



UNIVERSITÀ DI PERUGIA
Dipartimento di Matematica e Informatica



Appunti *Applied Image Processing*

Mr. Vinti

Anno Accademico 2022-2023

Last Update: June 15, 2023

Contents

1 Richiami	6
1.1 Spazio di Lebesgue	6
1.2 Prodotto di Convoluzione	8
1.3 Trasformata di Fourier	9
1.4 Teoria del Campionamento	13
1.4.1 Storia del Campionamento	13
1.4.2 Teorema del Campionamento	14
1.4.3 Svantaggi	17
1.4.4 Operatore Sampling Generalizzato	17
2 Le immagini digitali	20
2.1 Definizione di immagine	20
2.2 Rappresentazione di un'immagine	20
2.3 Discretizzazione	21
2.4 Campionamento di un segnale	21
2.5 Funzione Campionante	22
2.6 Quantizzazione	23
2.7 Immagine Digitale	24
2.8 Campionamento per Immagini	24
2.9 L'aliasing	26
2.10 La risoluzione	28
2.11 Alterazioni della risoluzione	29
2.12 Immagini in bianco e nero e immagini a colori	30
2.12.1 Immagini in bianco e nero	30
2.12.2 Immagini a toni di grigio	30
2.12.3 Immagini a colori	31
2.12.3.1 Lo spazio RGB	31
2.12.3.2 Lo spazio HSV	32
2.13 Relazioni di base fra pixel	33
2.13.1 Intorni	33
2.13.2 Connettività e adiacenza	33



2.13.3	Cammini	34
2.13.4	Cammini e regioni	34
2.13.5	Distanza tra pixel	34
2.14	Zoom e Shrink	35
2.14.1	ZOOM	35
2.14.2	SHRINK	36
3	Elaborazione con Filtri	38
3.1	Elaborazioni Puntuali	39
3.1.1	Thresholding	40
3.1.2	Stiramento	40
3.1.3	Negazione	41
3.1.4	Trasformazione Logaritmica	41
3.1.5	Trasformazione di Potenza	42
3.1.6	Lineare a Tratti	43
3.1.7	Altre Trasformazioni a Tratti	44
3.1.8	Finestramento - Windowing	45
3.2	Modelli delle Immagini	45
3.2.1	Il modello Probabilistico	46
3.2.1.1	L'istogramma dei toni di grigio	46
3.2.1.2	Interventi sull'istogramma	48
3.3	Il Rumore	51
3.3.1	Il modello del rumore	52
3.3.1.1	Il rumore Salt and Pepper	53
3.3.1.2	Rumore di Rayleigh	54
3.3.1.3	Rumore Gamma	55
3.3.1.4	Rumore Esponenziale	56
3.3.2	Signal-Noise Ratio	56
3.4	Filtri	57
3.4.1	Il filtraggio spaziale	57
3.4.2	Filtri lineari	57
3.4.2.1	Filtro medio	58
3.4.2.2	Filtro di media ponderata	58
3.4.2.3	Filtro Gaussiano	58
3.4.3	Filtri non lineari	59
3.4.3.1	Filtro mediano	59
3.4.3.2	Filtro di massimo	60
3.4.3.3	Filtro di minimo	61
3.4.3.4	Filtro di Punto Medio	61
3.4.3.5	Filtro medio alpha-trimmed	61
3.5	Filtri di Sharpening	61



3.5.1	Filtro del Gradiente	62
3.5.2	Filtro di Sobel	63
3.5.3	Filtro di Prewitt	64
3.5.4	Filtro di Roberts	65
3.5.5	Filtro Laplaciano	65
3.5.6	Filtro Differenziale	66
3.6	Operazioni aritmetiche e Image Enhancement	67
4	Image Segmentation	70
4.1	Thresholding Based Segmentation	72
4.1.1	Thresholding manuale	72
4.1.2	Thresholding automatico basilare	73
4.1.3	Metodo di Otsu (1979)	73
4.1.3.1	Descrizione	74
4.1.4	Edge-Gray Levels Histogram method (1977)	76
4.2	Clustering Based Segmentation	77
4.2.1	Fuzzy Selection	77
4.3	Region Based Segmentation	78
4.3.1	Operatori morfologici	78
4.3.1.1	Erosione	78
4.3.1.2	Dilatazione	79
4.3.1.3	Apertura	80
4.3.1.4	Chiusura	81
4.3.1.5	Applicazioni	82
5	Indici di Similarità	85
5.1	Mean Square Error (MSE)	85
5.2	Peak Signal to Noise Ratio (PSNR)	85
6	Image Registration	86
7	Approfondimenti	91
7.1	Compressione di immagini	91
7.1.1	Compressione	92
7.1.1.1	informazioni e dati	92
7.2	Ridondanza	93
7.2.0.1	Ridondanza nella codifica	93
7.2.0.2	Ridondanza interpixel	94
7.2.0.3	Ridondanza psicovisuale	94
7.2.0.4	Ridondanza	95
7.3	Il processo di compressione	95

*”Oي, con quanto sentimento
defeco sul tuo naso,
così che ti coli sul mento.”*

Wolfgang Amadeus Mozart

Chapter 1

Richiami

1.1 Spazio di Lebesgue

Definizione 1. Lo **Spazio di Lebesgue** è definito come:

$$L^1(\mathbb{R}) = \left\{ f : \mathbb{R} \rightarrow \mathbb{R} : \int_{\mathbb{R}} |f(t)| dt < +\infty \right\}$$

Ci riferiamo anche a questo spazio come

L'insieme di tutte le funzioni sempre o assolutamente integrabili in modulo.

N.B. Con la notazione " $< +\infty$ " si intende "*è finito*".

Definizione 2. Definiamo lo spazio delle funzioni sempre o assolutamente integrabili in modulo alla potenza p come:

$$L^p(\mathbb{R}) = \left\{ f : \mathbb{R} \rightarrow \mathbb{R} : \int_{\mathbb{R}} |f(t)|^p dt < +\infty \right\}$$

$L^p(\mathbb{R})$ è anche detto **Spazio di Lebesgue al variare di p** (con $p \geq 1$).

Definizione 3. Definiamo la **norma di L^p** come:

$$\|f\|_p = \left(\int_{\mathbb{R}} |f(t)|^p dt \right)^{\frac{1}{p}}$$

Prendiamo ora la seguente funzione:

$$f :]0, 1] \rightarrow \mathbb{R}, \quad f(x) = \frac{1}{\sqrt{x}}$$

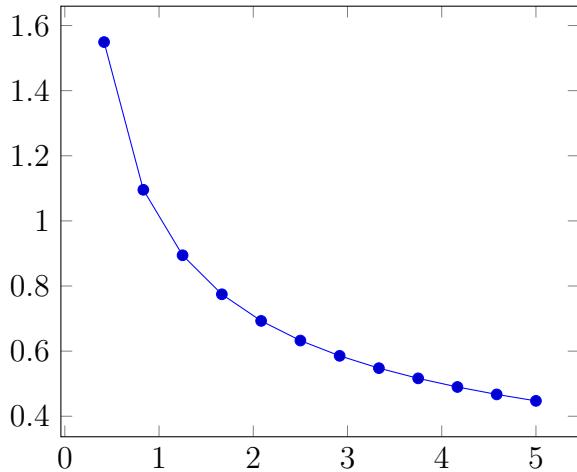
e proviamo a vedere se questa sta nello spazio $L^1([0, 1])$.

$$\begin{aligned}
 \int_0^1 \left| \frac{1}{\sqrt{x}} \right| dx &= \int_0^1 \frac{1}{\sqrt{x}} dx = \lim_{x \rightarrow 0^+} \int_x^1 \frac{1}{\sqrt{t}} dt = \\
 &= \lim_{x \rightarrow 0^+} \int_x^1 t^{-\frac{1}{2}} dt = \lim_{x \rightarrow 0^+} [2\sqrt{t}]_x^1 = \\
 &= \lim_{x \rightarrow 0^+} 2[1 - \sqrt{x}] = 2 < +\infty.
 \end{aligned} \tag{1.1}$$

Concludiamo quindi che $f(x) = \frac{1}{\sqrt{x}} \in L^1([0, 1])$.

Note:

- $\int_0^1 \left| \frac{1}{\sqrt{x}} \right| dx = \int_0^1 \frac{1}{\sqrt{x}} dx$: il modulo scompare perché $\frac{1}{\sqrt{x}}$ nell'intervallo $[0, 1]$ sarà sempre positivo !



Ora ci chiediamo se il prodotto di due funzioni che appartengono a L^1 sta ancora in L^1 . Prendiamo dunque:

$$f(x) \cdot f(x) = \frac{1}{x}$$

e verifichiamo se $\frac{1}{x} \in L^1([0, 1])$:

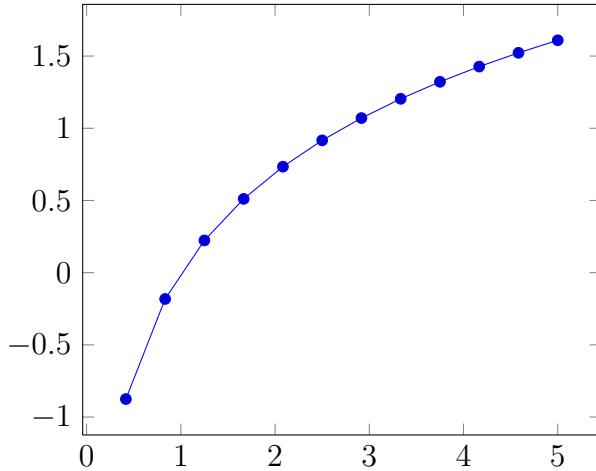
$$\begin{aligned}
 \int_0^1 \frac{1}{x} dx &= \lim_{x \rightarrow 0^+} \int_x^1 \frac{1}{t} dt = \lim_{x \rightarrow 0^+} [\log|t|]_x^1 = \\
 &= \lim_{x \rightarrow 0^+} [\log 1 - \log|x|] = +\infty.
 \end{aligned} \tag{1.2}$$

Deduciamo quindi che non sempre il prodotto di due funzioni che stanno in L^1 appartiene a L^1 .



Note:

- Ricordiamo che $\log 1 = 0$
- Ricordiamo che $\log|x|$ con x che tende a 0^+ (tende a 0 dalla destra) va a $-\infty$



- $\lim_{x \rightarrow 0^+} [\log 1 - \log|x|] = 0 - (-\infty) = +\infty$

Per far sì che il prodotto appartenga a L^1 dobbiamo introdurre il concetto di **Prodotto di Convoluzione**.

1.2 Prodotto di Convoluzione

Il **Prodotto di Convoluzione** è importante perché ci permette di mantenere l'appartenenza allo spazio della singola immagine. Questo, può essere espresso, in modo formale, nel seguente modo:

La *Convoluzione* è una tecnica che consente di regolarizzare le funzioni e di approssimarle in L^p .

Definizione 4. Siano $f \in L^1(\mathbb{R})$ e $g \in L^p(\mathbb{R})$ con $1 \leq p \leq +\infty$. È possibile definire il prodotto di convoluzione tra f e g come:

$$(f \star g)(x) = \int_{\mathbb{R}} f(x-y)g(y) dy$$

e

$$\|f \star g\|_p \leq \|f\|_1 \cdot \|g\|_p.$$

Ovvero, il Prodotto di Convoluzione sta in $L^p(\mathbb{R})$

$$f \star g \in L^p(\mathbb{R})$$

Risulta quindi vero che $\|f \star g\|_p < +\infty$.

N.B. Il Prodotto di Convoluzione eredita le migliori proprietà delle due funzioni.

Diamo ora il seguente lemma riassuntivo che raccoglie tutte le principali proprietà del prodotto di convoluzione.

Lemma 1. Date $f, g, h \in L^1(\mathbb{R})$, risulta:

1. $f \star g = g \star f$ (proprietà commutativa);
2. $f \star (g+h) = (f \star g) + (f \star h)$ (proprietà distributiva rispetto alla somma);
3. $(f \star g) \star h = f \star (g \star h)$ (proprietà associativa);
4. posto $\tau_a f(x) := f(x+a)$, $x \in \mathbb{R}$ (operatore di traslazione,) risulta:

$$\tau_a(f \star g) = (\tau_a f) \star g$$

(invarianza per traslazioni).

1.3 Trasformata di Fourier

In questa sezione daremo la definizione e richiameremo alcune importanti proprietà che riguardano la trasformata di Fourier.

La trasformata di Fourier viene introdotta per poter dare una rappresentazione simile a quella fornita dalle serie di Fourier per le funzioni non periodiche.

Definizione 5. Sia $f \in L^1(\mathbb{R})$. Allora la trasformata di Fourier di f è definita come:

$$\hat{f}(\lambda) := \int_{-\infty}^{+\infty} f(x) e^{-i\lambda x} dx, \quad \lambda \in \mathbb{R}$$



Note:

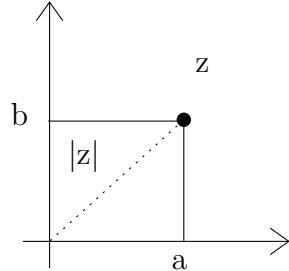
- Ricordiamo che $e^{-i\lambda x} = \cos(\lambda x) - i \sin(\lambda x)$
- Ricordiamo i numeri complessi definiti come:

$$z = a + ib$$

dove a viene detta *parte intera* e b la *parte immaginaria*.

- Il modulo di un numero complesso è definito come:

$$|z| = \sqrt{a^2 + b^2}$$



- Quindi

$$|e^{-i\lambda x}| = (\cos^2(\lambda x) + \sin^2(\lambda x))^{\frac{1}{2}} = 1$$

Ci domandiamo il perché sia così importante che $f \in L^1(\mathbb{R})$ (possibile domanda di esame).

Se $f \in L^1(\mathbb{R})$ allora $\hat{f} \in L^1(\mathbb{R})$.

Dimostrazione. Poiché $f \in L^1(\mathbb{R})$, la funzione $\hat{f}(\lambda)$ risulterà sicuramente ben definita, infatti:

$$\begin{aligned} |\hat{f}(\lambda)| &= \left| \int_{\mathbb{R}} f(t) e^{-i\lambda t} dt \right| \leq \int_{\mathbb{R}} |f(t) e^{-i\lambda t}| dt = \\ &= \int_{\mathbb{R}} |f(t)| \cdot |e^{-i\lambda t}| dt = \int_{\mathbb{R}} |f(t)| dt < +\infty, \forall \lambda \in \mathbb{R} \end{aligned} \tag{1.3}$$

Risulta essere $< +\infty$ se e solo se $f \in L^1(\mathbb{R})$

□



Ma, in generale non possiamo affermare che $\hat{f}(\lambda) \in L^1(\mathbb{R})$.

