passo 2 e il passo 3 fino a convergenza (nessun pixel cambia cluster).

K-means

La distanza viene normalmente calcolata come la differenza assoluta o al quadrato fra un pixel e il centro del cluster.

La differenza può essere basata sulla posizione, l'intensità, il colore, la texture o una combinazione pesata di questi fattori. È garantito che l'algoritmo converga, ma non è detto che restituisca la soluzione ottimale. La qualità della soluzione dipende dal set iniziale di clusters e dal valore di K.

K-means



Morfologia matematica

La morfologia matematica è una teoria e tecnica per l'analisi e il processing di strutture geometriche fondata sulla teoria degli insiemi, dei reticoli, sulla topologia e sulle funzioni random.

La morfologia matematica viene principalmente applicata alle immagini digitali, ma può essere impiegata anche sui grafi, sulle mesh di superficie, sui solidi e su molte altre strutture spaziali.

Morphological image processing

La morfologia matematica sta alla base delle tecniche di morphological image processing, che consistono nell'impiego di alcuni operatori per trasformare le immagini (solitamente di tipo binario, ma si può applicare anche ad immagini in scale di grigio). I quattro operatori principali sono: erosioni, dilatazione, apertura e chiusura.

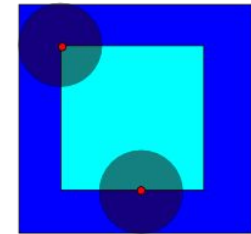
L'idea di base della morfologia binaria è di testare l'immagine con una struttura semplice predefinita, per verificare se questa struttura è compatibile o incompatibile con le forme presenti nell'immagine. La struttura di test viene chiamata structuring element.

Erosione

$$A \ominus B = \{z \in E \mid B_z \subseteq A\} = \bigcap_{b \in B} A_{-b}$$

$$B_z = \{b + z \mid b \in B\}, \forall z \in E$$

Traslazione di B tramite il vettore z

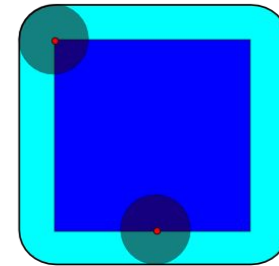


In uno spazio euclideo E , L'erosione dell'immagine A con lo structuring element B è la combinazione di tutti i punti che può raggiungere il centro di B quando si muove all'interno di A .

Esempio: l'erosione di un quadrato di dimensione 10 con un disco di raggio 2 è un quadrato di dimensione 6.

Dilatazione

$$A \oplus B = \bigcup_{b \in B} A_b = \bigcup_{a \in A} B_a$$

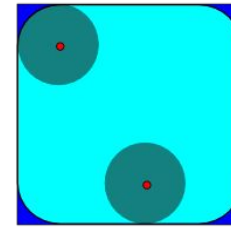


La dilatazione dell'immagine A con lo structuring element B è la combinazione di tutti i punti che può raggiungere B quando il centro di B si muove all'interno di A .

Esempio: la dilatazione di un quadrato di dimensione 10 con un disco di raggio 2 è un quadrato di dimensione 14 con gli angoli arrotondati.

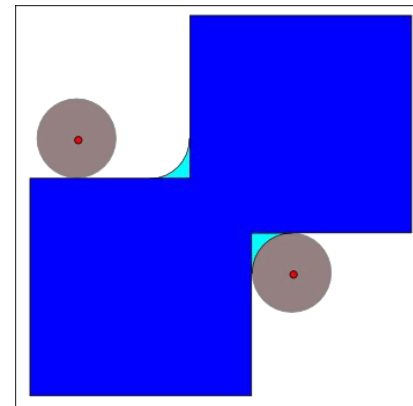
Apertura e chiusura

$$A \circ B = (A \ominus B) \oplus B$$



L'apertura è la traslazione dello structuring element B all'interno di A .

$$A \bullet B = (A \oplus B) \ominus B$$



La chiusura è il complemento della traslazione dello structuring element B all'esterno dall'immagine A .

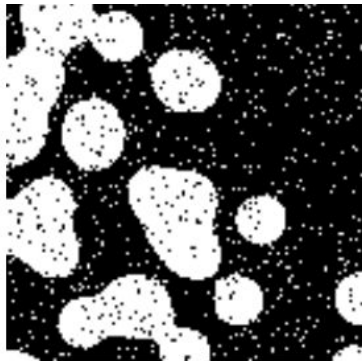
Morphological image processing

L'erosione permette di rendere più regolari i bordi poco precisi, riducendo la dimensione delle forme. Permette anche di individuare eventuali buchi all'interno delle forme.

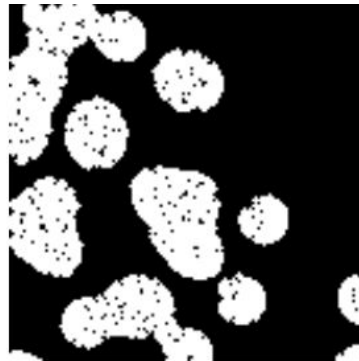
La dilatazione permette di rendere più spesse le forme fini e di coprire eventuali buchi.

Morphological image processing

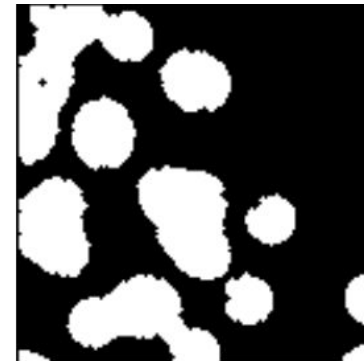
L'apertura permette di ristrutturare le figure rimuovendo le imperfezioni sui bordi e i piccoli buchi. La controparte è che arrotonda le forme. La chiusura permette di isolare le forme rimuovendo eventuale rumore di fondo.



Originale



Apertura



Chiusura