

## INTRODUZIONE

L'**elaborazione delle immagini** riguarda il **miglioramento** dell'informazione grafica e l'**elaborazione** del contenuto per la percezione autonoma delle macchine.

Un'immagine è definibile come una **funzione 2D  $f(x, y)$**  dove **x** e **y** sono le coordinate spaziali nel piano e **f** è definita **intensità** o **livello di grigio**.

Essa è composta da un numero finito di elementi (**pixel**), ciascuno dei quali con una posizione e valore. Il livello di dettaglio dipende sia dalla dimensione della matrice che dal numero di livelli di intensità rappresentabili.

In genere si hanno 256 possibili valori e quindi 8 bit.

Nel mondo reale, un'immagine è un **segnale analogico**, che, per essere analizzata da un calcolatore, deve diventare un **segnale discreto** attraverso un processo di **discretizzazione** che prevede due fasi:

- **Campionamento:** si misura il valore di un segnale continuo in vari istanti di tempo e lo si converte in discreto mediante un'approssimazione. In pratica il valore del segnale nell'istante di tempo analizzato corrisponde al valore discreto utilizzato. Un cattivo campionamento riduce la qualità dell'immagine. Esso si verifica quando non è rispettato il **teorema del campionamento**, ovvero quando si campiona ad un **tasso di campionamento** (numero di campioni prelevati per unità di tempo) minore di un valore pari almeno il doppio della frequenza massima della funzione;
- **Quantizzazione:** terminata la fase di campionamento, i valori del segnale catturati possono assumere un valore compreso in un determinato intervallo. La quantizzazione stabilisce un determinato intervallo che corrisponde ai valori assumibili dal segnale. Una cattiva quantizzazione porta alla creazione di **artefatti** nell'immagine (distorsioni o elementi non presenti nell'immagine originale).

## FASI CHIAVE NELL'ELABORAZIONE DELLE IMMAGINI

1. **Acquisizione:** generare immagini a partire da una sorgente (la luce);
2. **Miglioramento:** manipolare l'immagine migliorandone la qualità;
3. **Ripristino:** tecniche di ripristino dell'immagine originale basate su modelli matematici;
4. **Elaborazione morfologica:** estrapolare le componenti di un'immagine;
5. **Segmentazione:** dividere l'immagine nelle sue parti costituenti.

## TIPI DI IMMAGINI

- **Immagini di riflessione:** informazioni sulla superficie dell'oggetto (radar, sonar);
- **Immagini di emissione:** informazioni interne all'oggetto (infrarossi);
- **Immagini di assorbimento:** informazioni sulla struttura interna dell'oggetto (raggi X).

## MACHINE LEARNING

Il **machine learning** è un sottoinsieme dell'**Intelligenza Artificiale** che ha come idea quella di fornire ai computer la capacità di apprendere dall'esperienza.

Maggiori saranno gli esempi di apprendimento, maggiori le prestazioni.

L'**apprendimento** può essere di due tipi:

- **Supervisionato:** sviluppare un modello predittivo sulla base dei dati di input e output;
- **Non supervisionato:** raggruppare e interpretare i dati sulla base di quelli in input.

Il **reinforcement learning** è un paradigma di apprendimento finalizzato alla costruzione di agenti software in grado di intraprendere azioni al fine di massimizzare un profitto espresso in forma di premio, senza necessitare di dati di input/output o correzioni delle azioni.

Un **sistema esperto** è un programma che cerca di emulare la prestazione di una persona esperta in un determinato campo.

Esso è composto da una **base di conoscenza** (informazioni sul dominio del problema), un **motore inferenziale** (utilizzato per derivare le conoscenze da quelle fornite di base) e un'**interfaccia utente**.

L'intelligenza artificiale è utilizzata nell'elaborazione delle immagini per il riconoscimento di volti e oggetti e per la classificazione di immagini.

## APPRENDIMENTO SUPERVISIONATO

Un algoritmo di apprendimento supervisionato prende un insieme di dati di input (**feature**) ed un insieme di risposte note ad essi e formula un modello per generare predizioni accurate.

Le **variabili quantitative** rappresentano dati basati su valori calcolabili.

Le **variabili qualitative** rappresentano dati su proprietà e attributi e sono indicate con codici numerici: il caso più semplice si ha con due categorie, ad esempio "successo" e "fallimento", e per rappresentarli si possono usare i valori 0 e 1 o -1 e 1 (**target**).

Dato un vettore di input  $X$  ed un vettore di output qualificativi  $Y$ , l'**apprendimento** consiste nel fornire una buona predizione per i valori di  $Y$ .

Esistono due tipologie di apprendimento supervisionato:

- **Regressione:** predire output quantitativi;
- **Classificazione:** predire output qualitativi.

### REGRESSIONE

Si hanno a disposizione i valori di una variabile di input  $X$  (**training set**) e si vuole predire i valori della variabile di output (**target**)  $Y$  tale che  $Y = f(X) + \epsilon$  dove  $\epsilon$  rappresenta l'errore.

Si calcola una funzione che misuri l'**errore** analizzando la predizione effettuata e i reali output.

Supponiamo di voler determinare una curva che sia adatta all'interpolazione dei dati di training.

La fase fondamentale è il **model selection**: di che grado deve essere il polinomio della curva?

Un polinomio di grado alto aumenta l'accuratezza della predizione ma può causare **over fitting**.

L'**over fitting** si verifica quando si usa un modello altamente espressivo ma l'algoritmo riesce ad analizzare perfettamente i dati di training ma non il loro andamento e non riesce a fornire predizioni corrette.

Possibili soluzioni sono abbassare il grado del polinomio o fornire più dati.

Per capire se il modello fornisce buone predizioni, si utilizza un nuovo insieme di dati (**test set**).

Un esempio di regressione è predire il consumo di carburante di un'automobile avendo a disposizione sue informazioni come accelerazione, cilindrata, peso ecc.

### CLASSIFICAZIONE

L'obiettivo è prendere un campione di input  $X$  ed assegnarlo ad una di  $K$  classi discrete.

A differenza della regressione, il valore di output è un valore numerico che identifica la classe.

Si parla di **classi linearmente separabili** quando è possibile tracciare una linea di classificazione (come nel caso di una classificazione binaria) e si usano **vettori di feature** come input.

Per verificare l'esito della classificazione, si realizza una **matrice di confusione**, nella quale sono rappresentate le predizioni e gli effettivi risultati.

Essa ha tante righe e colonne quante sono le classi da predire e le risposte corrette sono sulla diagonale principale.

Per calcolare l'**accuratezza**, si sommano i risultati corretti e si divide per il numero di valori di input.

L'**apprendimento** è definito come la valutazione dell'aumento delle prestazioni con l'**esperienza**.

Esso è definito in funzione del **compito T**, delle **prestazioni P** e dell'**esperienza di addestramento E**.

Un esempio di classificazione è classificare correttamente le tre specie di Iris avendo a disposizione informazioni come la lunghezza e l'altezza del petalo e del setolo.

## APPRENDIMENTO NON SUPERVISIONATO

L'obiettivo è cercare delle similitudini all'interno dei **campioni** osservati così da raggrupparli in **cluster** in base a caratteristiche comuni.

Per permettere il **clustering**, occorre definire i criteri di similarità tra i campioni.

### CLUSTERING K-MEANS

Si cerca di partizionare le osservazioni in un numero predefinito di cluster, definito in partenza.

Il requisito è che tutti i campioni siano in un solo cluster e l'obiettivo è partizionare le osservazioni in K cluster tale che la **varietà intra-cluster totale** sia minima.

L'algoritmo segue una procedura iterativa.

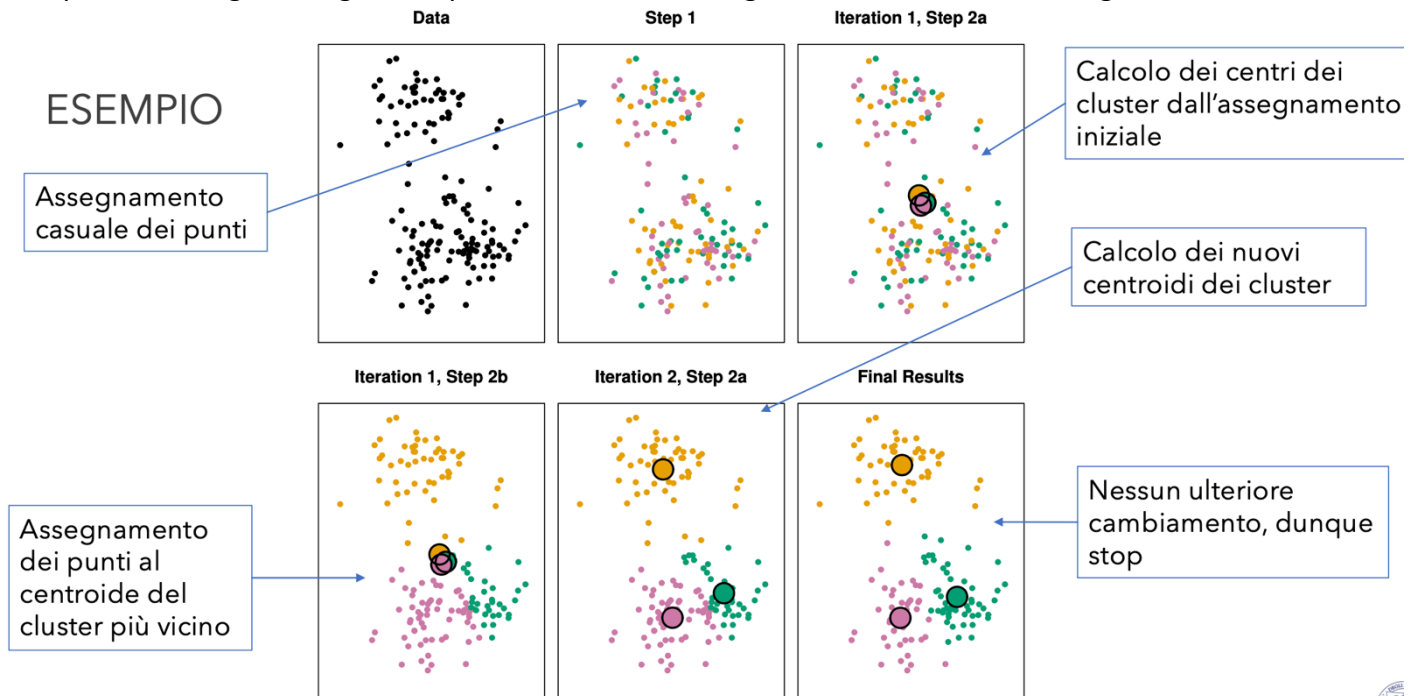
Inizialmente si assegna ogni osservazione ad un cluster casuale.