Ci accorgiamo ora che, nella realtà, molto spesso abbiamo a che fare con segnali *"in banda"* (espressi con la rispettiva trasformata di Fourier) e risulterebbe molto comodo poter riuscire a risalire alla funzione originale. In altre parole ci domandiamo se esiste l'inverso della Trasformata di Fourier.

Teorema 1. *Sia $f \in L^1(\mathbb{R}) \cap C^0(\mathbb{R})$ e tale che $\hat{f} \in L^1(\mathbb{R})$. Allora definiamo la Trasformata Inversa di Fourier come:*

$$f(x) = \int_{-\infty}^{+\infty} \hat{f}(\lambda) e^{i\lambda x} d\lambda, \quad \forall x \in \mathbb{R}$$

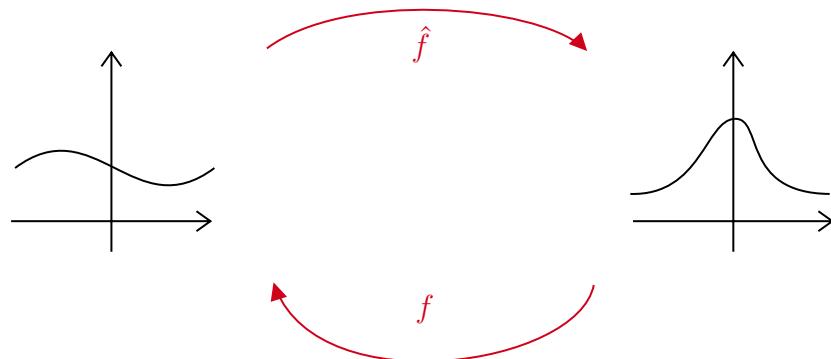
Dimostrazione. Poiché $\hat{f} \in L^1(\mathbb{R})$, la funzione $f(x)$ risulterà sicuramente ben definita, infatti:

$$|f(x)| = \left| \int_{\mathbb{R}} \hat{f}(\lambda) e^{i\lambda t} d\lambda \right| \leq \int_{\mathbb{R}} |\hat{f}(\lambda)| d\lambda < +\infty, \forall x \in \mathbb{R} \quad (1.4)$$

Risulta essere $< +\infty$ se e solo se $\hat{f} \in L^1(\mathbb{R})$

□

Il precedente Teorema permette di esprimere f in termini della sua Trasformata di Fourier. Questo rappresenta la Trasformata Inversa di Fourier e possiamo indicarla anche con \hat{f}^{-1} .



Note:

- $C^0(\mathbb{R}) = \{f : \mathbb{R} \rightarrow \mathbb{R} | f \text{ risulta continua in } \mathbb{R}\}$. Ovvero, $C^0(\mathbb{R})$ è l'insieme di tutte le funzioni che risultano continue in \mathbb{R} .

Analizziamo ora alcune proprietà della Trasformata di Fourier.



Teorema 2. Sia $f \in L^1(\mathbb{R})$, allora la funzione $\hat{f}(\lambda) \in C^0(\mathbb{R})$ e vale:

$$\lim_{\lambda \rightarrow +\infty} |\hat{f}(\lambda)| = 0.$$

Ricordiamo che le funzioni che sono in $L^1(\mathbb{R})$ non è detto che siano continue ! Una funzione può essere integrabile anche se ha un numero finito di punti di discontinuità. Il precedente teorema ci va a dire che se applico la Trasformata di Fourier ad una funzione che non è continua (ma comunque integrabile), allora il risultato sarà sicuramente una funzione continua e che quella funzione sarà integrabile ! Questo è anche visibile dal grafico che otterremmo in quanto avrà le "code" che si schiacciano su zero.

Questa proprietà viene anche chiamata "*Regolarizzazione*".

Teorema 3. Siano $f, g \in L^1(\mathbb{R})$ e $\alpha, \beta \in \mathbb{C}$ allora:

$$\widehat{(\alpha f + \beta g)}(\lambda) = \alpha \hat{f}(\lambda) + \beta \hat{g}(\lambda).$$

Questo teorema ci dice che la Trasformata di Fourier è anche un operatore **Lineare**.

Ci domandiamo ora se esiste qualche legame tra la Trasformata di Fourier e la sua Derivata Prima e se c'è la possibilità di legarle in qualche modo.

Teorema 4. Sia $f \in L^1(\mathbb{R}) \cap C^1(\mathbb{R})$ tale che $f' \in L^1(\mathbb{R})$. Allora:

$$\hat{f}'(\lambda) = i\lambda \hat{f}(\lambda), \quad \forall \lambda \in \mathbb{R}.$$

Note:

- $C^1(\mathbb{R}) = \{f : \mathbb{R} \rightarrow \mathbb{R} \mid \exists f' \text{ continua in } \mathbb{R}\}$. Questo è l'insieme delle funzioni la cui Derivata Prima è continua in \mathbb{R} .

Chiaramente, il risultato del precedente Teorema può essere generalizzato per le derivate di ordine superiore. Infatti:

Teorema 5. Sia $f \in L^1(\mathbb{R}) \cap C^k(\mathbb{R})$ con $k \in \mathbb{N}$, $k \geq 2$ e $f^{(j)} \in L^1(\mathbb{R})$ per ogni $j = 1, 2, \dots, k$, allora risulta banalmente che:

$$\widehat{f^k}(\lambda) = (i\lambda)^k \hat{f}(\lambda), \quad \lambda \in \mathbb{R}.$$



Note:

- $C^k(\mathbb{R}) = \{f : \mathbb{R} \rightarrow \mathbb{R} \mid \exists f^k \text{ continua in } \mathbb{R}\}$. Questo è l'insieme delle funzioni la cui Derivata k-esima e di conseguenza tutte le precedenti sono continue in \mathbb{R} .

Vale il seguente corollario.

Corollario 1. *Sia $f \in L^1(\mathbb{R}) \cap C^k(\mathbb{R})$ e $f^{(j)} \in L^1(\mathbb{R})$ per ogni $j = 1, 2, \dots, k$. Allora:*

$$\lim_{|\lambda| \rightarrow +\infty} |\hat{f}^k(\lambda)| = 0 \rightarrow \lim_{|\lambda| \rightarrow +\infty} |\lambda|^k |\hat{f}(\lambda)| = 0$$

Applicando questo corollario ad $f \in L^1(\mathbb{R}) \cap C^2(\mathbb{R})$ si ha che:

$$\lim_{\lambda \rightarrow \infty} |\lambda|^2 |\hat{f}(\lambda)| = \lim_{\lambda \rightarrow \infty} \frac{|\hat{f}(\lambda)|}{\frac{1}{|\lambda|^2}} = 0.$$

Questo ci dice che $|\hat{f}(\lambda)|$ è un infinitesimo più veloce di $|\lambda|^2$, per λ che tende a infinito, perciò sappiamo che \hat{f} è assolutamente integrabile. Quindi $\hat{f} \in L^1(\mathbb{R})$. Questo corollario ci fornisce anche una condizione sufficiente per l'integrabilità di $\hat{f}(\lambda)$.

Teorema 6. *Sia $f \in L^1(\mathbb{R}) \cap C^0(\mathbb{R})$. Se $\hat{f}(\lambda) = 0, \forall \lambda \in \mathbb{R}$, allora $f(x) = 0, \forall x \in \mathbb{R}$.*

Corollario 2. *Se due funzioni $f, g \in L^1(\mathbb{R}) \cap C^0(\mathbb{R})$ e $\hat{f}(\lambda) = \hat{g}(\lambda), \forall \lambda \in \mathbb{R}$, allora $f = g$ in \mathbb{R} .*

Teorema 7. *Siano $f, g \in L^1(\mathbb{R})$. Allora,*

$$\widehat{f * g}(\lambda) = \hat{f}(\lambda) \hat{g}(\lambda), \quad \forall \lambda \in \mathbb{R}.$$

Il precedente teorema ci mostra che la Trasformata di Fourier del prodotto di convoluzione è uguale al prodotto delle trasformate. Questo è un importante teorema perché, a livello computazionale, è molto costoso computare la trasformata di Fourier del prodotto di convoluzione, ma è molto più semplice effettuare una banale moltiplicazione tra due funzioni.

1.4 Teoria del Campionamento

1.4.1 Storia del Campionamento

Completare



1.4.2 Teorema del Campionamento

Diamo la definizione del classico *WKS-sampling theorem* o *Teorema del Campionamento*. Definiamo prima il **supporto** di una funzione f come:

$$\text{sup } f = \{x \in \mathbb{R} : f(x) \neq 0\}.$$

Teorema 8. *Sia $f \in L^2(\mathbb{R}) \cap C(\mathbb{R})$ una funzione continua, ad energia finita e a banda limitata, i.e., tale che $\text{sup } \hat{f} \subset [-\pi W, \pi W], W > 0$. Allora vale la seguente formula di interpolazione.*

$$f(t) = \sum_{k=-\infty}^{+\infty} f\left(\frac{k}{W}\right) \text{sinc}[\pi(Wt - k)], \text{ con } t \in \mathbb{R}$$

dove:

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

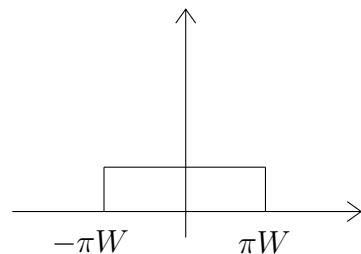


Figure 1.1: Una funzione a banda limitata

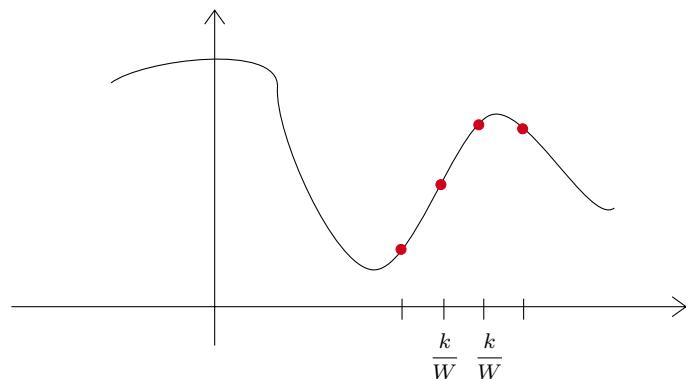


Figure 1.2: Esempio di applicazione del campionamento



In altre parole, il teorema del campionamento ci dice che, da una famiglia discreta di valori campione, assunti in modo uniformemente spaziato su tutto l'asse reale, è possibile ricostruire in modo esatto, un segnale ad energia finita (ovvero in $L^2(\mathbb{R})$) e a banda limitata, mediante la serie interpolante introdotta sopra.

$\Delta t = 1/W$ denota l'*intervallo temporale* tra i campioni, $\pi W/2\pi = w/2$ denota l'*ampiezza di banda* di f , che è misurata in cicli per unita di tempo (Hertz) e $R = 1/\Delta t$ denota l'*ordine di campionamento (sampling rate)*, misurato in campioni per unita di tempo.

H. Nyquist mise in rilievo il significato dell’“intervallo” (numero) $1/W$ nella telegrafia e Shannon lo chiamò *intervallo di Nyquist* corrispondente alla frequenza di banda $[\pi W, \pi W]$. L’ordine di campionamento R , legato all’intervallo di Nyquist, rappresenta un minimo teorico per la ricostruzione del segnale. Per essere più precisi, il teorema sampling costituisce un algoritmo che permette di rispondere alle due seguenti domande:

1. *Quanti campioni sono necessari per assicurare che l’informazione contenuta nel segnale venga preservata, durante il processo di ricostruzione?*

Se il segnale contiene componenti di alte frequenze, è necessario campionarlo almeno due volte la massima frequenza del segnale per evitare perdite di informazione, cioè $f_s \geq 2f_H$, dove f_s è la frequenza sampling, f_H è la più alta frequenza del segnale e $f_s = 2f_H$ è la frequenza (ordine) di Nyquist. Campionare ad un ordine più alto di $2f_H$ significa utilizzare una successione più fine di valori campione, e questo è il caso del cosiddetto sovraccampionamento (oversampling). Nella pratica il sovraccampionamento deve essere effettuato molto spesso, poiché un fattore correttivo deve essere introdotto per compensare il fatto che il campionamento e l’interpolazione non possono coincidere con i corrispondenti valori teorici; questo è il caso del cosiddetto *round-off error* o *errore di quantizzazione*. Inoltre un altro errore temporale si presenta quando i campioni non possono essere presi esattamente agli istanti temporali k/W ; questo errore è chiamato *time-jitter error*. Un esempio concreto di sovraccampionamento è rappresentato dal un lettore di compact disk. Infatti considerando che la più alta frequenza udibile è approssimativamente intorno a 18.000 cicli al secondo (Hertz), a seconda dell’ascoltatore, in base alla teoria precedente un segnale audio deve essere campionato almeno 36.000 volte al secondo per far sì che il segnale venga ricostruito completamente. L’ordine di campionamento dei primi lettori di compact disk era di circa 44.000 campioni al secondo.



Altre volte capita di non avere a disposizione un numero ragionevole di valori campione per ricostruire il segnale completamente. Questo avviene quando non conosciamo il segnale originale, ma dobbiamo ricostruirlo dalle informazioni che abbiamo a disposizione del segnale (cioè dai suoi valori campione). In questo caso, cioè quando il numero di informazioni (valori campione) è scarso, la distanza tra i nodi è più grande dell'intervallo di Nyquist ed abbiamo il cosiddetto sottocampionamento (undersampling). In questo caso si presenta il fenomeno di aliasing e sostanzialmente ciò che succede è che le replicanti spettrali si sovrappongono.

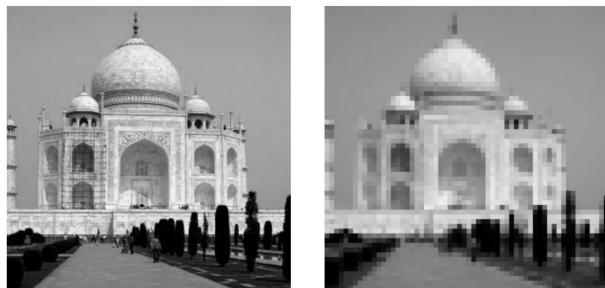


Figure 1.3: Nella figura possiamo apprezzare un esempio di aliasing.

2. *Dato un segnale, campionato ad un certo fissato ordine di campionamento, quali frequenze non devono essere contenute nel segnale al fine di preservare l'informazione?*

Questa è la situazione inversa rispetto a quella presentata nella prima domanda. La risposta è che un segnale campionato ad un fissato ordine di campionamento non può contenere componenti a frequenze superiori alla metà dell'ordine di campionamento, cioè: $f_H \leq f_s/2$, dove f_H è la frequenza più alta contenuta nel segnale e f_s è la frequenza sampling. In altre parole, la più alta frequenza che puo essere accuratamente rappresentata è la metà della frequenza sampling. Quindi il teorema sampling fornisce un limite superiore per la più alta frequenza f_H . Spieghiamolo con un esempio: supponiamo di utilizzare un CD-rom con un fissato ordine di campionamento di 44.1KHz. Nyquist afferma che la più alta frequenza che può essere riprodotta da questo campionamento è di 22.05KHz. Infatti le frequenze superiori a 22.05KHz producono aliasing. Per evitare questo, può essere utilizzato un filtro passa basso per bloccare le frequenze superiori a 22.05KHz. È importante notare che l'aliasing può essere prodotto sia da segnali campionati non corretta-

mente (con un numero di campioni non sufficiente) che da frequenze troppo alte del segnale averti un fissato ordine di campionamento.

1.4.3 Svantaggi

Benchè il teorema del campionamento venga laramente impiegato in ambito applicativo esso presenta svantaggi o limiti da questo punto di vista.

- Le ipotesi fatte sul segnare da ricostruire f , ovvero che sia ad energia finita e a banda limitata, implicano che ci riduciamo a considerare una famiglia di segnali estremamente regolari e ci consente di ricostruire solo sengali appartenenti alla classe $C^\infty(\mathbb{R})$. Nelle applicazioni reali, i segnali difficilmente sono così regolari. Per esempio, nelle immagini, forti salti di livello nella scala di grigi (vedi bordi degli oggetti), matematicamente possono essere visti come delle discontinuità.
- Per ricostruire un segnale f al tempo t è necessario un numero di campioni infinito e il loro valore viene rilevato sia all'istante precedente a t che a quelli successivi. Dal punto di vista applicativo, una prima difficoltà potrebbe essere quella di implementare una formula teorica che prevede infiniti termini. Il problema potrebbe essere risolto troncando la serie; tuttavia, questo introdurrebbe un errore (di troncamento) nella ricostruzione, che non garantirebbe piu l'interpolazione del segnale ma una mera approssimazione. Inoltre, nei problemi reali, i segnali sono noti solo nel “passato”, ovvero i valori campione di f sono noti, in generale, solo negli istanti di tempo k/W precedenti a t . I valori di $f(k/W)$ nel “futuro”, ovvero quando $k/W > t$, non sono in generale disponibili.
- Per il principio di Hisemberg i segnali non possono essere contemporaneamente a durata e banda limitata, quindi i segnali nel mondo reale, in generale, non soddisfano le ipotesi del teorema del campionamento.

1.4.4 Operatore Sampling Generalizzato

Per poter rendere applicabile il teorema del campionamento ed evitare le problematiche viste in precedenza, è stato studiato un *operatore di tipo Sampling Generalizzato*, con lo scopo di fornire una versione approssimata del teorema. L'idea di base è quella di sostituire la funzione *sinc* con una funzione φ a supporto continuo e compatto, il che implica che la serie si riduce ad una somma finita per ogni $t \in \mathbb{R}$ e dunque, per ricostruire un dato segnale in un certo istante di tempo, sono necessari solo un numero finito di valori



campione.

Ora per dimostrare che i campioni possono essere assunti nel passato solo rispetto a t forniremo due definizioni:

Definizione 6. Sia $\varphi \in C_c(\mathbb{R})$ e $f \in C(\mathbb{R})$, definiamo la famiglia di operatori $(S_W^\varphi f)_{W>0}$ con

$$(S_W^\varphi f)(t) := \sum_{k=-\infty}^{+\infty} f\left(\frac{k}{W}\right) \varphi(Wt - k), \quad t \in \mathbb{R}$$

L'operatore S_W^φ è detto *Operatore Sampling Generalizzato*, mentre φ è detta *Nucleo* dell'operatore.

Definizione 7. Data la funzione $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ e un numero $\beta > 0$, possiamo definire

$$m_\beta(\varphi) := \sup_{u \in \mathbb{R}} \sum_{k \in \mathbb{Z}} |u - k|^\beta |\varphi(u - k)|$$

che prende il nome di *Momento assoluto discreto di ordine β* della funzione φ .

Grazie alla definizione appena introdotta otteniamo la seguente stima:

$$|(S_W^\varphi f)(t)| \leq \sum_{k=-\infty}^{+\infty} |f\left(\frac{k}{W}\right)| |\varphi(Wt - k)| \leq \|f\|_\infty \sum_{k=-\infty}^{+\infty} |\varphi(Wt - k)| \leq \|f\|_\infty m_0(\varphi)$$

$\forall t \in \mathbb{R}$.

Poichè f è limitata e il momento $m_0(\varphi) \leq +\infty$ risulta che l'operatore di sampling generalizzato risulta ben definito in $C(\mathbb{R})$. Vale quindi il seguente risultato di approssimazione.

Teorema 9. Sia $\varphi \in C_c(\mathbb{R})$ tale che

$$\sum_{k=-\infty}^{+\infty} \varphi(u - k) = 1, \quad \forall u \in \mathbb{R} \tag{1.5}$$

Allora, se la funzione limitata $f : \mathbb{R} \rightarrow \mathbb{R}$ è continua in $t_0 \in \mathbb{R}$, risulta:

$$\lim_{W \rightarrow \infty} (S_W^\varphi f)(t_0) = f(t_0)$$

In particolare, se $f \in C(\mathbb{R})$ si ha:

$$\lim_{W \rightarrow \infty} \|S_W^\varphi f - f\|_\infty = 0$$



Da qui deriva il corollario che ci mostra com'è possibile effettuare un processo di ricostruzione grazie a campioni rilevati esclusivamente nel passato.

Corollario 3. *Sia $\varphi \in C_c(\mathbb{R})$, con $\sup \varphi \subset \mathbb{R}^+$ e tale che sia soddisfatta la condizione (1.5), allora per ogni funzione limitata $f : \mathbb{R} \rightarrow \mathbb{R}$, nei punti t_0 di continuità si ha:*

$$\lim_{W \rightarrow \infty} (S_W^\varphi f)(t_0) = \lim_{W \rightarrow \infty} \sum_{\substack{k \\ W < t_0}} f\left(\frac{k}{W}\right) \varphi(Wt_0 - k) = f(t_0)$$

Dimostrazione. Come conseguenza del fatto che $\varphi(Wt_0 - k) = 0$ per $\frac{k}{W} \geq t_0$ risulta che

$$(S_W^\varphi f)(t_0) = \sum_{\substack{k \\ W < t_0}} f\left(\frac{k}{W}\right) \varphi(Wt_0 - k)$$

Quindi la tesi segue banalmente per il teorema precedente. \square

Note

- φ si legge *Phi*.
- $C_c(\mathbb{R})$ indica l'insieme delle funzioni continue a supporto compatto.



Chapter 2

Le immagini digitali

2.1 Definizione di immagine

Definizione 8. *Un'immagine è una rappresentazione grafica di valori numerici.*

In dettaglio un'immagine è una funzione bi-dimensionale $f(x, y)$, dove le variabili (spaziali) x e y sono valori reali che definiscono la posizione dei punti nell'immagine e $f(x, y)$ è in genere un valore reale che definisce l'intensità dell'immagine nel punto (x, y) .

Ad ogni punto che andiamo a definire con le coordinate (x, y) viene associata, a seconda del tipo dell'immagine, una tonalità di grigio o i livelli di intensità dei tre colori principali: Rosso, Verde e Blu.

Tutti i colori rappresentati dal calcolatore possono essere scomposti in combinazioni di 3 colori principali: **Rosso, Verde e Blu (RGB)**. Dove:

$$R = f_1, \quad G = f_2, \quad B = f_3$$

Note:

- In natura i tutti i colori si ottengono a partire da **Rosso, Giallo e Blu (RYB)**, ma al computer possiamo ottenere un *"giallo sintetico"* partendo dal Verde.

2.2 Rappresentazione di un'immagine

La funzione f che rappresenta l'immagine può essere a valori in \mathbb{R} , in \mathbb{R}^2 o in \mathbb{R}^3 , a seconda del tipo di immagine.

- **Immagine in scala di grigi:** $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ (funzione scalare)
- **Immagine a colori:** $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ (funzione vettoriale)

Ovvero:

$$f(x, y) = [f_1(x, y), f_2(x, y), f_3(x, y)]$$

dove le componenti f_i , con $i = 1, 2, 3$ si dicono canali.

Se vogliamo rappresentare una scena in movimento, ottenendo cioè un' **immagine dinamica**, è necessario introdurre una terza variabile, quella **temporale** (t), per cui si lavora con una funzione $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$.

$$f(x, y, t) = [f_1(x, y, t), f_2(x, y, t), f_3(x, y, t)].$$

Nelle immagini **Analogiche** conosco l'intensità di ogni livello di grigio in ogni punto. Le immagini mostrate al calcolatore invece vanno **DISCRETIZZATE!!**

2.3 Discretizzazione

Se si vuole utilizzare un calcolatore elettronico per lo studio di un segnale, è necessario **discretizzare** la funzione $s(t)$ che rappresenta il segnale. Infatti un calcolatore elettronico è in grado di trattare solo segnali discreti, cioè successioni di campioni i cui valori sono rappresentati con precisione finita. Se si lavora con un segnale continuo $s(t)$, per implementarne lo studio al calcolatore è necessario passare ad un opportuno segnale discreto.

Ciò avviene utilizzando il procedimento di **campionamento**, che consiste nel discretizzare la variabile temporale t . Inoltre, è anche necessario discretizzare i valori che la funzione $s(t)$ assume (**quantizzazione**).

Nel caso delle immagini applicare i processi di **campionamento** e **quantizzazione** significa passare da un'immagine **analogica** ad un'immagine **digitale**.

2.4 Campionamento di un segnale

Il campionamento di un segnale può essere fatto in 2 diversi modi:

1. **Nel tempo:** Il campionamento di un segnale si ottiene prelevando i valori che il segnale assume soltanto in istanti temporali fissati, in genere individuati tramite una funzione periodica (**funzione campionante**). La successione dei valori campionati di s fornisce una rappresentazione **discreta** (nel tempo) di $s(t)$.
2. **Nello spazio:** Un'immagine può essere vista come una funzione $f(x, y, t)$ dello spazio e del tempo e dunque è necessario discretizzare anche le variabili spaziali. Si ottiene in questo modo una matrice a tre dimensioni, delle quali due sono spaziali ed una è temporale.

2.5 Funzione Campionante

In genere, si assume che il campionamento sia **uniforme**, sia dal punto di vista spaziale che temporale, ovvero che la funzione campionante sia periodica di periodo costante.

Fissiamo gli intervalli di campionamento Δx , Δy , Δt appropriati (dal Teorema Sampling e dalla teoria di Nyquist), ovvero la distanza tra due campioni successivi lungo le coordinate x , y e t .

Indichiamo con N , M , T le dimensioni della matrice dei valori campionati dell'immagine. Infine possiamo dare le seguenti

Definizione 9. *La funzione campionante è*

$$s_c(x, y, t) = \sum_{j=1}^M \sum_{k=1}^N \sum_{h=1}^T \delta(x - j\Delta x, y - k\Delta y, t - h\Delta t)$$

Definizione 10. *L'immagine campionata è*

$$\begin{aligned} s_c(x, y, t) &= f(x, y, t)s_c(x, y, t) = \\ &= f(x, y, t) \sum_{j=1}^M \sum_{k=1}^N \sum_{h=1}^T \delta(x - j\Delta x, y - k\Delta y, t - h\Delta t) \end{aligned} \quad (2.1)$$

Lo scopo della funzione campionante $s_c(x, y, t)$ è di prelevare i valori campionati dal segnale continuo di partenza e pertanto ha un caratteristico andamento **pulsante**.

Ricontrollare le seguenti rime!

- Il segnale **non va mai letto** quando x cade nel nodo della funzione in quanto non si sarebbe in grado di leggerlo.



- Il segnale **va letto** soltanto in $\frac{j}{w}$ ovvero la funzione campionante parallela ai campioni.

Ricontrolare questi 2 punti.

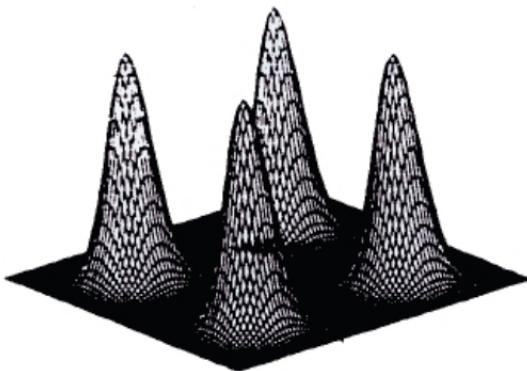


Figure 2.1: Nell'immagine possiamo osservare l'andamento pulsante della funzione campionante.

2.6 Quantizzazione

Per ottenere una completa discretizzazione di un'immagine è necessario discretizzare, oltre al dominio, anche l'insieme immagine (insieme dei valori).

Definizione 11. Si definisce **quantizzazione** il procedimento di discretizzazione dei valori della funzione che rappresenta un'immagine, cioè il passaggio da valori continui a valori discreti.

Per le immagini a toni di grigio si parla di **grey level quantization**, mentre per le immagini a colori si parla di **color depth**, in riferimento al numero di bit utilizzati per ciascun canale di colore (8, 16, 24, 32 bit).

- **Esempio 1:** Le immagini che siamo abituati a vedere tutti i giorni sui nostri cellulari sono immagini a colori a 8bit.
- **Esempio 2:** Nelle immagini mediche, di solito in formato **DICOM**, le immagini vengono rappresentate a 16bit ma gli ultimi 4 bit dell'immagine sono riservati ad informazioni personali che servono ad identificare il paziente che ha sostenuto l'esame.

2.7 Immagine Digitale

Tramite il campionamento e la quantizzazione è possibile definire un'immagine digitale come segue:

Definizione 12. Una immagine digitale è una rappresentazione di matrici di elementi immagine, detti anche pixel (pixel = picture elements). Dove

- Il **pixel** costituisce la componente elementare della matrice, dove gli indici di riga e colonna indicano i valori delle due variabili spaziali, cioè la posizione di un punto nell'immagine.
- Ogni elemento della matrice contiene i valori che rappresentano l'intensità dei corrispondenti punti nell'immagine, anche detta **luminanza**.