Si itera l'algoritmo fin tanto che l'assegnamento al cluster rimane stabile (non cambia).

Per ognuno dei cluster si calcola il **centroide** (punto medio) e si assegna ogni osservazione al cluster il cui centroide è più vicino (distanza Euclidea).

Ad ogni passo è garantito che il valore della funzione obiettivo decresce ma non che l'algoritmo fornisca il minimo globale, in quanto la funzione obiettivo è **non convessa**.

È importante eseguire l'algoritmo più volte e con un assegnamento iniziale diverso ogni volta.



## CLUSTERING GERARCHICO

Il clustering K-means ha lo svantaggio di richiedere in anticipo il numero di cluster K.

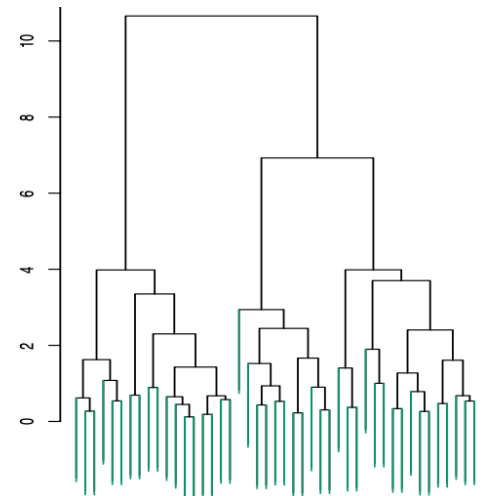
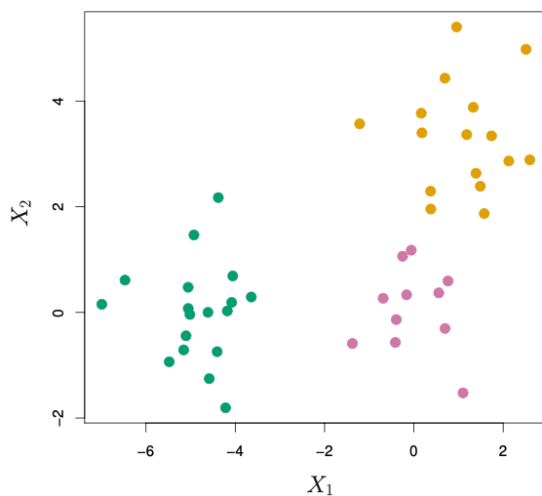
Il **clustering gerarchico** genera un **dendrogramma** che permette di vedere i raggruppamenti ottenuti per ogni possibile numero di cluster.

L'algoritmo comincia dall'inserimento di ogni osservazione in un cluster differente e si procede all'accoppiamento graduale di cluster a due a due, per poi terminare quando tutti i punti sono in un unico cluster.

Per quantificare la distanza tra coppie di elementi e definirne la **dissimilarità** si usano le seguenti tecniche:

- **Complete linkage:** distanza massima tra le osservazioni;
- **Single linkage:** distanza minima tra le osservazioni;
- **Average linkage:** distanza media tra ogni coppia di osservazioni;
- **Centroide:** distanza tra i centroidi delle osservazioni.

Un'alternativa alla distanza Euclidea è la **correlazione**, che considera due osservazioni simili se le loro feature sono correlate.



## FORMAZIONE DELLE IMMAGINI

### L'OCCHIO UMANO

L'occhio umano è circondato da tre membrane.

La **sclera** è il continuo della **cornea** ed è lo strumento di ricezione delle informazioni.

La **coroide** si trova sotto la sclera e riduce la quantità di luci esterne che penetra nel bulbo oculare.

Il **cristallino** è una lente biconvessa capace di modificare la sua forma per regolare la messa a fuoco dei raggi luminosi sulla retina in base alla **distanza focale** dall'oggetto.

La **retina** è la membrana più interna.

È la superficie sulla quale viene impressa l'immagine e sono distribuiti:

- **Coni:** sensibili ai colori e capaci di discriminare i dettagli. Si trovano attorno alla **fovea**;
- **Bastoncelli:** poco sensibili ai colori e attivi in condizione di buio.

L'assenza di ricettori in un punto (20 gradi dalla fovea) implica un **punto cieco**.

L'occhio umano è in grado di percepire un vasto intervallo di livelli di intensità di luce.

Tuttavia, tale intervallo non è coperto simultaneamente e solo un sottoinsieme può essere discriminato.

La luminosità percepita non dipende solo dall'intensità ma anche dalla luminosità ambientale: al di sotto di una certa soglia si percepisce il nero.

## IMMAGINI DIGITALI

Le immagini digitali sono ottenute a partire da un dispositivo di imaging.

Le **immagini ad intensità** misurano la quantità di luce che incide un dispositivo fotosensibile.

I **parametri geometrici**, come tipo di proiezione, posizione ed orientamento, determinano la posizione dell'immagine su cui è proiettato un punto 3D.

I **parametri ottici**, come tipo di lenti e campo di visione, caratterizzano le ottiche del sensore.

I **parametri fotometrici** descrivono tipo, intensità e direzione di illuminazione.

La messa a fuoco di tutti i raggi da un punto di una scena su un singolo punto dell'immagine è realizzabile in due modi:

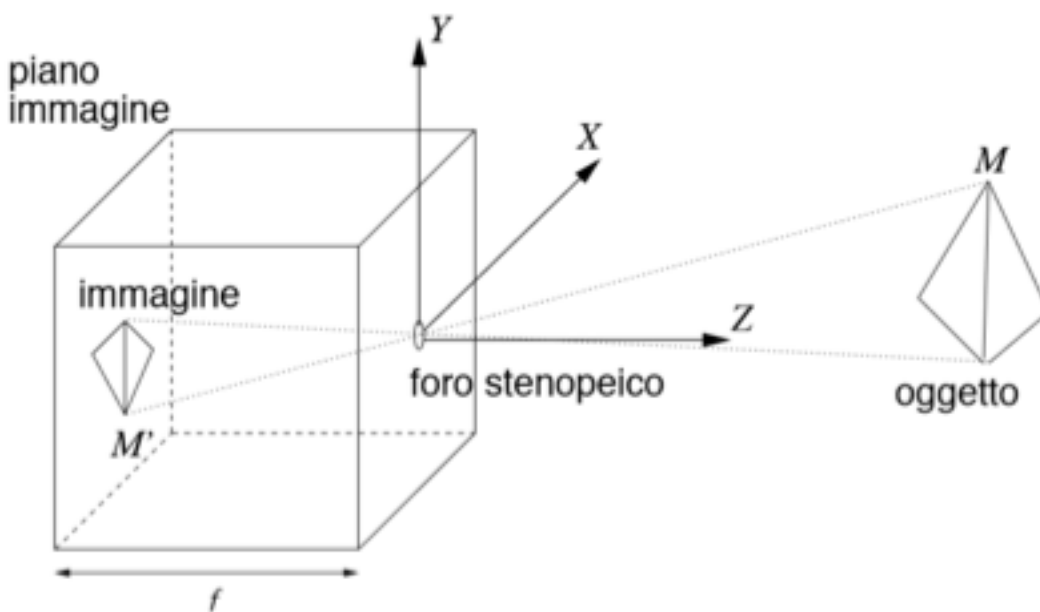
- riducendo l'apertura della camera ad un punto definito **pinhole**: solo un raggio proveniente da un qualsiasi punto della scena entra nella camera;
- introducendo un sistema ottico composto da lenti.

## TELECAMERA PINHOLE

Il modello geometrico più semplice è la **telecamera pinhole**.

Sia **P** un punto della scena di coordinate  $(X, Y, Z)$  e sia **P'** la sua proiezione sul piano immagine di coordinate  $(X', Y', Z')$ . Se **f** è la distanza dal **centro di proiezione O** (**distanza focale**) del piano immagine, si ottiene che i punti O e P sono **collineari**, ovvero giacciono sulla stessa retta.

$$\vec{OP'} = \lambda \vec{OP} \Rightarrow \begin{cases} X' = \lambda X \\ Y' = \lambda Y \\ Z' = -f \end{cases} \Rightarrow \lambda = \frac{X'}{X} = \frac{Y'}{Y} = -\frac{f}{Z} \Rightarrow \begin{cases} X' = -f * \frac{X}{Z} \\ Y' = -f * \frac{Y}{Z} \end{cases}$$



Come si evince dal segno meno, l'immagine è invertita rispetto alla scena, perciò il piano immagine è posto davanti al centro di proiezione.

La Z determina l'**effetto di scorcio**, ovvero la dimensione dell'immagine che varia in funzione della sua distanza dall'osservatore.

## PROIEZIONE PROSPETTICA

Le linee parallele convergono in un punto sull'orizzonte detto **punto di fuga**.

Definiamo una linea che nella scena attraversa il punto  $(X_0, Y_0, Z_0)$  in direzione  $(U, V, W)$  come un insieme di punti  $(X_0 + \lambda U, Y_0 + \lambda V, Z_0 + \lambda W)$ .

La proiezione di un qualsiasi punto da tale linea sul piano di immagine è:

$$p_\lambda = \left( f \frac{X_0 + \lambda U}{Z_0 + \lambda W}, f \frac{Y_0 + \lambda V}{Z_0 + \lambda W} \right)$$

Per  $\lambda \rightarrow \pm\infty$  si ha  $p_\infty = \left( f \frac{U}{W}, f \frac{V}{W} \right)$  e se  $W \neq 0$  è detto **punto di fuga**.

## PROIEZIONE ORTOGRAFICA

Se l'oggetto inquadrato è abbastanza sottile, confrontato con la sua distanza media dalla telecamera, si può approssimare la proiezione prospettica con quella **ortografica**.

Se la profondità  $Z$  di tutti i punti varia in un intervallo  $Z_0 = \pm \Delta Z$ , con  $\frac{\Delta Z}{Z_0} \ll 1$ , le equazioni di proiezione diventano  $X' = sX$  e  $Y' = sY$  con un **fattore di scala prospettico** di  $f/Z$  approssimabile a  $s = f/Z_0$ .

## LENTI SOTTILI

Una lente è in grado di raccogliere più luce di un foro.

Tuttavia, non tutta la scena può essere a fuoco allo stesso tempo.

I raggi paralleli all'asse ottico incidenti sulla lente sono fatti passare per un punto chiamato **fuoco F**.

La distanza del fuoco  $F$  dal **centro C** della lente prende il nome di **distanza focale D**.

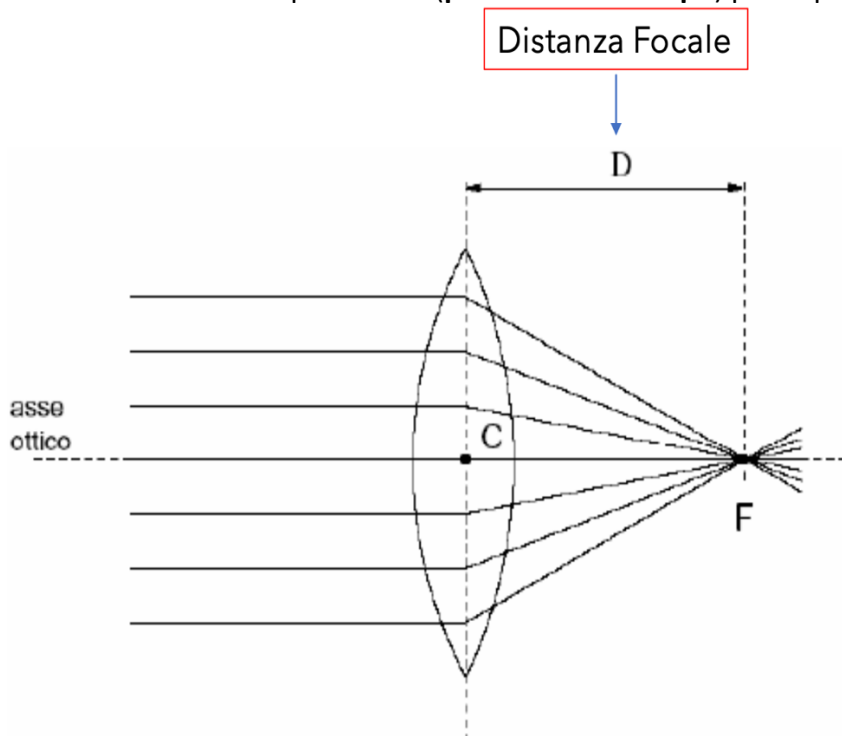
I raggi che passano per il centro  $C$  della lente restano inalterati.

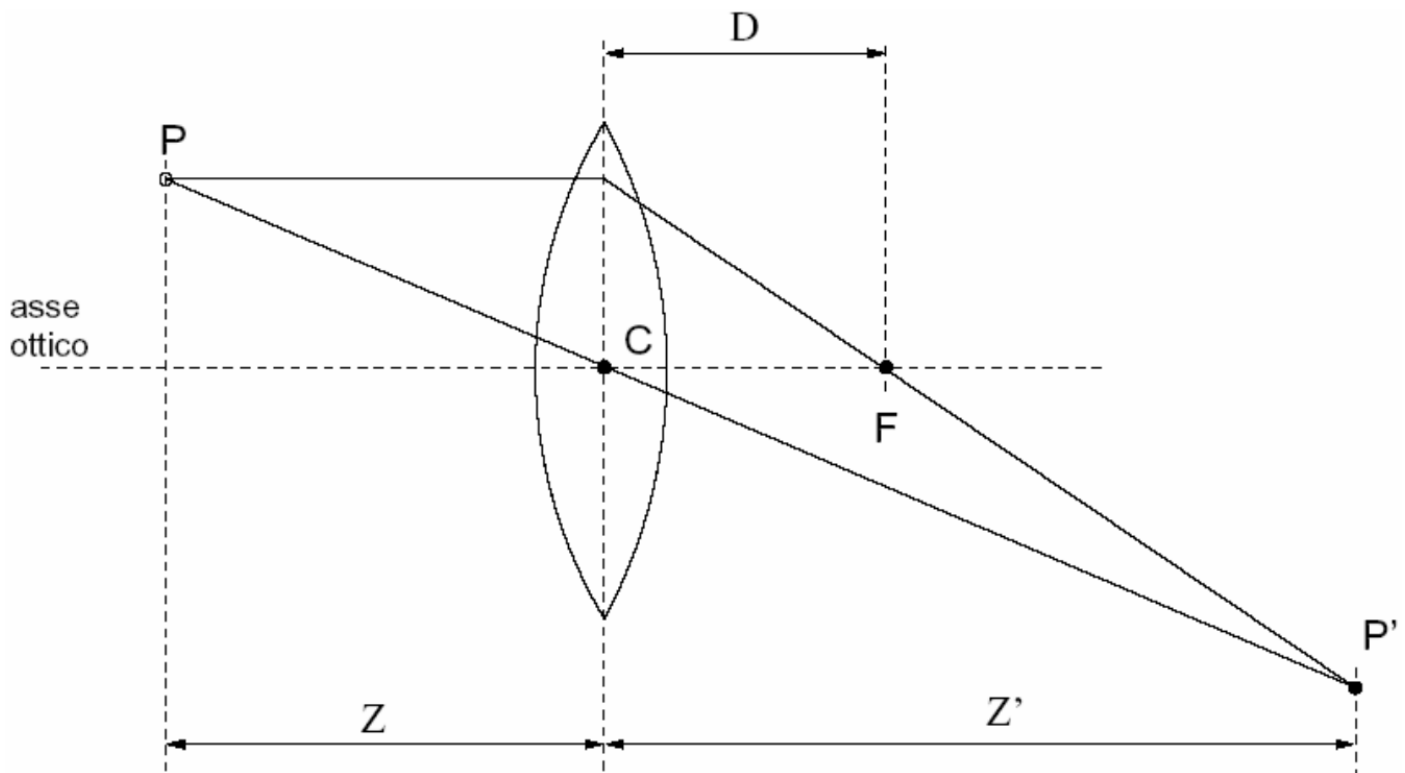
Dato un punto  $P$ , è possibile costruirne l'immagine  $P'$  servendosi del raggio parallelo all'asse ottico che passa per  $F$  e al raggio che passa per  $C$ .

L'**equazione della lente** è data da  $\frac{1}{Z} + \frac{1}{Z'} = \frac{1}{D}$ .

Cambiando  $Z$ , che rappresenta la distanza tra  $P$  e  $C$ , si sfoca l'immagine (**cerchio di confusione**).

Esiste un intervallo di profondità (**profondità di campo**) per il quale i punti sono a fuoco.





## ACQUISIZIONE E RAPPRESENTAZIONE DELLE IMMAGINI

La luce priva di colore è definita **monocromatica** ed è caratterizzata dalla sua intensità, il cui range (nero al bianco) è chiamato **scala di grigio**.

Un sistema di acquisizione di immagini consiste di una **telecamera**, un **frame grabber** e un **calcolatore host**. La telecamera è composta da una lente sottile e da una **matrice Charge-Coupled Device** che costituisce il piano immagine.

Si tratta di una matrice  $n \times m$  di celle ognuna delle quali converte l'intensità della radiazione luminosa incidente in un potenziale elettrico e la cui uscita è un segnale analogico.

Il segnale video analogico è letto dal **frame grabber**, che lo digitalizza, convertendolo in una matrice  $n \times m$  di valori interi definiti **pixel**, memorizzati in un'area di memoria definita **frame buffer**.

Siccome la dimensione della matrice CCD non è strettamente la stessa dell'immagine sottoforma di pixel, la posizione di un punto del piano immagine è diversa se misurata in CCD o in pixel.

Per comodità, si assume una relazione uno ad uno tra pixel ed elementi CCD.

Il calcolatore host estrae l'immagine dal frame buffer e la salva in memoria.

Il numero di bit necessari a rappresentare un'immagine  $M \times N$  a  $L$  livelli di grigio è  $b = MN \log L$ .

Il **range dinamico** di un sistema di imaging è il rapporto tra l'intensità massima misurabile e il minimo livello di intensità rilevabile nel sistema.

La **saturazione** è il fenomeno per cui oltre una certa soglia tutte le intensità sono rappresentate come bianco.

Il **rumore** è l'errore di misurazione del sensore, che esegue una misurazione dell'intensità di luce (pixel neri che diventano grigio scuro).

Il **contrasto** è la differenza di intensità tra il più alto e il più basso livello di intensità di un'immagine.