2.8 Campionamento per Immagini

L'Immagine campionata è rappresentata tramite la seguente formula:

$$f_c(x, y) = f(x, y)s_c(x, y) = f(x, y) \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} \delta(x - j\Delta x, y - k\Delta y)$$

dove $s_c(x, y)$ è la **funzione campionante** e δ rappresenta funzione **delta di Dirac**.

Definire
funzione di
Dirac

Si può provare che c'è una relazione tra \hat{f}_c e \hat{f} . Per questo è importante assumere che lo spettro del segnale f sia a **banda limitata**, cioè:

$$\hat{f}(\omega_x, \omega_y) = 0 \text{ per } |\omega_x| > \bar{\omega}_x \text{ e } |\omega_y| > \bar{\omega}_y$$

dove $\bar{\omega}_x$ e $\bar{\omega}_y$ definiscono la banda rettangolare dell'immagine.

Così lo spettro dell'immagine campionata consiste nello spettro dell'immagine continua infinitamente ripetuta nel piano delle frequenze, in una griglia di risoluzione $(\frac{2\pi}{\Delta x}, \frac{2\pi}{\Delta y})$, dove:

$$(\frac{2\pi}{\Delta x}, \frac{2\pi}{\Delta y}) = (w_{xc}, w_{yc})$$

sono le **frequenze Sampling**.

Per ricostruire esattamente un segnale campionato, la frequenza del campionamento non deve essere inferiore ad una **frequenza minima** (ovvero



frequenza sampling), che corrisponde ad un valore massimo per ciascuno degli intervalli Δx , Δy .

Tale valore minimo deve essere almeno pari al doppio della banda massima di f , cioè:

$$\omega_{xc} \geq 2\bar{\omega}_x \text{ e } \omega_{yc} \geq 2\bar{\omega}_y \quad (2.2)$$

o equivalentemente:

$$\frac{1}{\Delta x} \geq 2\bar{\omega}_x \text{ e } \frac{1}{\Delta y} \geq 2\bar{\omega}_y$$

Se nella (2.2) vale l'uguaglianza, allora si dice che l'immagine è **campionata alla sua frequenza di Nyquist**.

Se Δx e Δy sono più piccoli del richiesto criterio di Nyquist, l'immagine risulta sovraccampionata (**oversampling**). Nel caso contrario, l'immagine non può essere ricostruita esattamente: si parla di sottocampionamento (**undersampling**) e si presenta un fenomeno di distorsione detto **aliasing**.

Note:

- il valore minimo è un valore puramente teorico. Nella pratica, non potendo in generale determinare con precisione la banda massima del segnale, si utilizzano frequenze di campionamento più alte. Spesso si campiona con una frequenza pari a 4 volte quella misurata.

Teorema 10 (Teorema del Campionamento per Immagini). *Sia $f(x, y)$ un'immagine*

- *a banda limitata e ad energia finita, soddisfa quindi*

$$\hat{f}(\omega_x, \omega_y) = 0 \text{ per } |\omega_x| > \bar{\omega}_x \text{ e } |\omega_y| > \bar{\omega}_y;$$

- *con f uniformemente campionata in una griglia rettangolare con intervalli spaziali Δx , Δy ,*
- *che abbia l'ordine di campionamento più grande dell'ordine di Nyquist, cioè*

$$\omega_{xc} \geq 2\bar{\omega}_x, \omega_{yc} \geq 2\bar{\omega}_y$$

allora la f può essere ricostruita dai suoi valori campione $f(j\Delta x, k\Delta y)$. Inoltre, l'immagine ricostruita è data dalla seguente formula di interpolazione:



$$f(x, y) = \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} f(j\Delta x, k\Delta y) \left(\frac{\sin(xw_{xc} - j)\pi}{(xw_{xc} - j)\pi} \right) \left(\frac{\sin(yw_{yc} - k)\pi}{(yw_{yc} - k)\pi} \right)$$

2.9 L'aliasing

Per ricostruire esattamente una immagine, è necessario limitare in banda l'immagine che deve essere campionata, campionando all'ordine di campionamento di Nyquist o più grande e interpolando appropriatamente i valori immagine.

Se c'è sovrapposizione di spettri, risultante dal sottocampionamento, vuol dire che componenti spettrali spurie sono state introdotte nel processo di ricostruzione. L'effetto che si ottiene è chiamato aliasing.

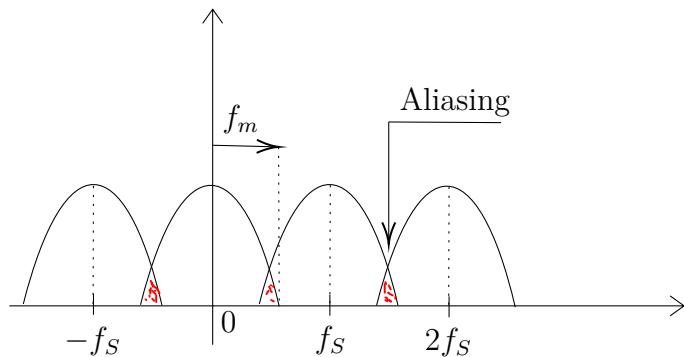


Figure 2.2: Visualizzazione della sovrapposizione degli spettri che genera aliasing

Quindi l'aliasing è la presenza di componenti spettrali (frequenze) indesiderate nella ricostruzione dell'immagine, componenti che non erano presenti quando l'immagine originale era stata campionata.

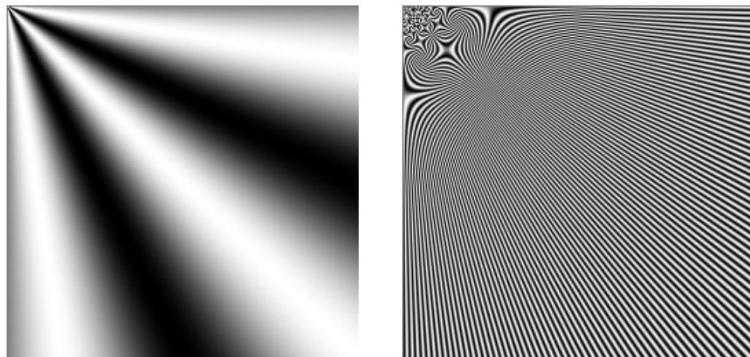


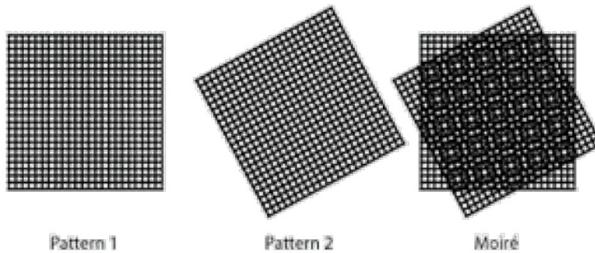
Figure 2.3: Esempio di frequenze indesiderate prodotte dall'aliasing.

L'aliasing deriva dal sottocampionamento e causa perdita di risoluzione dell'immagine campionata (effetto scacchiera).



Figure 2.4: Esempio di effetto a scacchiera generato dall'aliasing

- Per prevenire aliasing di queste componenti, è possibile filtrarle via (eliminarle) prima di campionare il segnale. Eliminare certe frequenze e lasciare passare le basse frequenze, è una operazione nota come **filtraggio passa-basso**.
- Ogni attenuazione relativa a questo processo di filtraggio rappresenta una perdita di risoluzione dell'immagine campionata.
- Come risultato, mentre da un lato c'è una perdita della risoluzione dell'immagine campionata, dall'altro c'è una attenuazione dell'aliasing error.
- **Effetto Moirè:** ovvero la distorsione visiva che si manifesta quando due griglie si sovrappongono



2.10 La risoluzione

Il campionamento e la quantizzazione determinano la **risoluzione** dell'immagine.

La **risoluzione** di un segnale è un indice del grado di qualità dell'immagine: misura il grado di oggetti distinguibili nell'immagine. Esistono differenti definizioni di risoluzione:

Definizione 13. *La **Risoluzione Spaziale** indica la densità dei campioni, ovvero è data dal numero di campioni per unità di area.*

Spesso è espressa come numero di pixel nell'unità di lunghezza e viene misurata in pixel per pollice (ppi).

Un'immagine ad alta risoluzione contiene più pixel di una delle stesse dimensioni con una risoluzione inferiore, quindi è in grado di riprodurre un maggior numero di dettagli. Un'elevata risoluzione comporta tuttavia un aumento considerevole delle dimensioni (quantità di dati) dell'immagine.

Esempio: Un'immagine di $1\text{cm} \times 1\text{cm}$ con una risoluzione di 72 ppi contiene 5184 pixel (72×72). La stessa immagine di $1\text{cm} \times 1\text{cm}$ a 300 ppi conterrebbe 90.000 pixel.

Definizione 14. *La **Risoluzione spettrale** indica la banda passante del sensore.*

Definizione 15. ***Risoluzione radiometrica** indica il numero di livelli di quantizzazione.*

Definizione 16. ***Risoluzione temporale** indica la frequenza di acquisizione dei frames di un'immagine in movimento.*

2.11 Alterazioni della risoluzione

Alterando i vari tipi di risoluzione, l'immagine presenterà di volta in volta un diverso tipo di distorsione.

Risoluzione spaziale: diminuendo la risoluzione spaziale (nell'esempio di un quarto) si ottiene il tipico effetto "quadrettato", detto anche a scacchiera, dovuto all'aliasing.



Risoluzione spettrale: Diminuendo la banda passante del sensore di acquisizione dell'immagine si ottiene un'immagine più "sfocata", in quanto i dettagli ad alta frequenza spaziale vanno persi.



Risoluzione radiometrica: Diminuendo la profondità di colore, si distinguono in maniera più marcata i passaggi da un colore ad un altro; essi risultano pertanto sempre più accentuati e meno graduati, fino a produrre dei "falsi contorni" (variazioni di ombreggiatura).



2.12 Immagini in bianco e nero e immagini a colori

2.12.1 Immagini in bianco e nero

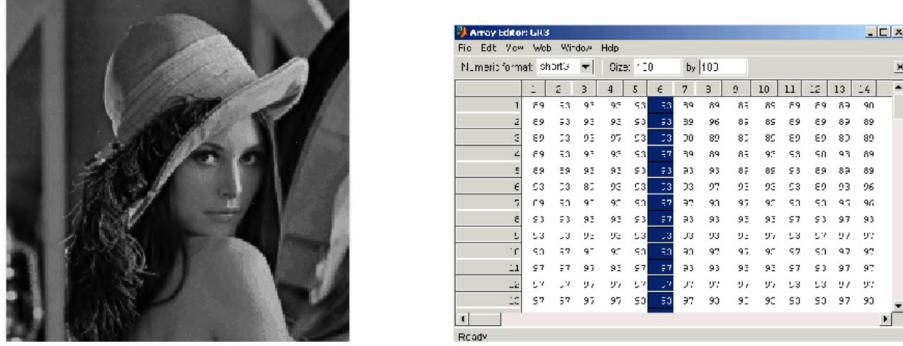
Un'immagine in bianco e nero (b/w) è caratterizzata da una rappresentazione binaria, ovvero la funzione che la rappresenta in ogni punto (x, y) può assumere solo due valori: 0 e 1. In genere, ad 1 si associa il bianco, mentre a 0 il nero.



Figure 2.5: Esempio di un'immagine in bianco e nero (binaria).

2.12.2 Immagini a toni di grigio

Un'immagine a toni di grigio è rappresentata da una matrice le cui entrate sono i valori che la funzione f assume in ogni punto.



In genere, si assume che i **livelli di grigio** siano discreti ed equispaziati in un intervallo di valori normalmente **tra 0 e 255**: esistono allora un massimo di 256 livelli di grigio.

La funzione $f(x, y)$ può essere rappresentata come una superficie nello spazio.



2.12.3 Immagini a colori

Per rappresentare un'**immagine a colori** è necessario ricorrere ad una funzione vettoriale. Un colore infatti può essere sempre decomposto come **somma dei tre colori fondamentali (rosso, verde, blu)**, ciascuno con un'opportuna intensità. Un'immagine a colori, dunque, può essere rappresentata da una funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ del tipo

$$f(x, y) = [R(x, y), G(x, y), B(x, y)]$$

Questo tipo di rappresentazione viene detta **RGB (Red,Green,Blue)**.

2.12.3.1 Lo spazio RGB

Lo spazio RGB è uno spazio cartesiano, con tre assi ortogonali. Il colore di ciascun pixel viene rappresentato da un vettore $[R(x, y), G(x, y), B(x, y)]$ nello spazio RGB; ogni componente indica la quantità di rosso, verde e blu,

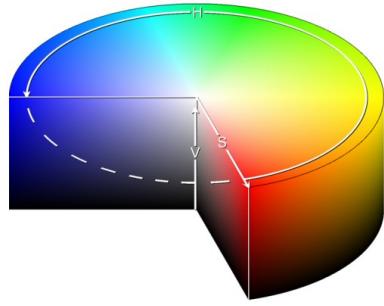


rispettivamente, necessari ad ottenere quel colore. In base alla rappresentazione RGB, un’immagine a colori viene rappresentata da una terna di matrici, ognuna delle quali contiene i valori relativi ad un canale di colore. Ogni canale, preso a sè, non è altro che un’immagine a toni di grigio.



2.12.3.2 Lo spazio HSV

Oltre alla RGB esistono anche altri tipi di rappresentazioni (che possono essere in genere derivate da essa). Una di queste è la rappresentazione **HSV**. Lo spazio HSV ha un sistema di coordinate cilindrico con due assi ortogonali ed un angolo di rotazione intorno ad uno dei due assi. L’altezza del cono rappresenta la **luminosità (Value)**, con valori da 0 (nero) a 1 (bianco). La **saturazione (Saturation)** indica l’intensità e la purezza del colore, con valori da 0 (sull’asse del cono) a 1 (sulla superficie del cono). La terza coordinata rappresenta la **tonalità di colore (Hue)** e viene misurata da un angolo intorno all’asse verticale (rosso a 0 gradi, verde a 120 e blu a 240).



2.13 Relazioni di base fra pixel

2.13.1 Intorni

Un pixel p di coordinate (x, y) ha quattro **neighbors** (vicini) orizzontali e verticali:

$$(x + 1, y), (x - 1, y), (x, y - 1), (x, y + 1)$$

Questi punti formano il **4-intorno (4-neighboor)** di (x, y) che indicheremo con $N_4(p)$

I quattro **neighbors (vicini) diagonali** di p ($N_D(p)$) hanno invece coordinate

$$(x + 1, y + 1), (x + 1, y - 1), (x - 1, y + 1), (x - 1, y - 1)$$

I vicini diagonali, insieme a quelli orizzontali e verticali, formano l'**8-intorno (8-neighbor)** di (x, y) che indicheremo con $N_8(p)$

2.13.2 Connattività e adiacenza

Due pixel sono connessi se sono vicini e presentano livelli di intensità con una certa relazione (ad esempio hanno lo stesso livello di grigio). Fissato un insieme V di valori di intensità, due pixel p e q si dicono

- **4-adiacenti** se i valori di entrambi appartengono a V e $q \in N_4(p)$
- **8-adiacenti** se i valori di entrambi appartengono a V e $q \in N_8(p)$
- **m-adiacenti** (adiacenza mista) se i valori di entrambi appartengono a V e $q \in N_4(p)$ oppure $q \in N_D(p)$ e l'insieme $N_4(p) \cap N_4(q)$ non contiene pixel a valori in V .

2.13.3 Cammini

Un **cammino (path) digitale** dal pixel $p = (x, y)$ al pixel $q = (x', y')$ è una sequenza di pixel

$$(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$$

dove $(x_0, y_0) = (x, y)$, $(x_n, y_n) = (x', y')$ e i pixel (x_i, y_i) , (x_{i+1}, y_{i+1}) sono adiacenti per ogni $i = 0, \dots, n - 1$, dove n è la **lunghezza del cammino**. Se $(x_0, y_0) = (x_n, y_n)$ si parla di **cammino chiuso**. Si puo' definire in particolare un 4-, 8-, o m - cammino restringendo l'adiacenza alla corrispondente tipologia.

2.13.4 Cammini e regioni

Fissato un sottoinsieme S di pixel di un'immagine digitale, due pixel p e q si dicono **connessi** se esiste un cammino tra p e q che consiste di pixel tutti contenuti in S .

Se tutti i pixel di S sono connessi, S si dice un insieme connesso: in tal caso diciamo che S è una regione dell'immagine. Il **bordo (boundary, border, contour)** di una regione R è l'insieme dei pixel di R che hanno uno o più vicini che non appartengono ad R . Nel caso in cui R sia l'intera immagine, il bordo si definisce come la prima e l'ultima riga e la prima e l'ultima colonna. Se l'immagine contiene k regioni distinte R_1, \dots, R_k , nessuna delle quali tocca i bordi dell'immagine, allora l'unione $R = \bigcup_{i=1}^n R_i$ si dice **primo piano (foreground)**, mentre il complementare R^c viene detto **sfondo (background)**.

2.13.5 Distanza tra pixel

Una distanza (o metrica) tra pixel è una funzione $D(p, q)$ tale che per ogni p, q, z

- $D(p, q) \geq 0$
- $D(p, q) = 0$ se e solo se $p = q$
- $D(p, q) = D(q, p)$
- $D(p, z) \leq D(p, q) + D(q, z)$

Esempi

- Distanza euclidea: $D_e(p, q) = \{(x_1 - x_2)^2 + (y_1 - y_2)^2\}^{\frac{1}{2}}$, se $p = (x_1, y_1), q = (x_2, y_2)$



- Distanza D_4 , o city-block: $D_4(p, q) = |x_1 - x_2| + |y_1 - y_2|$.
- Distanza D_8 , o a scacchiera: $D_s(p, q) = \max\{|x_1 - x_2|, |y_1 - y_2|\}$

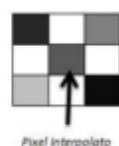
2.14 Zoom e Shrink

Questa parte dovrebbe far parte del Chapter Elaborazione Immagini ed essere messa con le elaborazioni locali

2.14.1 ZOOM

Si crea una nuova griglia, ovvero delle nuove "locazioni" per i pixel, sovrapponendola a quella originale; si assegnano poi i livelli di grigio a queste nuove locazioni.

- **Nearest Neighbor:** assegna ad ogni nuova locazione l'intensità del pixel più vicino dell'immagine originale
- **Interpolazione bilineare:** prevede l'utilizzo dei quattro pixel più vicini per stimare l'intensità da assegnare ad ogni nuova locazione



$$f(x, y) = ax + by + cxy + d$$

dove i coefficienti a, b, c, d sono determinati dal seguente sistema lineare di 4 equazioni in 4 incognite:

$$\begin{cases} f(x_0, y_0) = ax_0 + by_0 + cx_0y_0 + d \\ f(x_0, y_1) = ax_0 + by_1 + cx_0y_1 + d \\ f(x_1, y_0) = ax_1 + by_0 + cx_1y_0 + d \\ f(x_1, y_1) = ax_1 + by_1 + cx_1y_1 + d \end{cases}$$

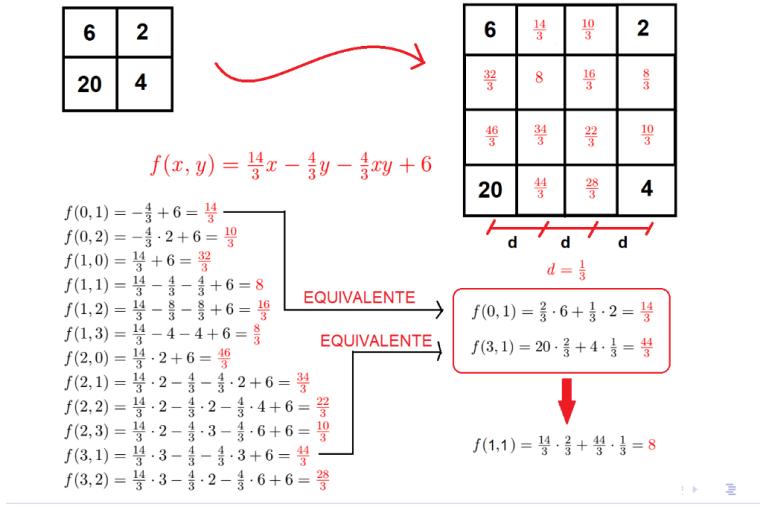
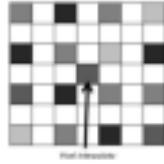


Figure 2.6: Esempio di svolgimento di uno zoom con interpolazione bilineare.

- **Interpolazione bicubica:** prevede l'utilizzo dei sedici pixel più vicini per stimare l'intensità da assegnare ad ogni nuova locazione.



2.14.2 SHRINK

Lo stesso procedimento dello zoom, con una griglia immaginaria di dimensioni inferiori all'originale. Per fattori interi si procede come nel caso dello zoom, ma per cancellazione di righe e colonne.

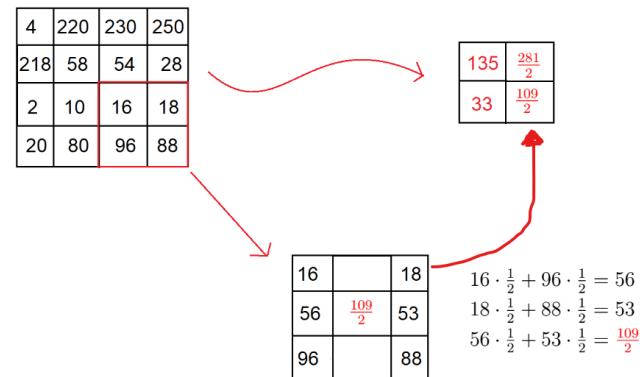


Figure 2.7: Esempio di svolgimento di uno shrink.

Chapter 3

Elaborazione con Filtri

L'elaborazione delle immagini è una disciplina che prevede l'utilizzo di algoritmi i quali operano sui pixel che compongono l'immagine e, applicando trasformazioni numeriche, restituiscono un'immagine modificata.

Le tecniche di elaborazione delle immagini hanno vari scopi, fra cui:

- il miglioramento della qualità dell'immagine (**image enhancement**)
- il ripristino della qualità dell'immagine (**image restoration**)
- l'estrazione di informazioni sul contenuto dell'immagine (**image analysis**)

Note:

- L'**image analysis** è una parte fondamentale della computer vision e precede l'**image recognition**.

Essa può richiedere elaborazioni differenti a seconda del tipo di informazione che si vuole estrarre: tra queste, le elaborazioni nel dominio spaziale, nel dominio delle frequenze, con riduzione dei dati tra ingresso e uscita (compressione), etc.

Esempio: Un'elaborazione nel dominio spaziale, può essere espressa come

$$g(x, y) = T(f(x, y))$$

- f è l'immagine di ingresso
- g è l'immagine di uscita
- T è un operatore su f , definito in un intorno di (x, y) .



La natura dell'intorno definisce il tipo di elaborazione e si distingue, in particolare, fra: **elaborazioni puntuali**, **locali** e **globali**.

Definizione 17. *Le elaborazioni **puntuali** trasformano il valore di un pixel sulla base del valore del pixel stesso.*

Definizione 18. *Le elaborazioni **locali** lavorano sulla base dei valori assunti dai pixel in un intorno di quello preso in esame.*

Definizione 19. *Le elaborazioni **globali** trasformano il valore di un pixel sulla base dei valori assunti da tutti i pixel dell'immagine.*

3.1 Elaborazioni Puntuale

Definizione 20. *Un'elaborazione puntuale si dice **omogenea** se il risultato dipende solo dal valore (in scala di grigi) del pixel a cui è applicata. Se invece il risultato dipende anche dalla posizione del pixel, l'elaborazione puntuale è **non omogenea**.*

Note:

- Elaborazioni puntuale sono anche dette **manipolazioni della scala dei grigi**

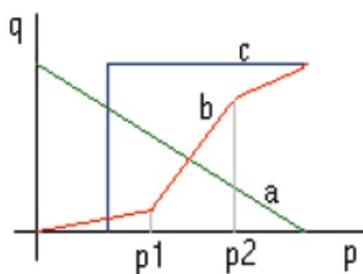
Un'elaborazione puntuale omogenea può essere rappresentata da una trasformazione

$$s = T(r)$$

dove:

- r è il livello di grigio dell'immagine di ingresso
- s è il livello di grigio dell'immagine in uscita

- In base al tipo di funzione T si ottiene un tipo diverso di trasformazione: a **gradino** (threshold), a **rampa**, **lineare** a **tratti**,...



3.1.1 Thresholding

Consideriamo una funzione T a gradino, ottenendo così una **elaborazione threshold (soglia)**.

Tale elaborazione fa sì che i valori dei pixel che non superano la soglia fissata vengano portati a 0, mentre i valori dei pixel che superano la soglia siano posti pari a 1.

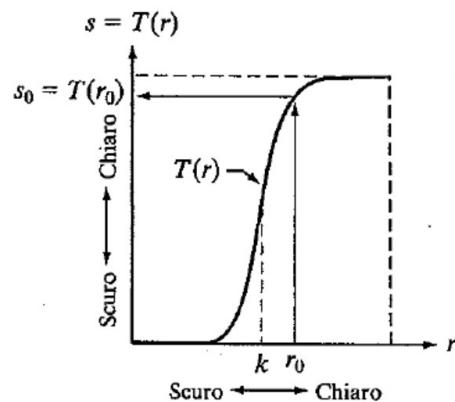
Si produce così un'immagine **binaria**.



Questo è un tipico esempio di **binarizzazione**. Si può ottenere una binarizzazione anche scegliendo una qualsiasi altra **funzione di discriminazione**, invece di una soglia costante.

3.1.2 Stiramento

Consideriamo una T del tipo:



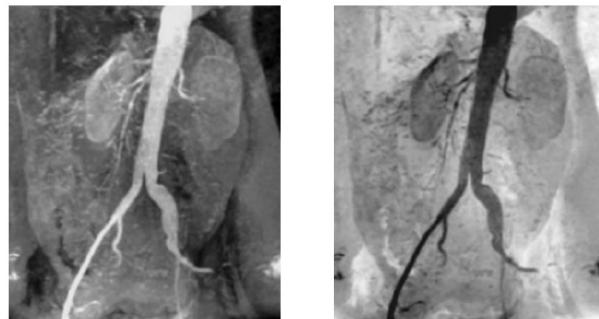
Vengono scuriti i livelli di grigio al di sotto di k e schiariti quelli al di sopra di k . Si ottiene così uno **stiramento dell'immagine (image stretching)**. L'elaborazione threshold può essere riguardata come caso limite di questo tipo di operazione.

3.1.3 Negazione

Consideriamo una T del tipo:

$$T(r) = L - 1 - r$$

se il range dinamico dell'immagine è $[0, L - 1]$ la scala di grigi viene invertita, ottenendo così una negazione.



3.1.4 Trasformazione Logaritmica

Consideriamo una T del tipo:

$$T(r) = c \log(1 + r), \quad c \in \mathbb{R}, r \geq 0$$

che prende il nome di **trasformazione logaritmica**: associa ad una stretta gamma di valori a bassa intensità dell'immagine originale una più ampia gamma nell'immagine in output.

Per livelli ad alta intensità, invece, si verifica il contrario.

Un tipico caso in cui è utile applicare questa trasformazione è per rappresentare la trasformata di Fourier, che spesso presenta una gamma molto ampia di intensità, difficilmente riproducibile senza perdere un significativo livello di dettaglio.

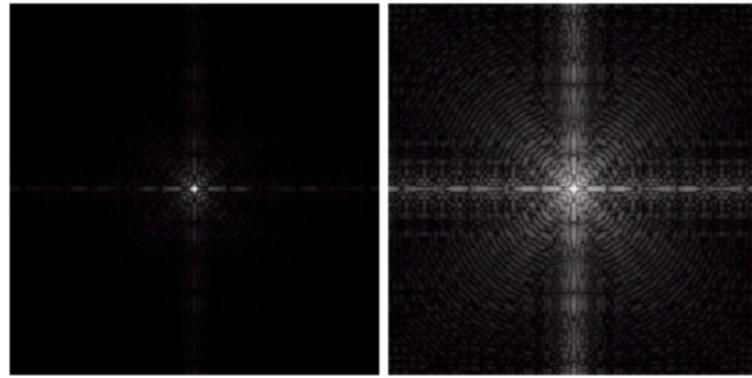


Figure 3.1: Nella figura sinistra viene mostrato lo spettro di Fourier con valori in $[0, 1.5 \times 10^6]$. Mentre in quella di destra il risultato dell'applicazione della trasformazione logaritmica con $c = 1$ (valori in $[0, 6.2]$).

3.1.5 Trasformazione di Potenza

Consideriamo una T del tipo:

$$T(r) = cr^\gamma, c, \gamma > 0,$$

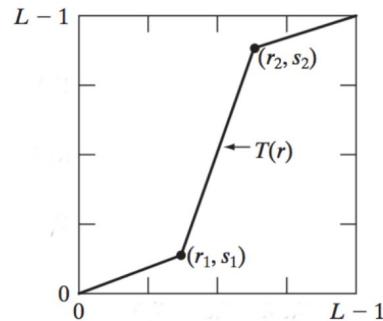
che prende il nome di **trasformazione di potenza (gamma)**. Se $\gamma < 1$ le curve potenza corrispondenti trasformano una stretta gamma di valori scuri in una gamma più ampia di valori in output, mentre se $\gamma > 1$, si verifica la trasformazione opposta.