La **risoluzione spaziale** di un'immagine è la dimensione del più piccolo dettaglio riconoscibile.

La funzione  $f(x, y)$  che rappresenta un'immagine è caratterizzata da:

- **Illuminazione:** quantità di luce incidente sulla scena;
- **Riflettanza:** quantità di luce riflessa dagli oggetti della scena.

$$f(x, y) = i(x, y)r(x, y)$$

Nella pratica, l'intensità è compresa in una **scala di grigio** convenzionalmente scalata in  $[0, L - 1]$  (nero-bianco).

## INTERPOLAZIONE

Spesso è necessario stimare il valore di un'immagine in una posizione non intera.

Il metodo più utilizzato è l'**interpolazione** basata sui valori delle posizioni vicine.

In base al numero di vicini, si parla di interpolazione **nearest neighbour** (valore più vicino), **bilineare** (4 più vicini) e **cubica** (16 più vicini).

All'aumentare dei vicini, aumentano costi e prestazioni.

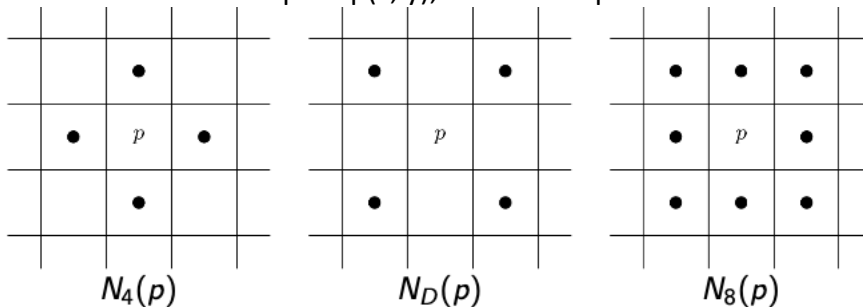
Quando si vuole effettuare il **resize** di un'immagine, per ogni pixel si calcola il valore dell'immagine di input dalla posizione corrispondente e se essa non è intera, si interpola.

## RELAZIONI TRA PIXEL

La struttura di un'immagine digitale permette di definire relazioni tra pixel, legandoli secondo qualche caratteristica.

Un esempio è identificare in un'immagine i pixel che rappresentano il fiume ed ignorare i restanti.

- **Vicinanza:** definito un pixel  $p(x, y)$ , i suoi vicini possono essere



- **Adiacenza:** viene utilizzata l'intensità dei pixel per definire la relazione di similarità. Si divide in 4-adiacenza, 8-adiacenza e m-adiacenza (mixed). L'm-adiacenza è utilizzata per evitare l'ambiguità della 8-adiacenza, che potrebbe portare a confusione nel caso un pixel sia legato a più pixel;
- **Percorso:** il percorso dal pixel  $p(x_p, y_p)$  al pixel  $q(x_q, y_q)$  è un insieme di pixel distinti tali che per ogni  $i$   $(x_{i-1}, y_{i-1})$  e  $(x_i, y_i)$  sono adiacenti.

Definendo  $S$  come sottoinsieme di pixel di un'immagine, due pixel  $p$  e  $q$  sono **connessi** se esiste un **percorso** che li collega, formato interamente da pixel di  $S$ .

Per ogni pixel  $p$  in  $S$ , l'insieme dei pixel ad esso connessi rappresenta una **componente connessa**.

Se in  $S$  c'è solo una componente e questa è connessa,  $S$  è detto **insieme connesso**.

In un'immagine, ogni sottoinsieme connesso è definito **regione**.

0	1	1	0	0	0	0	0	0	3	3	3	0	0	0	0	0
0	0	1	1	0	0	0	0	0	3	3	3	3	3	3	3	3
1	1	1	0	0	0	0	0	0	3	3	3	3	0	0		
0	0	1	1	1	0	0	2	2	2	0	0	0	3	3	3	
0	0	1	1	1	1	1	2	2	2	2	0	3	3	3	0	
0	0	0	1	1	0	0	2	2	2	0	0	3	3	3	0	

## RELAZIONI TRA REGIONI

Due regioni  $R_i$  e  $R_j$  sono **adiacenti** se la loro unione è un insieme connesso.

Regioni non adiacenti sono dette **disgiunte**.

L'unione di tutte le regioni di un'immagine è chiamata **foreground** e il suo complemento, ovvero l'insieme dei pixel che non appartengono ad alcuna regione, **background**.

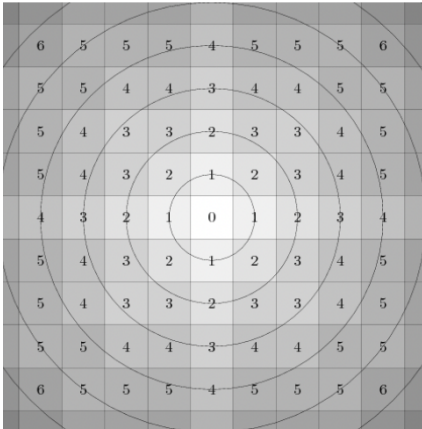
Il **bordo** di una regione  $R$  è composto da pixel che hanno almeno uno dei propri vicini nel background.

L'**edge** è una sottosequenza del bordo.

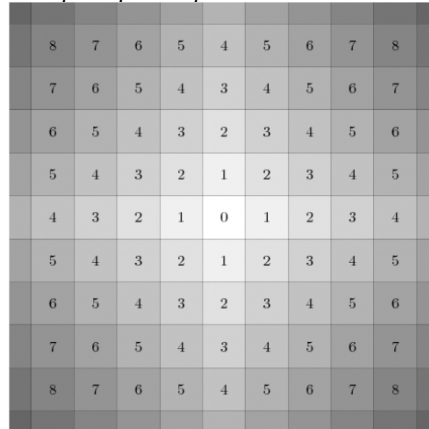


## DISTANZA

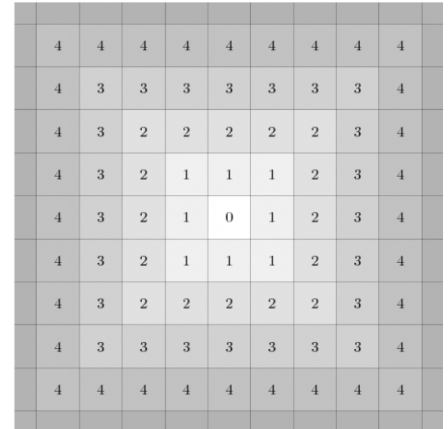
- **Distanza Euclidea:**  $D(p, q) = [(x_p - x_q)^2 + (y_p - y_q)^2]^{\frac{1}{2}}$ ;
- $D_4(p, q) = |x_p - x_q| + |y_p - y_q|$ ;
- **Scacchiera:**  $D_8(p, q) = \max(|x_p - x_q|, |y_p - y_q|)$ .



$D_e$



$D_4$



$D_8$

Un altro modo per definire la distanza è usando la **m-adiacenza**.

In questo caso la distanza tra i due pixel dipende dal valore dei pixel lungo il percorso e da quello dei vicini. La distanza tra i due punti è definita come il **percorso più breve** tra i due punti.

## OPERAZIONI TRA PIXEL

Le immagini sono rappresentate da matrici e le operazioni puntuali su più immagini sono fatte pixel per pixel.

Le operazioni aritmetiche fra immagini sono operazioni puntuali eseguite tra coppie di pixel corrispondenti. Un esempio comune è l'utilizzo della somma per ridurre il rumore delle immagini.

Se fissato un operatore  $H$ , che produce un'immagine di output  $g(x, y)$  a partire da una di input  $f(x, y)$ :  $H[f(x, y)] = g(x, y)$ , e date due costanti arbitrarie  $a$  e  $b$  e due immagini  $f_1(x, y)$  e  $f_2(x, y)$ , si ha

**$H[af_1(x, y) + bf_2(x, y)] = aH[f_1(x, y)] + bH[f_2(x, y)] = ag_1(x, y) + bg_2(x, y)$** , allora  $H$  è detto **operatore lineare**.

Un esempio di operatore **lineare** è la somma, mentre un operatore **non lineare** è trovare il pixel con valore di intensità massimo.

Le restanti operazioni sono definibili come operazioni puntuali (addizione tra immagini per riduzione del rumore e sottrazione per il confronto tra immagini).

## OPERAZIONI SPAZIALI

- **Operazioni sui pixel:** si altera l'intensità dei singoli pixel di un'immagine usando una **funzione di trasformazione**.  $s = T(z)$ , con  $z$  intensità del pixel nell'immagine originale;
- **Operazioni di vicinanza:** usare un filtro su di un insieme di pixel in un intorno (**operazioni locali**). Sono le più utilizzate in quanto poco costose rispetto ad applicare un filtro ad un'intera immagine e il loro risultato è un'immagine con perdita di informazione (livello di grigio non recuperabile);
- **Trasformazioni geometriche:** usate per modificare la disposizione spaziale dei pixel di un'immagine, consistono in una trasformazione delle coordinate e in un'interpolazione, in quanto dopo la trasformazione potrebbe mancare la corrispondenza 1 a 1.

Tutte le altre trasformazioni, definite **affini** (scaling, traslazione, rotazione, distorsione) sono esprimibili attraverso una matrice 3\*3 nella forma:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = A \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Completata la trasformazione, l'interpolazione può essere realizzata mediante:

- **Mapping diretto:** scansione dei pixel dell'immagine di input e calcolo ad ogni posizione della nuova posizione del corrispondente pixel nell'immagine di output. Due o più pixel dell'immagine di input potrebbero essere trasformati nella stessa posizione nell'immagine di output e alcune posizioni potrebbero non corrispondere;
- **Mapping inverso:** scansione dei pixel di output, calcolo della posizione corrispondente all'immagine di input con  $(x, y) = A^{-1}(x', y')$  e interpolazione.