La correzione tramite il fattore γ è importante per la corretta visualizzazione di immagini sullo schermo di un computer: immagini non corrette nel modo giusto possono apparire sbiadite o, al contrario, troppo scure.

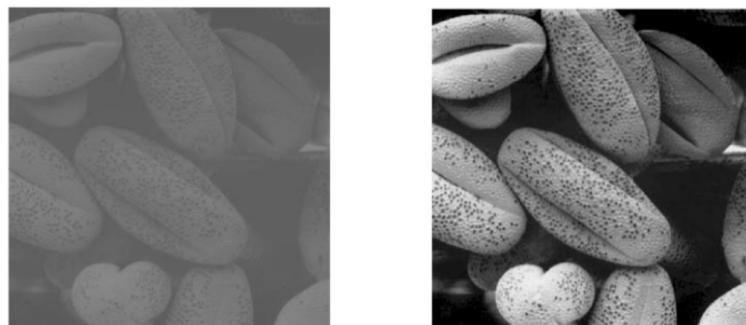


3.1.6 Lineare a Tratti

Considerando una **T lineare a tratti** del tipo si ottiene uno stretching del contrasto



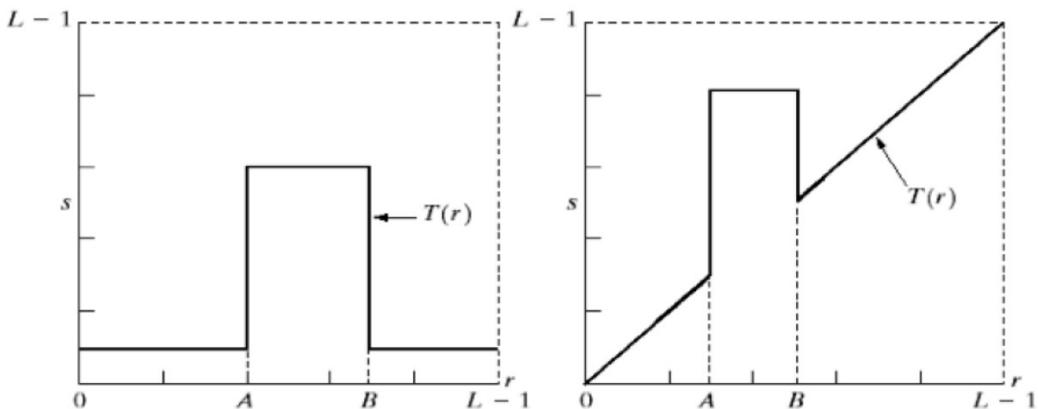
In questo caso invece si ha uno stretching del contrasto differente perché si sceglie come $(r_1, s_1) = (r_{min}, 0)$ e $(r_2, s_2) = (r_{max}, L - 1)$



Note:

- In questo caso r_{min} e r_{max} stanno ad indicare il più basso e alto valore nei livelli di grigio dell'immagine utilizzata.

3.1.7 Altre Trasformazioni a Tratti



In questo caso prendiamo queste due diverse trasformazioni ed analizziamo il loro effetto sull'immagine.

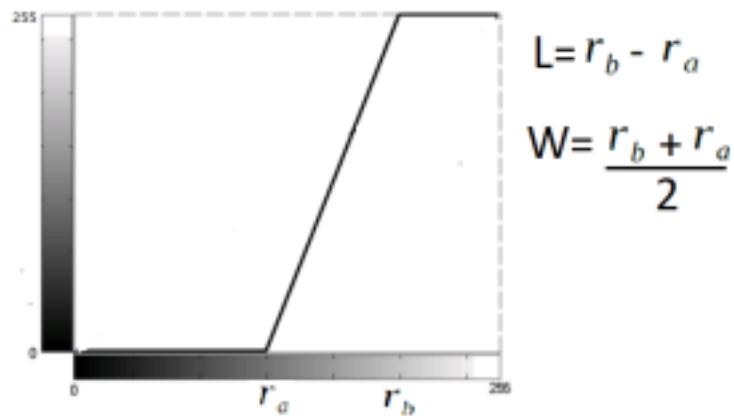
- La prima crea una binarizzazione ma in questo caso utilizza due differenti gradazioni di grigio e sono precisamente bianco e nero.
- La seconda mette in **risalto** una porzione della scala di grigi alzandogli il livello e schiarendoli.



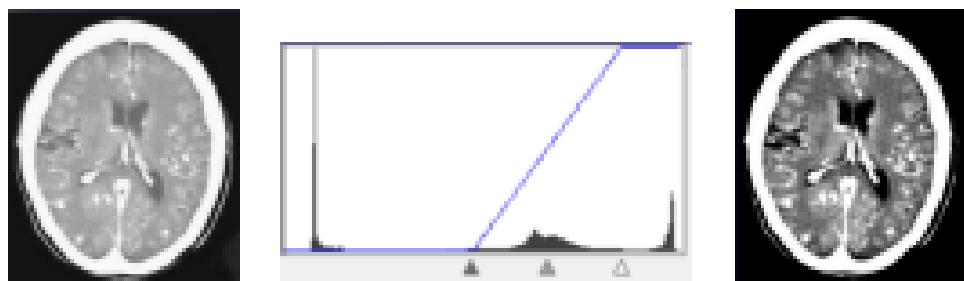
- Nella prima immagine vediamo una normale Angiografia aortica.
- Nella seconda abbiamo il risultato della selezione di intensità del primo tipo (banda di interesse $[A, B]$ selezionata sulla parte alta della scala di grigi).
- Nella terza abbiamo il risultato della selezione di intensità del secondo tipo (banda $[A, B]$ sulle tonalità medio-grigie impostata sul nero, così da preservare le tonalità di grigio dei vasi e dei reni).

3.1.8 Finestramento - Windowing

Definizione 21. Il *windowing* consiste nel mostrare solo una parte del range dei valori di grigio dell'immagine.



Applicazione: in molte immagini mediche il numero di valori della scala dei grigi utili dal punto di vista diagnostico è sensibilmente minore di tutti quelli disponibili. Usando sul monitor tutto il range possibile si diminuisce il contrasto visibile nella zona di interesse.



- Finestramento con $L = 169$ e $W = 97$
- Numero di grigio: 238 (Immagine di sx) e 97 (immagine di dx)
- Non viene aggiunta informazione, si aumenta solo il contrasto visibile

3.2 Modelli delle Immagini

In base al tipo di elaborazione che si vuole effettuare, può essere conveniente adottare diversi modelli per le immagini. Ad esempio:

- **Modello Deterministico**
- **Modello Probabilistico:**
 - sui **pixel** prevedendo che il loro valore sia considerato una variabile aleatoria,
 - sull'**immagine** riguardando cioè l'immagine stessa come un processo stocastico

3.2.1 Il modello Probabilistico

Nel modello probabilistico per i pixel, i valori assunti nei vari pixel ($N \times M$) di un'immagine vengono considerati come valori assunti da una variabile aleatoria in una successione di $N \times M$ esperimenti. Vado quindi a modificare i pixel in base ad un valore di probabilità.

È dunque possibile analizzare ed elaborare l'immagine utilizzando gli strumenti del calcolo delle probabilità e del calcolo stocastico. Un esempio di questo processo è **l'analisi dell'istogramma**.

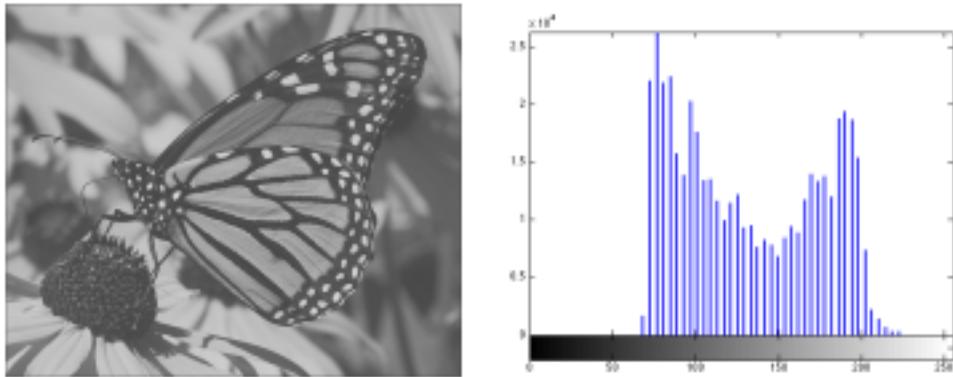
3.2.1.1 L'istogramma dei toni di grigio

Definizione 22. *L'istogramma dei toni di grigio si ottiene contando, per ogni valore del codominio dell'immagine (spazio di tutti i valori che possono essere assunti dai pixel), il numero di volte che tale valore compare nell'immagine.*

Il grafico che si ricava è un istogramma, cioè un grafico a barre dove l'asse delle ascisse è suddiviso in tanti punti quanti sono i possibili toni di grigio dell'immagine.

In termini probabilistici, l'istogramma rappresenta la distribuzione di probabilità della variabile aleatoria r ed indica il valore di grigio di un pixel, e quindi può essere visto come la **carta d'identità di un'immagine**. Saper leggere l'istogramma ci da un valore **qualitativo** dell'Immagine.

Esempio: il codominio di un'immagine a 256 toni di grigio sarà formato da tutti i numeri interi da 0 a 255 → l'asse orizzontale sarà suddiviso in 256 parti.



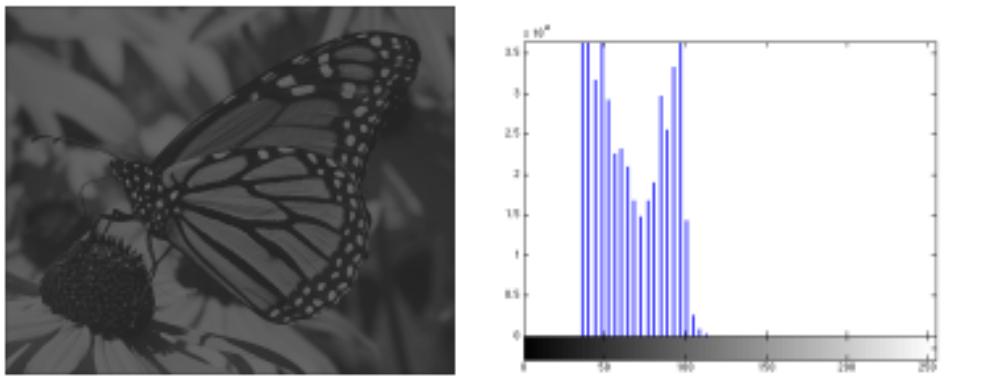
Definizione 23. In base alla definizione di distribuzione l’istogramma andrebbe normalizzato in modo da assumere valori tra 0 e 1. Se i toni di grigio sono r_k , $k = 0, \dots, 255$ allora l’altezza della barra in r_k è pari alla frequenza relativa di r_k , ovvero:

$$p(r_k) = \frac{n_k}{n}$$

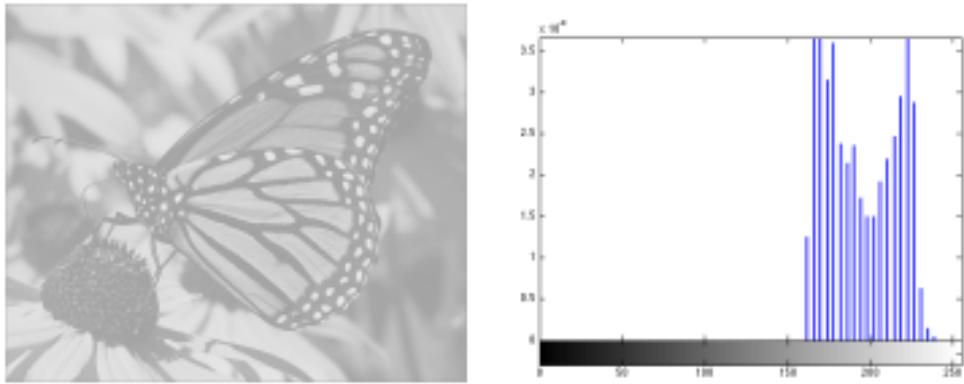
dove n_k è il numero di pixel in cui viene assunto r_k , n è il numero totale di pixel dell’immagine.

L’istogramma quindi fornisce una raffigurazione sintetica del contenuto cromatico o di luminosità dell’immagine, dunque una descrizione della qualità dell’immagine.

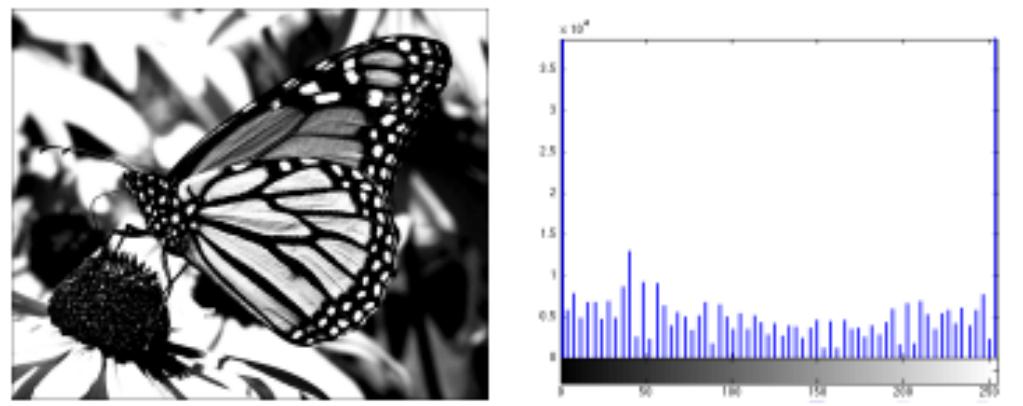
- **Immagine troppo scura** la distribuzione è concentrata su toni bassi di grigio



- **Immagine troppo chiara** la distribuzione è concentrata su toni alti di grigio



- **Immagine con alto contrasto** la distribuzione è concentrata su valori vicini a 0 e a 255. L'alto contrasto dell'immagine è quindi dovuto all'uniformità dell'istogramma.



3.2.1.2 Interventi sull'istogramma

Qualora le caratteristiche della distribuzione dei toni di grigio nell'immagine non siano ottimali, è consigliabile applicare all'istogramma opportune trasformazioni, basate su sistemi stocastici.

Tra queste:

Equalizzazione dell'istogramma

Definizione 24. *Ha lo scopo di uniformare l'istogramma dell'immagine lungo tutto il suo dominio.*

Il risultato è un nuovo istogramma in cui il numero di pixel ad ogni tono di grigio è il più possibile costante.



Algoritmo:

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

con $k = 0, \dots, L - 1$.

Il risultato è l'aumento del contrasto.

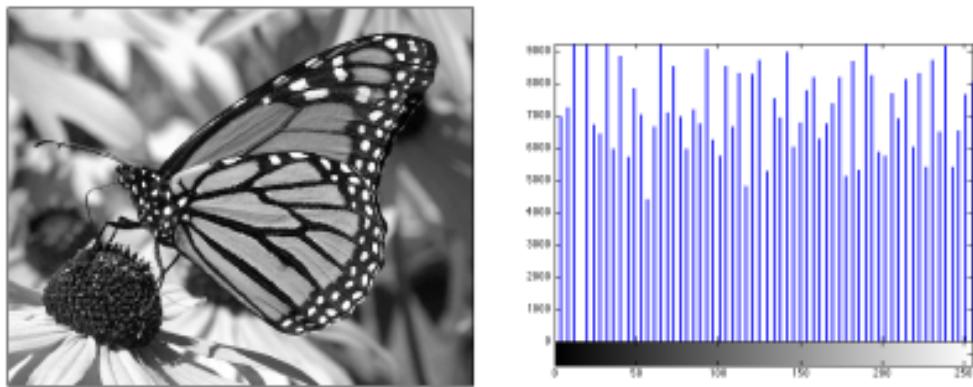


Figure 3.2: L'immagine mostra il risultato dell'equalizzazione dell'istogramma.

Shift dell'istogramma

Definizione 25. Lo Shift dell'istogramma consiste nel traslare i valori dell'istogramma.

Algoritmo:

$$\begin{aligned} T(r) &= \alpha r & 0 < \alpha < 1 \\ T(r) &= \alpha r + (L - 1)(1 - \alpha) & 0 < \alpha < 1 \end{aligned}$$

Andrò quindi a shiftare verso sinistra o verso destra l'istogramma schiarendo o scurendo la mia immagine.

Il **risultato** è un'immagine più scura o schiarita.





Figure 3.3: L'immagine mostra il risultato dello Shift dell'isogramma.

Stretching dell'istogramma

Definizione 26. Lo Stretching dell'istogramma consiste in uno stiramento dell'istogramma, in modo da distanziarne i picchi.

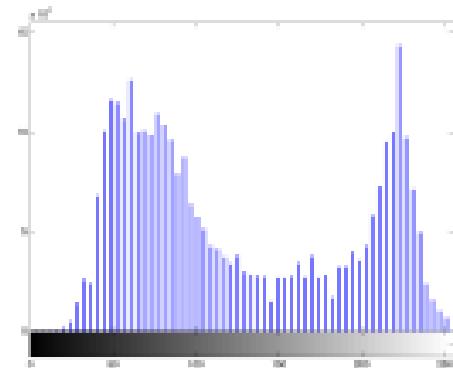
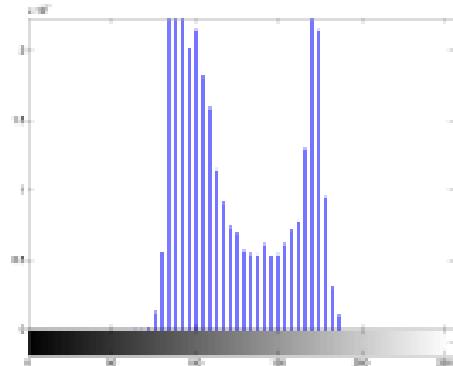
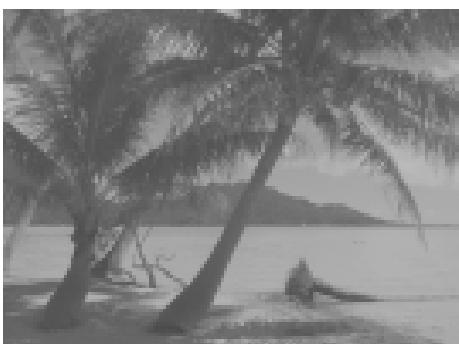
Si usa quando l'istogramma presenta dei picchi abbastanza ravvicinati, provocando un'immagine piuttosto uniforme.

Algoritmo:

$$T(r) = \begin{cases} 0, & 0 \leq r \leq r_{min} \\ (r - r_{min}) \frac{L-1}{r_{max} - r_{min}}, & r_{min} \leq r \leq r_{max} \\ L - 1, & r_{max} \leq r \leq L - 1 \end{cases}$$

dove $[r_{min}, r_{max}]$ è il range osservato nell'immagine originale, individuato dal minimo e dal massimo livello di grigio che presenta l'immagine.

Il risultato è l'aumento di contrasto



3.3 Il Rumore

Definizione 27. si intende l'insieme dei segnali indesiderati che si sovrappongono al segnale utile oggetto di studio (ad esempio un'immagine) causandone una degenerazione.

Si possono avere diverse forme di degrado tra cui:

- **Il rumore di quantizzazione**
- **Il rumore introdotto da condizioni esterne**
- **Il rumore introdotto dal sensore**
- **Il rumore introdotto dai dispositivi di amplificazione/condizionamento del segnale**

In base alle sue cause, il rumore si distingue tra:

- **Rumore indipendente dal segnale (additivo)**

- **Rumore dipendente dal segnale** (la relazione tra il segnale corrotto e quello originale è non lineare)

Rumore indipendente dal segnale: (caso più comune) la funzione che descrive il segnale corrotto è:

$$f(x, y) = g(x, y) + v(x, y)$$

dove $g(x, y)$ è il segnale e $v(x, y)$ è il rumore, che sarà di tipo **additivo**

Rumore dipendente dal segnale: l'intensità del rumore dipende dal segnale. Supponendo anche che esso sia molto più grande del segnale, la funzione è

$$f(x, y) = g(x, y) + v(x, y)g(x, y) = g(x, y)(1 + v(x, y)) \approx g(x, y)v(x, y)$$

dove $g(x, y)$ è il segnale e $v(x, y)$ è il rumore, che sarà di tipo **moltiplicativo**

Essendo di natura intrinsecamente stocastica, il rumore viene in genere analizzato usando la teoria dei processi stocastici ed è caratterizzato in base alla:

- **Distribuzione:** descrive la probabilità che il rumore assuma certi valori di intensità
- **Distribuzione spettrale:** ha a che fare con l'energia ad esso associata, al variare della frequenza

3.3.1 Il modello del rumore

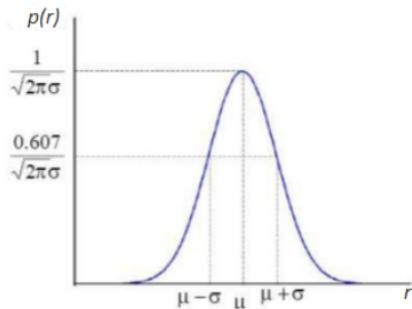
Per poter caratterizzare e studiare il rumore in un'immagine si fanno in genere delle ipotesi semplificative, in modo che il rumore abbia una qualche distribuzione di probabilità: si costruisce in questo modo un modello del rumore.

Una modellizzazione tipica prevede che:

- **lo spettro abbia distribuzione uniforme (rumore bianco)**

- il rumore abbia distribuzione gaussiana, cioè

$$p(r) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(r-\mu)^2}{2\sigma^2}}$$



dove μ è il valor medio e σ la deviazione standard.

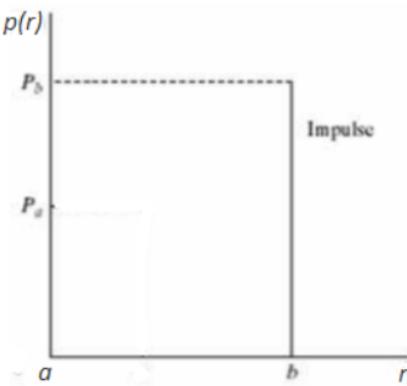


Figure 3.4: Immagine originale e immagine corrotta da rumore gaussiano

3.3.1.1 Il rumore Salt and Pepper

Nel caso in cui il rumore abbia distribuzione spaziale di tipo impulsivo, si parla di rumore salt and pepper (sale e pepe): agisce corrompendo in maniera casuale i pixel dell'immagine, portandone il valore a 0 (valore minimo) oppure a 255 (valore massimo).

$$p(r) = \begin{cases} p_a, & r = 0 \\ p_b, & r = 255 \\ 0, & \text{altrimenti} \end{cases}$$



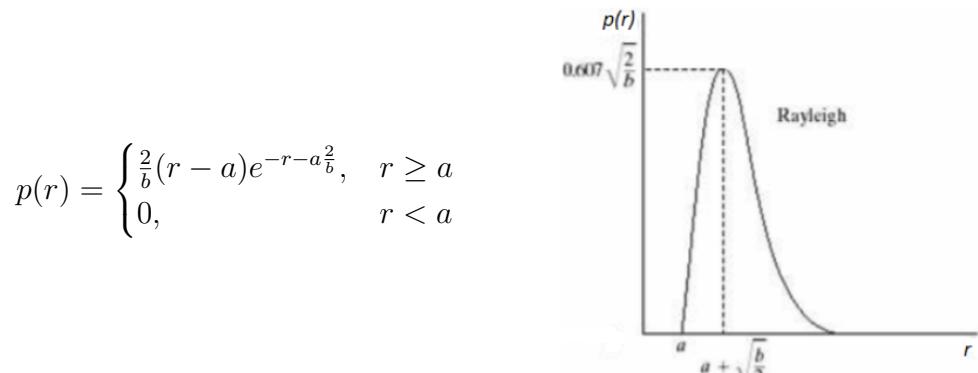
Ovviamente il rumore impulsivo, pur essendo additivo, **non è lineare**.



Figure 3.5: Immagine originale e immagine corrotta da rumore salt and pepper

3.3.1.2 Rumore di Rayleigh

Distribuzione del **rumore di Rayleigh**



dove il valor medio e la deviazione standard sono dati da:

$$\mu = a + \sqrt{\pi b / 4} \quad \sigma = \sqrt{\frac{b(4 - \pi)}{4}}$$

Nel grafico lo scostamento dall'origine e la forma inclinata verso destra rende questa densità utile per l'approssimazione di histogrammi non simmetrici e può essere utilizzata per rappresentare fenomeni di rumori tipici di alcuni sensori di range.

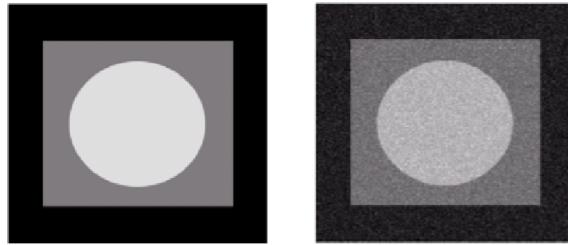
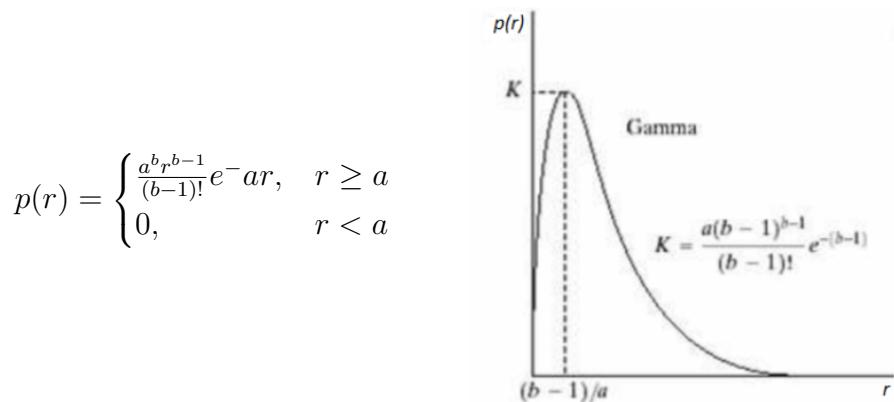


Figure 3.6: Immagine originale e immagine corrotta da rumore Rayleigh

3.3.1.3 Rumore Gamma

Distribuzione del **rumore gamma**:



dove $a > 0$, b è un intero positivo e il valor medio e la deviazione standard sono dati da:

$$\mu = \frac{b}{a} \quad \sigma = \frac{\sqrt{b}}{a}$$

Questo rumore è presente nelle immagini laser.

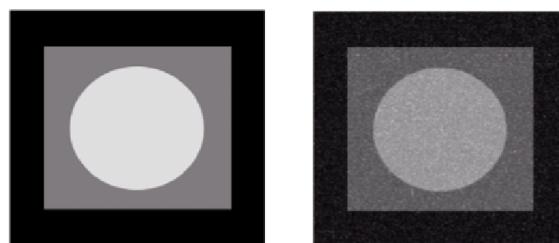
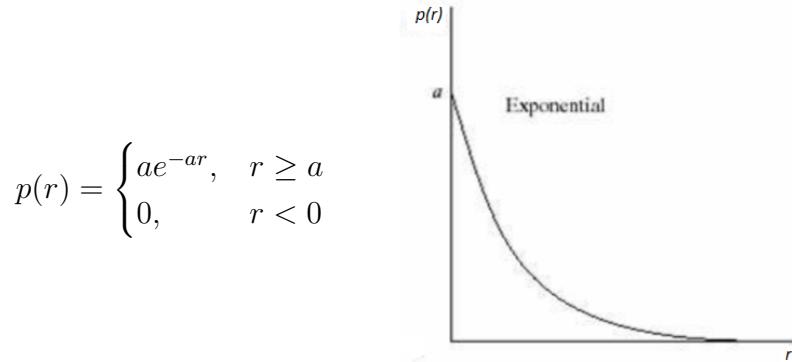


Figure 3.7: Immagine originale e immagine corrotta da rumore Gamma

3.3.1.4 Rumore Esponenziale

Distribuzione del rumore Esponenziale:



dove $a > 0$, b è il valor medio e la deviazione standard sono dati da:

$$\mu = \frac{1}{a} \quad \sigma = \frac{1}{a}$$

Questo rumore è un caso particolare del rumore di Gamma con $b = 1$.