Quando si ruota un'immagine, potrebbe uscire dai contorni.

In questo caso, si può tagliare l'immagine o ingrandirla, con i pixel in background che saranno neri.

## TRASFORMAZIONI DI INTENSITA

Il **dominio spaziale** di un'immagine è il piano che ne contiene i pixel.

Le tecniche del dominio spaziale, definibili come  $g(x, y) = T[f(x, y)]$ , operano direttamente sui pixel dell'immagine (fanno uso delle informazioni contenute nella matrice) e sono meno gravose.

L'**operatore trasformazione T** è definito in un'opportuna area di vicinanza  $(x, y)$  (in genere un rettangolo).

Quando la vicinanza fuoriesce dall'immagine, sono richiesti criteri di estensione (padding).

Se il raggio della vicinanza è 0, la trasformazione coinvolge solo il pixel considerato e dipende dalla sua intensità  $s = T(r)$ .

Le **trasformazioni di intensità** sono utilizzate per migliorare immagini per l'elaborazione visuale.

- **Negativa:** a volte i dettagli sono più evidenti quando l'intensità dei pixel è invertita (ad esempio bianco su nero). Si utilizza la trasformazione negativa  $s = (L - 1) - r$ ;
- **Trasformazione logaritmica:** utilizzata per migliorare il livello di dettaglio nelle zone scure di un'immagine, espandendo il range dei valori di bassa intensità e comprimendo quello dei valori di alta intensità.  $s = c \log(1 + r)$ ,  $c = \frac{L-1}{\log L}$ ;
- **Trasformazione gamma:** usate per correggere l'output del dispositivo di visualizzazione e per la correzione del contrasto.  $s = cr^\gamma$  con un  $\gamma$  opportunamente scelto. In base alla scelta di  $\gamma$ , si può avere un chiarimento o un oscuramento ( $\gamma = 1$  identità,  $\gamma < 1$  schiarimento,  $\gamma > 1$  oscuramento);
- **Trasformazioni di slicing del livello di intensità:** evidenziano un range di intensità (alcune caratteristiche particolari dell'immagine) e impostano tutte le intensità che non sono nel range ad un valore basso, preservando quelle che invece si trovano nel range. Il **bit-plane slicing** considera un'immagine come uno strato di bit, il più alto dei quali contiene la maggior parte delle informazioni. Maggiore sarà il numero di strati, maggiore sarà il livello di dettaglio;
- **Stretching:** estende il range dei valori di intensità dei pixel di un'immagine in modo da aumentare il contrasto e rendere i dettagli più evidenti.

## ISTOGRAMMA DI UN'IMMAGINE

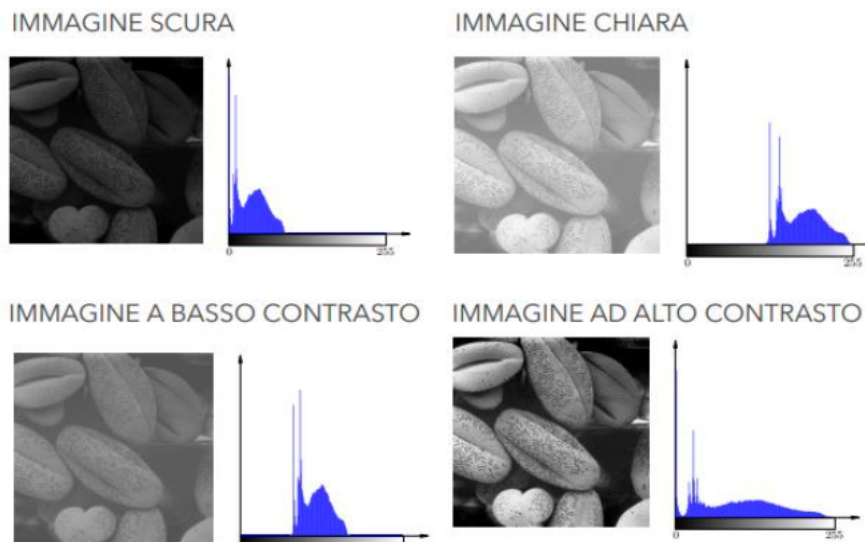
L'**istogramma** di un'immagine a  $L$  valori fornisce uno strumento intuitivo per valutarne alcune proprietà ed è una **funzione discreta**  $h(k) = n_k, k \rightarrow [0, L - 1]$ , con  $n_k$  numero dei pixel con intensità  $k$ .

È utile, ad esempio, se si vuole sapere quanti pixel assumono uno specifico valore di intensità.

Spesso è preferibile considerare l'**istogramma normalizzato**  $p(k) = \frac{n_k}{MN}$ , che è una stima della probabilità di che il pixel assuma un determinato valore di intensità, rispetto al numero di pixel.

L'obiettivo è avere un istogramma con una distribuzione uniforme e pochi picchi, così da avere un'immagine ad alto contrasto e rendere i dettagli facilmente individuabili.

In un'immagine scura, i componenti dell'istogramma sono localizzati a bassi valori d'intensità, viceversa in un'immagine chiara.



## TRASFORMAZIONI MONOTONICHE

Supponendo di avere:

1.  $s = T(r), 0 \leq r \leq L - 1$ , ovvero  $T(r)$  produce un output di intensità  $s$  per ogni pixel dell'immagine di input con intensità  $r$ , con  $r$  compreso tra 0 e  $L - 1$ ;
2.  $T(r_2) \geq T(r_1), r_2 > r_1$ , ovvero  $T(r)$  è una funzione monotona non decrescente;
3.  $0 \leq T(r) \leq L - 1, 0 \leq r \leq L - 1$ , ovvero  $T(r)$  e  $r$  compresi tra 0 e  $L - 1$ .

Si dice che  $T$  è **strettamente monotona crescente** se esiste  $T^{-1}$  tale che:

$$r = T^{-1}(s), 0 \leq s \leq L - 1$$

Se  $s = T(r)$  e  $T(r)$  è **continua** e **differenziabile**, si ha:

$$p_s(s) = p_r(r) \left| \frac{dr}{ds} \right|$$

Applicando la **trasformazione**  $s = T(r) = (L - 1) \int_0^r p_r(w) dw$  si ottiene un'immagine **equalizzata**, ovvero un'immagine la cui distribuzione di intensità è stata modificata per migliorare il contrasto.

Si ha:

$$\frac{ds}{dr} = \frac{dT(r)}{dr} = (L - 1) \frac{d}{dr} \left| \int_0^r p_r(w) dw \right|$$

E quindi:

$$p_s(s) = p_r(r) \left| \frac{dr}{ds} \right| = p_r(r) \left| \frac{1}{(L - 1)p_r(r)} \right| = \frac{1}{L - 1}$$

$S$  sarà uniforme, indipendentemente da  $p_r$ .

Ricordando che gli integrali nel discreto diventano sommatorie e sapendo che  $r_k$  è il livello di intensità in  $[0, L - 1]$ , un'immagine può essere equalizzata assegnando l'intensità  $s_k$  a tutti i pixel di intensità  $r_k$  con:

$$s_k = T(r_k) = (L - 1) \sum_{j=0}^k p_r(r_j) = \frac{L - 1}{MN} \sum_{j=0}^k n_j$$

Questa trasformazione è detta **equalizzazione di istogramma** e soddisfa le condizioni del caso continuo.

## SPECIFICA DELL'ISTOGRAMMA

L'equalizzazione dell'istogramma permette dopo la trasformazione di ottenere un'immagine con una distribuzione di intensità approssimativamente uniforme.

Tuttavia, a volte la distribuzione delle intensità è nota non essere uniforme.

È per tanto utile ottenere un'immagine con una **specificata distribuzione (matching dell'istogramma)**.

Data un'immagine di input, i cui pixel sono distribuiti con una densità di probabilità  $p_r$ , e data la distribuzione di intensità desiderata  $p_z$ , si vuole trovare la trasformazione  $F$  tale che  $z = F(r)$ .

In pratica,  $F$  è la trasformazione che mappa l'intensità di input  $r$  in quella di output  $z$  secondo la specifica desiderata.

Sia  $s$  una variabile casuale tale che  $s = T(r) = (L - 1) \int_0^r p_r(w) dw$ .

Definiamo una funzione  $G$  su  $z$  tale che  $G(z) = (L - 1) \int_0^z p_z(t) dt = s$ .

Si ha  $G(z) = s = T(r)$  e il mapping cercato  $F$  ottenibile come  $z = G^{-1}(T(r))$ .

Manipolati tutti i pixel, la pdf dell'immagine di output sarà approssimativamente uguale a quella trattata. Nel discreto:

$$s_k = \frac{L - 1}{MN} \sum_{j=0}^k n_j, G(z_q) = (L - 1) \sum_{i=0}^q p_z(z_i) = s_k, z_q = G^{-1}(s_k)$$

Non siamo interessati a calcolare  $G^{-1}$ , bensì a calcolare tutti i valori di  $G$  per  $0 \leq q \leq L - 1$ , per poi memorizzarli in una tabella e cercare per ogni  $s_k$  il valore nella tabella più vicino  $z_q$  (soluzione approssimata).

Poiché  $G$  sia invertibile, occorre che sia **strettamente monotona**.

## PROCEDURA DI MATCHING DELL'ISTOGRAMMA

- Ottenere  $p_r$  dall'immagine di input;
  - Ottenere il mapping  $T$  con la relazione di equalizzazione  $s_k = \frac{L-1}{MN} \sum_{j=0}^k n_j$ ;
  - Ottenere il mapping  $G$  dalla  $p_z$  specificata in  $G(z_q) = (L - 1) \sum_{i=0}^q p_z(z_i)$ ;
  - Per ogni  $s_k$  usare i valori  $G$  in tabella per trovare i  $z_q$  corrispondenti, in modo che  $G(z_q)$  sia vicino a  $s_k$ .  
Se ci sono più valori, si sceglie il più piccolo;
  - Applicare la trasformazione  $F$  all'immagine originale tramite il mapping trovato.
- Poiché sia invertibile,  $G$  deve essere strettamente monotona (raro).

## TRASFORMATA DI FOURIER

A volte è più facile lavorare su di una classe di funzioni se ogni funzione è espressa in uno specifico dominio, cioè trasformata in un'altra funzione su un dominio diverso come il **dominio delle frequenze**. Una volta elaborata, la funzione può essere **retro-trasformata** al dominio originale, ovvero il **dominio dello spazio**.

Qualsiasi funzione periodica può essere espressa come somma di funzioni seno e coseno di frequenze diverse, ciascuna moltiplicata per un diverso coefficiente.

Più termini si aggiungono alla serie, più l'approssimazione è precisa.

## NUMERI COMPLESSI

Nell'ambito dell'analisi di Fourier entrano in gioco i **numeri complessi**, definiti come  $C = R + jI$ .

Essi sono visualizzabili geometricamente come punto in un **piano complesso** di cui l'asse delle ascisse è l'**asse reale** (valori di  $R$ ) e l'asse delle ordinate è l'**asse immaginario** (valori di  $I$ ).

## IMPULSO

$$\delta(t) = \begin{cases} \infty, & t = 0 \\ 0, & t \neq 0 \end{cases} \quad e \quad \int_{-\infty}^{\infty} \delta(t) dt = 1$$

In generale, definiamo un **impulso** in un punto  $t_0$  arbitrario come:

$$\delta(t - t_0) = \begin{cases} \infty, & t = t_0 \\ 0, & t \neq t_0 \end{cases} \quad e \quad \int_{-\infty}^{\infty} \delta(t - t_0) dt = 1$$

L'impulso possiede la **proprietà di sifting** che fornisce il valore di  $f(t)$  nella posizione dell'impulso  $t_0$ :

$$\int_{-\infty}^{\infty} f(t) \delta(t) dt = f(0) \quad e \quad \int_{-\infty}^{\infty} f(t) \delta(t - t_0) dt = f(t_0)$$

Per lavorare con funzioni definite solo in un insieme discreto di punti, definiamo l'**impulso discreto**:

$$\delta(x) = \begin{cases} 1, & x = 0 \\ 0, & x \neq 0 \end{cases} \quad e \quad \sum_{x=-\infty}^{\infty} \delta(x) = 1$$

Possiamo introdurre anche un **treno di impulsi**, definito come la somma di infiniti impulsi periodici unitari e separati:

$$s_{\Delta T}(t) = \sum_{n=-\infty}^{\infty} \delta(t - n\Delta T)$$

## SERIE E TRASFORMATA DI FOURIER

La **serie di Fourier** prevede che anche le funzioni non periodiche possano essere espresse come somme di seni e coseni moltiplicati per una funzione peso.

$$f(t) = \sum_{n=-\infty}^{\infty} c_n e^{i\frac{2\pi n}{T}t}, \quad c_n = \frac{1}{T} \int_{-T/2}^{T/2} f(t) e^{-i\frac{2\pi n}{T}t} dt$$

Per la **Trasformata di Fourier**, ogni funzione periodica con periodo  $T$  è rappresentabile come combinazione lineare di seni e coseni.

Sotto certe condizioni, per ogni funzione continua  $f(t)$ , la **trasformata di Fourier** di  $f(t)$  è:

$$F\{f(t)\} = \int_{-\infty}^{\infty} f(t) e^{-i2\pi vt} dt = F(v)$$

Per poter applicare la Trasformata di Fourier, la funzione deve essere **integrabile**, ovvero l'area sottesa sotto la curva deve essere finita.

Data  $F(v)$  è possibile risalire a  $f(t)$  mediante la **trasformata di Fourier inversa** o **antitrasformata**:

$$f(t) = F^{-1}\{F(v)\} = \int_{-\infty}^{\infty} F(v) e^{i2\pi vt} dv$$

Usando la trasformata di Fourier si passa dal dominio dello spazio, in cui le componenti della funzione continua sono descritte da coordinate cartesiane, a quello delle frequenze, in cui le componenti della funzione discreta sono descritte in funzione della loro frequenza, e viceversa per l'antitrasformata.

$F\{f(t)\}$  e  $F^{-1}\{F(v)\}$  sono chiamate **coppia di trasformata di Fourier**.

## TRASFORMATA DI FOURIER DI UN TRENO DI IMPULSI

$$F\{s_{\Delta T}(t)\} = S(v) = \frac{1}{\Delta T} \sum_{n=-\infty}^{\infty} \delta(v - \frac{n}{\Delta T})$$

In pratica, la trasformata di Fourier di un treno di impulsi è un treno di impulsi.

$\frac{1}{\Delta T}$  rappresenta il periodo di campionamento.

## CONVOLUZIONE E TRASFORMATA DI FOURIER

La **convoluzione** tra due funzioni continue,  $f$  e  $h$ , è definita come:

$$f(t) * h(t) = \int_{-\infty}^{\infty} f(\tau)h(t - \tau)d\tau$$

La trasformata della convoluzione di due funzioni continue consiste nel prodotto delle trasformate delle due funzioni:

$$F\{f(t) * h(t)\} = F(v)H(v)$$

Vale anche l'opposto:

$$F\{f(t)h(t)\} = F(v) * H(v)$$

La dimostrazione prende il nome di **teorema di convoluzione**.

Usando il teorema, possiamo calcolare la **trasformata di Fourier di una funzione campionata**:

$$\tilde{F}(v) = F\{\tilde{f}(t)\} = F\{f(t)s_{\Delta T}(t)\} = F(v) * S(v)$$

Dalla proprietà di sifting:

$$\tilde{F}(v) = \frac{1}{\Delta T} \sum_{n=-\infty}^{\infty} F(v - \frac{n}{\Delta T})$$

## CAMPIONAMENTO

Il **campionamento** è un passo di elaborazione richiesto per operare su una funzione con un computer.

Permette di convertire una funzione continua in una sequenza di valori discreti.

La **funzione campionata**  $\tilde{f}(t)$  è modellata come moltiplicazione della funzione considerata  $f$  con un treno di impulsi di opportuno periodo  $\Delta T$ , chiamato **passo di campionamento**:

$$\tilde{f}(t) = f(t)s_{\Delta T}(t) = \sum_{n=-\infty}^{\infty} f(t)\delta(t - n\Delta T)$$

Ciascun componente della sommatoria è un impulso pesato al quale è assegnato il valore assunto da  $f(t)$  nella posizione dell'impulso.

Il valore di ogni campione sarà:

$$f_k = \int_{-\infty}^{\infty} f(t)\delta(t - k\Delta T)dt = f(k\Delta T)$$

## TEOREMA DI CAMPIONAMENTO

Una funzione  $f(t)$  è detta a **banda limitata** se la sua trasformata è 0 per i valori di frequenza esterni ad un intervallo detto **banda**.

Il **tasso di Nyquist** è il tasso di campionamento critico.

Il **teorema del campionamento** afferma che per rappresentare correttamente un segnale continuo nello spazio dimensionale il **tasso di campionamento** deve essere almeno il doppio della frequenza massima in  $f(t)$ .

Se questo requisito non è soddisfatto, si presenta l'**aliasing**.

Si distinguono vari risultati di campionamento:

- **Sovracampionamento**: frequenza di campionamento abbastanza alta da permettere un sufficiente distacco tra i periodi, riuscendo a preservare l'integrità di  $F(v)$ ;
- **Sottocampionamento**: frequenza di campionamento più bassa del minimo richiesto e conseguente sovrapposizione dei periodi;
- **Campionamento critico**: frequenza di campionamento alta appena da permettere di preservare l'integrità di  $F(v)$ . I periodi non sono distaccati ma non arrivano a sovrapporsi.

## ALIASING

Se la funzione non è campionata con il tasso di Nyquist, i periodi si sovrappongono e sono indistinguibili.

Questo effetto è chiamato **aliasing** e nei casi pratici è quasi inevitabile.

Gli effetti di aliasing possono essere ridotti attenuando le componenti ad elevata frequenza prima del campionamento (smoothing prima del campionamento).

L'aliasing è introdotto anche quando un'immagine è **ricampionata** (resize).

Un esempio di aliasing è l'**effetto Moirè**, che vede la sovrapposizione di due pattern periodici.

L'**anti-aliasing** è un filtro utilizzato nelle immagini digitali dopo il campionamento.

## RICOSTRUZIONE DI UNA FUNZIONE CAMPIONATA

Si moltiplica  $\tilde{F}$  per una funzione box  $H(v)$ , in modo da ottenere un periodo di  $F$ :

$$H(v) = \begin{cases} \Delta T, & -v_{max} \leq v \leq v_{max} \\ 0, & \text{altrimenti} \end{cases}, F(v) = H(v)\tilde{F}(v)$$

Si applica la trasformata di Fourier inversa per ottenere  $f$ :

$$f(t) = F^{-1}\{F(v)\} = F^{-1}\{H(v)\tilde{F}(v)\} = h(t) * \tilde{f}(t)$$

La **funzione sinc** è definita come:

$$\text{sinc}(t) = \begin{cases} 1 & \text{se } t = 0 \\ \frac{\sin \pi t}{\pi t} & \text{altrimenti} \end{cases}$$

## TRASFORMATA DI FOURIER DISCRETA

La trasformata di Fourier di una funzione campionata a banda limitata è una funzione continua e periodica che si estende da -infinito a +infinito, il che la rende inutilizzabile nei casi reali.

Pertanto, definiamo la **trasformata di Fourier discreta** per lavorare con un numero finito di campioni.

Ricordando la definizione della trasformata di Fourier:

$$\tilde{F}(v) = \frac{1}{\Delta T} \sum_{n=-\infty}^{\infty} F(v - \frac{n}{\Delta T})$$

$\tilde{F}$  può essere espressa in termini di  $\tilde{f}$ :

$$\tilde{F}(v) = \int_{-\infty}^{\infty} \tilde{f}(t) e^{-I2\pi vt} dt$$

$\tilde{F}$  è una **funzione periodica** e **simmetrica** (funzione pari) e tutte le sue informazioni sono contenute in un singolo periodo. L'obiettivo è campionare un periodo.

Considerando  $M$  campioni di  $\tilde{F}$  in un periodo  $\Delta T$ , sono ispezionate le seguenti frequenze:

$$v_m = \frac{m}{M\Delta T}, \text{ che sostituite in } \tilde{F}(v) = \sum_{n=-\infty}^{\infty} f_n e^{-I2\pi v n \Delta T} \text{ risulta } F_m = \sum_{n=0}^{M-1} f_n e^{-I2\pi m n / M}$$

Questa trasformazione è detta **Trasformazione di Fourier Discreta**, e risulta **simmetrica** e **periodica**.

È possibile ricostruire gli  $M$  campioni usando la **Trasformazione di Fourier Discreta inversa**:

$$f_n = \frac{1}{M} \sum_{m=0}^{M-1} F_m e^{I2\pi m n / M}$$

## CONVOLUZIONE CIRCOLARE

La **convoluzione** di una sequenza finita di  $M$  elementi è definita come:

$$g(x) = f(x) * h(x) = \sum_{m=0}^{M-1} f(m)h(x-m)$$

Questa operazione è definita **convoluzione circolare** e la **periodicità** di  $g$  deriva da quella di  $f$  e  $h$ . Attraverso questa operazione, il **teorema di convoluzione** può essere esteso alla DFT.

$$F\{f(x, y) * h(x, y)\} = F(u, v) * H(u, v)$$

$f(x)$  è composta di  $M$  campioni di  $f(t)$  presi a  $\Delta T$  pezzi.

Pertanto, la sequenza copre un intervallo  $T = M\Delta T$ .

Nel dominio della frequenza, i campioni sono a pezzi di  $\Delta u = \frac{1}{M\Delta T} = \frac{1}{T}$ .

La **DFT** è definita su di un intervallo di frequenze dato da  $\Omega = M\Delta u = \frac{1}{\Delta T}$ .

## IMPULSO IN DOMINIO BIDIMENSIONALE

La Trasformata di Fourier va estesa al dominio bidimensionale perché lavoriamo con immagini, che corrispondono a funzioni bidimensionali.

$$\delta(t, z) = \begin{cases} \infty, & t = z = 0 \\ 0, & t \neq 0, z \neq 0 \end{cases} e \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(t, z) dt dz = 1$$

Vale la **proprietà di sifting**:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(t, z) \delta(t, z) dt dz = 1$$

Anche in questo caso si ha una versione discreta dell'impulso:

$$\delta(t, z) = \begin{cases} 1, & t = z = 0 \\ 0, & \text{altrimenti} \end{cases}$$

Un treno di impulsi è definito come:

$$s_{\Delta T \Delta Z}(t, z) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \delta(t - m\Delta T, z - n\Delta Z)$$

## COPPIA DI TRASFORMATE DI FOURIER CONTINUE 2D

$$F(v, u) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(t, z) e^{-I2\pi(vt+uz)} dt dz$$

$$f(t, z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(v, u) e^{I2\pi(vt+uz)} dv du$$

## TEOREMA DI CAMPIONAMENTO 2D

$$\tilde{f}(t, z) = f(t, z) s_{\Delta T \Delta Z}(t, z) = \sum_{m,n=-\infty}^{\infty} f(t) \delta(t - m\Delta T, z - n\Delta Z)$$

## COPPIA DI DFT 2D

$$F(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-I2\pi(\frac{ux}{M} + \frac{vy}{N})}$$
$$f(x, y) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{I2\pi(\frac{ux}{M} + \frac{vy}{N})}$$



## SPETTRO DI FOURIER E ANGOLO DI FASE

La trasformata di Fourier discreta è esprimibile in forma polare:

$$F(u, v) = |F(u, v)|e^{i\phi(u, v)}$$

$|F(u, v)|$  è detto **spettro di Fourier** e deriva dalla radice quadrata della somma delle parti reali e immaginarie al quadrato:

$$|F(u, v)| = \sqrt{[R^2(u, v) + I^2(u, v)]}$$

$\phi(u, v)$  è l'**angolo di fase** e deriva dall'arcotangente del rapporto tra parte immaginaria e reale:

$$\phi(u, v) = \arctan\left(\frac{I(u, v)}{R(u, v)}\right)$$

Lo **spettro di potenza** è definito come:

$$P(u, v) = |F(u, v)|^2 = R^2(u, v) + I^2(u, v)$$

Poiché si ha:

$$F(0, 0) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y)$$

$$F(0, 0) = MN \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) = MN|\tilde{f}(x, y)| \text{ e quindi } |F(0, 0)| = MN|\tilde{f}|$$

$\tilde{f}$  è il **valore medio** di  $f$ .

**F(0, 0)**, detta **componente direct current** (a frequenza zero) è molto più grande degli altri termini di  $F$  ed è necessaria una trasformazione logaritmica per visualizzarla.

Per quanto riguarda il **teorema di campionamento**, avendo ora un dominio bidimensionale, il treno di impulsi non considera solo un  $\Delta T$  ma anche un  $\Delta Z$ , che corrisponde alla distanza tra i campioni lungo l'asse  $z$  nella funzione continua  $f(t, z)$ .

Lo spettro di Fourier è insensibile alla **traslazione** ma sensibile alla **rotazione** e contiene informazioni sulle intensità dell'immagine.

L'angolo di fase contiene informazioni sulla posizione degli oggetti dell'immagine.

Esso rappresenta lo spostamento della sinusoide rappresentante la frequenza del segnale dell'immagine.

## FILTRAGGIO NEL DOMINIO DELLE FREQUENZE

La Trasformata di Fourier di un'immagine non fornisce una relazione esplicita sulla controparte spaziale, in quanto **F(u, v)** contiene tutti i valori di  $f(x, y)$  modificati da valori di termini esponenziali.

Per questo motivo, è impossibile associare direttamente le componenti di un'immagine e la sua trasformata e si devono considerare determinate relazioni tra le componenti in frequenza della trasformata di Fourier e le caratteristiche spaziali dell'immagine.

Le alte frequenze corrispondono a variazioni repentine, mentre le basse a variazioni lente.

Informazioni sulla scena (direzione dei contorni) possono essere dedotte dallo spettro.

Il **filtraggio nel dominio delle frequenze** è ottenuto modificando la trasformata di Fourier di un'immagine e trasformando all'indietro per ottenere l'immagine elaborata:

$$g(x, y) = F^{-1}\{H(u, v)F(u, v)\}$$

**H(u, v)** è la **funzione filtro**.

**g(x, y)** è l'**immagine filtrata**.

**F, H** e **g** sono array della stessa dimensione  $M \times N$  dell'immagine originale.

## ESEMPIO

La componente **F(0, 0)** corrisponde all'intensità media di  $f$ . Costruiamo un semplice filtro **H(u, v)** tale che sia uguale a 0 quando  $u = v = 0$  e uguale a 1 altrimenti. Impostare a 0 solo  $F(0, 0)$  sposta a 0 la media di  $g$  e la rende più scura.

## FILTRO PASSA BASSO E PASSA ALTO

Un **filtro passa basso** attenua le alte frequenze e lascia inalterate quelle basse, producendo una copia sfocata dell'immagine.

Un **filtro passa alto** fa l'inverso, porta un miglioramento dei dettagli ma un abbassamento del contrasto. Aggiungere una costante al filtro permette che le info associate ad  $F(0, 0)$  siano parzialmente preservate.

Un filtro passa alto si ottiene dal corrispondente filtro passa basso dall'equazione:

$$H_{HP}(u, v) = 1 - H_{LP}(u, v).$$

- Filtro passa basso **ideale**:  $H(u, v) = \begin{cases} 1, & D(u, v) \leq D_0 \\ 0, & \text{altrimenti} \end{cases}$ . Questo filtro lascia passare tutte le frequenze all'interno di un cerchio di raggio  $D_0$  centrato sull'origine e taglia fuori tutte le altre;
- Filtro passa basso **Gaussiano**:  $H(u, v) = e^{-\frac{D^2(u, v)}{2D_0^2}}$ . Minore effetto di smoothing rispetto al filtro di Butterworth ma non introduce ringing;
- Filtro passa basso **Butterworth**:  $H(u, v) = \frac{1}{1 + [\frac{D(u, v)}{D_0}]^{2n}}$ . Maggiore è il valore di  $n$ , maggiore è la presenza di ringing, dovuto al fatto che la discontinuità non è marcata come nel filtro ideale ma graduale;
- Filtro passa alto **ideale**:  $H(u, v) = \begin{cases} 0, & D(u, v) \leq D_0 \\ 1, & D(u, v) > D_0 \end{cases}$ ;
- Filtro passa alto **Butterworth**:  $H(u, v) = \frac{1}{1 + [D_0/D(u, v)]^{2n}}$ ;
- Filtro passa alto **Gaussiano**:  $H(u, v) = 1 - e^{-\frac{D^2(u, v)}{2D_0^2}}$ .

## WRAPAROUND, PADDING E RINGING

La convoluzione circolare di due funzioni periodiche può causare l'**errore di wraparound**, a causa dell'eccessiva vicinanza e la conseguente sovrapposizione dei periodi.

Ad esempio, nel caso 1D la  $x$  potrebbe essere non sufficiente da permettere lo scorrimento di  $f$  su tutto  $h$ . Può essere risolto usando lo **zero padding**: date due sequenze  $A$  e  $B$ , aggiungere gli zero ad esse in modo che entrambe abbiano  $P$  elementi con  $P \geq A + B - 1$ .

Se una funzione non è zero alla fine dell'intervallo, lo zero padding introduce **artefatti** (componenti ad alta frequenza nella trasformata).

Una buona norma è moltiplicare la funzione campionata per una Gaussiana prima del padding.

Il padding, applicato sia all'immagine che al filtro nel dominio spaziale, può evitare errori di wraparound.

Se il filtro è specificato nel dominio delle frequenze, può essere trasformato in quello spaziale, padded e trasformato indietro in quello delle frequenze, ma può verificarsi un fenomeno detto **ringing** (oscillazione).

Esso è dovuto ai troncamenti spaziali del filtro per applicare lo zero padding che introducono una discontinuità.

Se consideriamo la TF della funzione box, ovvero una funzione sinc con frequenze che si estendono all'infinito, anche l'inversa dovrebbe avere componenti spaziali che si estendono all'infinito.

In definitiva, non si può effettuare il padding della rappresentazione spaziale di un filtro definito nel dominio delle frequenze per evitare gli errori di wraparound.

Inoltre, non è possibile utilizzare un filtro ideale nel dominio delle frequenze.

Una migliore opzione è eseguire il padding dell'immagine e definire il filtro nel dominio della frequenza la cui dimensione è la stessa dell'immagine con padding.

Questa soluzione causa errori di wraparound, in quanto non è stato applicato padding per il filtraggio, ma sono preferibili al ringing.

Il **filtro Gaussiano** è l'unico non sottoposto a ringing.

## FILTRI DI SHIFT ZERO PHASE

Filtri che preservano l'angolo di fase per evitare effetti sull'output del filtraggio.

$$F(u, v) = R(u, v) + iI(u, v)$$

$$g(x, y) = F^{-1}\{H(u, v)R(u, v) + iH(u, v)I(u, v)\}$$

## PROCEDURA PER IL FILTRAGGIO DELLE FREQUENZE

- Data un'immagine  $f(x, y)$  di dimensioni  $M \times N$ , si ricavano i parametri di padding  $P$  e  $Q$ . Di solito,  $P = 2M$  e  $Q = 2N$ ;
- Si forma un'immagine padded  $f_p(x, y)$  di dimensioni  $P \times Q$  estendendo  $f(x, y)$  con un adeguato numero di zeri;
- Opzionalmente è possibile moltiplicare  $f_p(x, y)$  per  $(-1)^{x+y}$  per centrare la trasformata;
- Si calcola la DFT di  $f_p(x, y)$ ;
- Si genera una funzione filtro simmetrica e reale  $H(u, v)$  di dimensioni  $P \times Q$  e centro di coordinate  $(P/2, Q/2)$ ;
- Si calcola  $G(u, v) = H(u, v)F(u, v)$ ;
- Si ottiene l'immagine  $g_p(x, y) = \{real[F^{-1}\{G(u, v)\}]\}$ ;
- Opzionalmente si moltiplica  $g_p(x, y)$  per  $(-1)^{x+y}$ ;
- Il risultato finale  $g(x, y)$  è ottenuto mediante cropping della regione  $M \times N$  di  $g_p(x, y)$ .

## PROGETTAZIONE DI FILTRI SPAZIALI

Dato il filtro Gaussiano 1D nel dominio della frequenza:

$$H(u) = Ae^{-\frac{u^2}{2\sigma^2}}$$

Il corrispondente filtro nel dominio spaziale si ottiene prendendo l'antitrasformata di  $H(u)$ :

$$h(x) = \sqrt{2\pi\sigma} Ae^{-2\pi^2\sigma^2 x^2}$$

## LAPLACIANO

Il filtro Laplaciano nel dominio delle frequenze è:

$$H(u, v) = -4\pi^2(u^2 + v^2)$$

L'immagine Laplaciana è ottenuta con:

$$\nabla^2 f(x, y) = F^{-1}\{H(u, v)F(u, v)\} \text{ (anti trasformata del prodotto tra il filtro e l'immagine).}$$

L'immagine filtrata con:

$$g(x, y) = f(x, y) + c\nabla^2 f(x, y), c = -1 \text{ (somma dell'immagine con quella Laplaciana).}$$

## MASCHERAMENTO DI UNSHARP

La tecnica di **mascheramento di unsharp** aumenta nitidezza e contrasto nelle immagini:

$$g_{mask}(x, y) = f(x, y) - f_{LP}(x, y) \text{ (sottrarre all'immagine originale una sua versione sfocata)}$$

$$f_{LP}(x, y) = F^{-1}\{H_{LP}(u, v)F(u, v)\} \text{ (prodotto tra la trasformata di Fourier dell'immagine e un filtro passa basso)}$$

L'immagine filtrata nel dominio spaziale è:

$$g(x, y) = f(x, y) + kg_{mask}(x, y)$$

Nel dominio della frequenza:

$$g(x, y) = F^{-1}\{(1 + k(1 - H_{LP}(u, v)))F(u, v)\}$$

Che espresso in termini di filtro passa alto (sostituendo  $1 - H_{LP}$  con  $H_{HP}$ ):

$$g(x, y) = F^{-1}\{(1 + kH_{HP}(u, v))F(u, v)\}$$

Utilizzando due variabili:

- $k_1 \geq 0$ : distanza dall'origine;
- $k_2 \geq 0$ : contributo delle alte frequenze.

Si generalizza la formula:

$$g(x, y) = F^{-1}\{(k_1 + k_2 H_{HP}(u, v))F(u, v)\}$$

## FILTRAGGIO OMOMORFICO

Si tratta di una tecnica utilizzata per migliorare la qualità di un'immagine correggendone l'effetto di illuminazione.

La funzione di intensità di una scena  $f$  può essere modellata nella scomposizione dell'illuminazione  $i$  e della riflettanza  $r$ :

$$f(x, y) = i(x, y)r(x, y)$$

Questa relazione non può essere sfruttata direttamente per il filtraggio nel dominio della frequenza, ma è necessaria una trasformazione in grado di separare le due componenti, in quanto la trasformata di Fourier di un prodotto non è uguale al prodotto delle trasformate (proprietà valida nel dominio spaziale).

L'idea è utilizzare il logaritmo per scomporre il prodotto in una somma.

$$z(x, y) = \log f(x, y) = \log i(x, y) + \log r(x, y)$$

Per la linearità della DFT si ha:

$$Z(u, v) = F\{z(x, y)\} = F\{\log i(x, y)\} + F\{\log r(x, y)\}$$

Applico il filtro  $H$  ad entrambe le componenti:

$$S(u, v) = H(u, v)Z(u, v) = H(u, v)F_i(u, v) + H(u, v)F_r(u, v)$$

Il processo è realizzato mediante un **filtro omomorfico**:

$$H(u, v) = (\gamma_H - \gamma_L)(1 - \exp(-c \frac{D^2(u, v)}{D_0^2})) + \gamma_L$$

## FILTRI BAND-REJECT E PASSA BANDA

Sono filtri che operano solo su certe bande, escludendo le restanti (filtraggio selettivo).

- **Filtro ideale:**  $H(u, v) = \begin{cases} 0 & \text{se } C_0 - \frac{W}{2} \leq D(u, v) \leq C_0 + \frac{W}{2}; \\ 1 & \text{altrimenti} \end{cases}$

- **Filtro Gaussiano:**

$$H(u, v) = 1 - e^{-[\frac{D^2(u, v) - C_0^2}{D(u, v)W}]^2}$$

- **Filtro Butterworth:**

$$H(u, v) = \frac{1}{1 + [\frac{D(u, v)W}{D^2(u, v) - C_0^2}]^{2n}}$$

$D(u, v)$  è la distanza dal centro del rettangolo della frequenza,  $D_0$  è il raggio del cerchio della banda e  $W$  è la larghezza della banda.

Per ottenere il filtro band-pass dal corrispettivo band reject:

$$H_{BP}(u, v) = 1 - H_{BR}(u, v)$$

## FILTRI NOTCH

Opera su una regione di vicinanza limitata da qualche frequenza e rigetta o lascia passare frequenze specifiche. Un filtro notch con centro  $(u, v)$  deve avere un corrispondente punto notch in  $(-u, -v)$ .

Un **filtro notch di rigetto** è costruito come prodotto di filtri passa alto i cui centri sono traslati nei centri notch:

$$H_{NR}(u, v) = \prod_{k=1}^q H_k(u, v)H_{-k}(u, v)$$

dove  $H_k(u, v)H_{-k}(u, v)$  sono i filtri passa alto con centri  $(u_k, v_k)$  e  $(-u_k, -v_k)$ .

Un **filtro notch passante** è ottenuto da un filtro notch di rigetto mediante:

$$H_{NP}(u, v) = 1 - H_{NR}(u, v)$$

Ad esempio se voglio rimuovere il rumore periodico, posso progettare un filtro notch di valore 0 nella posizione di tali frequenze.

## VANTAGGI DEL FILTRAGGIO NEL DOMINIO DELLA FREQUENZA

La complessità della DFT è  $O((MN)^2)$ , con MN dimensione dell'immagine.

Tuttavia, essa è implementabile mediante FFT (trasformata di Fourier veloce), che ha complessità  $O(MN \log(M))$  (applicabile solo se M e N sono potenze di 2).

Il costo della convoluzione nel dominio spaziale è  $O((mn)(MN))$ , ovvero dal prodotto tra la dimensione dell'immagine e del filtro.

La convoluzione è più conveniente nel dominio della frequenza e risulta meno onerosa per kernel grandi.

In pratica, con kernel piccoli conviene lavorare nel dominio spaziale e viceversa.