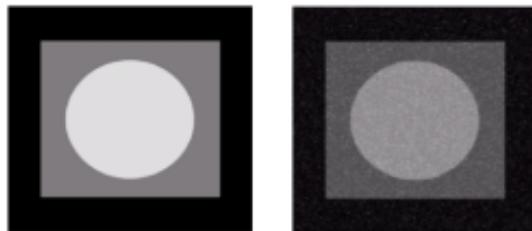


Figure 3.8: Immagine originale e immagine corrotta da rumore Esponenziale

3.3.2 Signal-Noize Ratio

Per avere una valutazione numerica dell'entità del rumore associato ad una data immagine si può ricorrere al **rappporto segnale-rumore (Signal Noise Ratio - SNR)**, definito come:

$$SNR = \frac{\sum_{(x,y)} f^2(x,y)}{\sum_{(x,y)} v^2(x,y)}$$

dove $f(x, y)$ è l'intensità del pixel (x, y) dell'immagine (già corrotta dal rumore), $v(x, y)$ è il rumore e le sommatorie sono al variare di tutti i pixel (x, y) dell'immagine.

3.4 Filtri

Definizione 28. Il rumore può essere corretto applicando opportune trasformazioni ai pixel dell'immagine, dette filtri.

Per ogni tipo di rumore esiste un filtraggio differente, che risulterà il più adatto ad attenuare o eliminare quel particolare rumore. In genere, la scelta del filtro dipende dalla linearità o meno della relazione fra l'immagine corrotta e quella originale. È possibile combinare più filtri, in modo da avere effetti più complessi.

3.4.1 Il filtraggio spaziale

Un filtro spaziale è caratterizzato da:

- Un intorno (**maschera**), in genere di dimensioni dispari;
- Un'operazione predefinita che viene applicata ai pixel nell'intorno

se l'operazione è lineare si parla di **filtro lineare**.

L'intensità dell'immagine filtrata nel pixel (x,y) sarà:

$$g(x, y) = \sum_{s=-a}^a \sum_{t=-b}^b w(s, t) f(x + s, y + t),$$

con $x = 0, \dots, M - 1$, $y = 0, \dots, N - 1$ (**correlazione di f e w**), dove i valori w sono i **coefficienti della maschera** avente dimensione $m \times n$, con $m = 2a + 1$ e $n = 2b + 1$

Ad esempio, se la maschera ha dimensione 3×3 , l'intensità dell'immagine filtrata nel pixel (x, y) sarà:

- $$g(x, y) = w(-1, -1)f(x-1, y-1) + w(-1, 0)f(x-1, y) + \dots + w(0, 0)f(x, y) + \dots + w(1, 1)f(x+1, y+1)$$

3.4.2 Filtri lineari

I filtri maggiormente utilizzati per la rimozione del rumore sono i cosiddetti **filtri di smoothing**, i quali eliminano picchi e increspature (passa-basso).



3.4.2.1 Filtro medio

Il filtro medio sostituisce il valore di ogni pixel prefissato con il valor medio dei pixel in un suo intorno di dimensioni fissate.

$$\frac{1}{mn} \times \begin{bmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{bmatrix}$$

Il **filtro medio** è molto utilizzato ad esempio per correggere il **rumore gaussiano**.

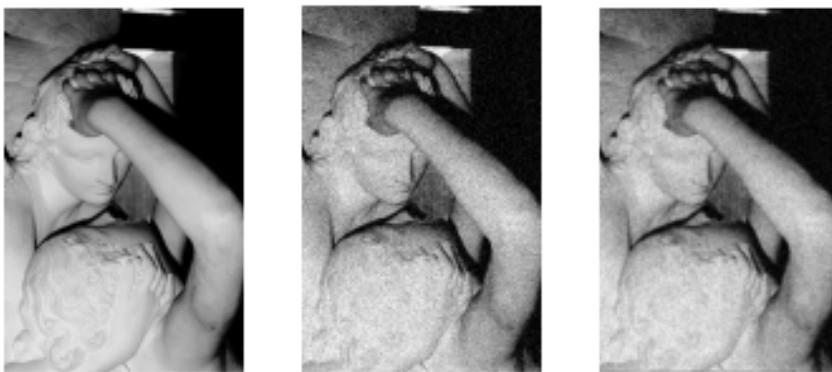


Figure 3.9: Immagine originale, immagine corrotta da rumore gaussiano e immagine filtrata con filtro medio

3.4.2.2 Filtro di media ponderata

Il filtro media ponderata sostituisce il valore di ogni pixel prefissato con la media ponderata dei pixel in un suo intorno di dimensioni fissate.

$$g(x, y) = \frac{\sum_{s=-a}^a \sum_{t=-b}^b w(s, t) f(x+s, y+t)}{\sum_{s=-a}^a \sum_{t=-b}^b w(s, t)}$$

3.4.2.3 Filtro Gaussiano

Il filtro gaussiano: sostituisce al valore di ogni pixel prefissato la media pesata dei valori dei pixel in un suo intorno. I pesi sono distribuiti secondo una funzione gaussiana.

$$\frac{1}{16} \times \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$



Figure 3.10: Possiamo vedere l'applicazione della sfocatura con filtro medio tramite maschere di dimensioni 3, 5, 9, 15, 35 su un'immagine di 500×500 pixel.

3.4.3 Filtri non lineari

I rumori non lineari (ad esempio quello impulsivo) non possono essere corretti con i filtri lineari. Possono essere trattati invece efficacemente con **filtri non lineari**

3.4.3.1 Filtro mediano

Il filtro Mediano sostituisce al valore di ogni pixel prefissato la mediana (valore centrale della lista ordinata) dei valori dei pixel nell'intorno fissato.

Il filtro mediano risulta particolarmente adatto a correggere il rumore **salt and pepper**, a differenza del filtro medio che, come si vede dall'esempio seguente, non risulta invece particolarmente efficace.



Figure 3.11: Immagine originale, immagine corrotta da rumore **salt and pepper** e immagine filtrata con **filtro mediano**



Figure 3.12: Immagine originale, immagine corrotta da rumore **salt and pepper** e immagine filtrata con **filtro medio**

3.4.3.2 Filtro di massimo

Sia S_p una maschera di dimensioni $m \times n$, dove $p = (x, y)$ è il pixel centrale della maschera. Il filtro di massimo sostituisce a p il massimo dei valori dei pixel in S_p , cioè

$$g_{max}(x, y) = \max_{(x,y) \in S_p} f(x, y)$$

Questo filtro è utile per **diminuire il rumore di tipo "pepper"**.

3.4.3.3 Filtro di minimo

Il **filtro di minimo** sostituisce a p il minimo dei valori dei pixel in S_p cioè:

$$g_{\min}(x, y) = \min_{(x,y) \in S_p} f(x, y)$$

Questo filtro è utile per diminuire il rumore di tipo "salt"

3.4.3.4 Filtro di Punto Medio

Il filtro di punto medio sostituisce a p il punto medio tra il massimo ed il minimo dei valori dei pixel in S_p , cioè

$$g(x, y) = \frac{1}{2} \left[\max_{(x,y) \in S_p} f(x, y) + \min_{(x,y) \in S_p} f(x, y) \right]$$

3.4.3.5 Filtro medio alpha-trimmed

Fissato un valore $0 \leq d \leq mn - 1$, supponiamo di cancellare in S_p i $d/2$ valori più chiari e i $d/2$ valori più scuri; delle restanti intensità ne calcoliamo la media aritmetica, cioè:

$$g(x, y) = \frac{1}{mn - d} \sum_{(x,y) \in \bar{S}_p} f(x, y)$$

dove \bar{S}_p è l'insieme dei pixel S_p rimanenti. Questo filtro risulta utile in situazioni in cui si sovrappongono diversi tipi di rumore, ad esempio nel caso di una combinazione tra rumore salt an pepper e rumore gaussiano.

3.5 Filtri di Sharpening

L'obiettivo di questi filtri è quello di **marcare i bordi** (**edge detection**), attribuendo meno importanza alle aree che hanno variazione lenta a livello di intensità. Si basano sull'operatore di derivazione ed hanno l'effetto di evidenziare soltanto i bordi dell'immagine. In forma discreta, se f rappresenta l'intensità associata a ogni pixel (x, y) , le derivate possono essere espresse come differenze finite:

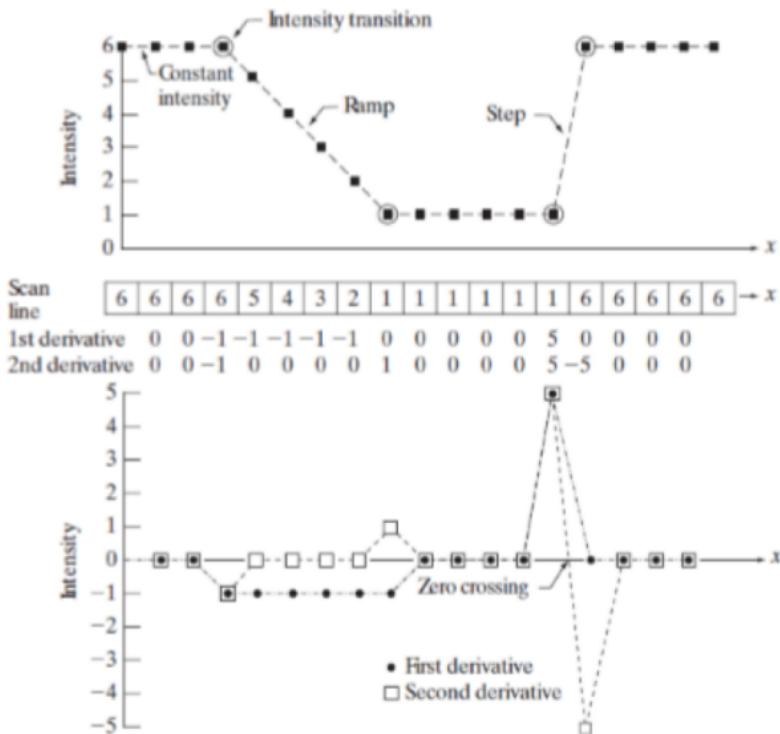
$$\begin{aligned} \frac{\partial f}{\partial x}(x, y) &\sim f(x+1, y) - f(x, y), \quad \frac{\partial f}{\partial y}(x, y) \sim f(x, y+1) - f(x, y) \\ \frac{\partial^2 f}{\partial x^2}(x, y) &\sim f(x+1, y) + f(x-1, y) - 2f(x, y) \\ \frac{\partial^2 f}{\partial y^2}(x, y) &\sim f(x, y+1) + f(x, y-1) - 2f(x, y) \end{aligned}$$

Il filtro di sharpening ha l'effetto di evidenziare soltanto i bordi dell'immagine.



Figure 3.13: Applicazione di un filtro di sharpening

Da ricontrolare e aggiungere spiegazioni



3.5.1 Filtro del Gradiente

Operatore Gradiente:

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right),$$

Una comune implementazione della derivata prima nell'ambito dell'image processing è costituita dal modulo del gradiente

$$|\nabla f(x, y)| \sim \left| \frac{\partial f}{\partial x} \right| + \left| \frac{\partial f}{\partial y} \right|$$

Maschere:

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



Figure 3.14: Applicazione del Filtro del Gradiente

3.5.2 Filtro di Sobel

Un'altra approssimazione discreta del gradiente è data da:

$$\begin{aligned} \nabla f(x, y) &\sim \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} \sim \\ &\sim |[f(x+1, y-1) + 2f(x+1, y) + f(x+1, y+1)] + \\ &\quad - |[f(x-1, y-1) + 2f(x-1, y) + f(x-1, y+1)]| + \\ &\quad + |[f(x-1, y+1) + 2f(x, y+1) + f(x+1, y+1)] + \\ &\quad - |[f(x-1, y-1) + 2f(x, y-1) + f(x+1, y-1)]| \end{aligned}$$

Maschere:

$$\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

che sono detti **operatori di Sobel**.

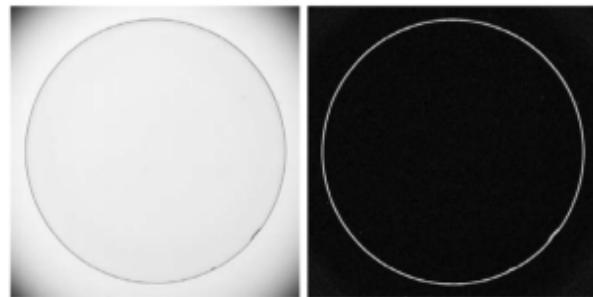


Figure 3.15: Immagine originale e immagine filtrata tramite il gradiente di Sobel.

3.5.3 Filtro di Prewitt

Analogo a quello di Sobel eccetto per le maschere di convoluzione che in questo caso sono:

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

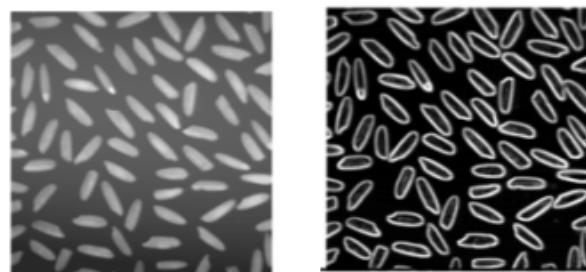


Figure 3.16: Immagine originale e immagine filtrata tramite il gradiente di Prewitt.

3.5.4 Filtro di Roberts

Si ottiene mediante la seguente approssimazione del gradiente:

$$|\nabla f(x, y)| \sim \left| \frac{\partial f}{\partial x} \right| + \left| \frac{\partial f}{\partial y} \right| \sim \\ \sim |f(x, y) - f(x+1, y+1)| + |f(x+1, y) - f(x, y+1)|$$

Calcola il gradiente lungo le direzioni diagonali.

Maschere:

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

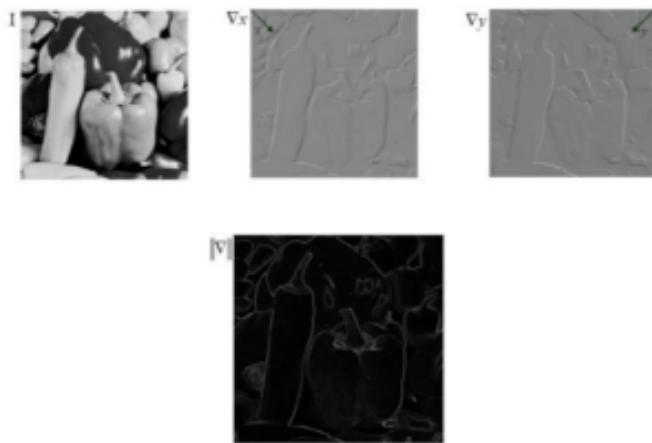


Figure 3.17: Nella figura possiamo osservare l'applicazione delle singole matrici ∇_x , ∇_y e poi l'applicazione di tutte e due in contemporanea (filtro di Roberts).

3.5.5 Filtro Laplaciano

Operatore Laplaciano:

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \sim \\ \sim f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1) - 4f(x, y)$$

Maschera:

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

oppure, tenendo conto delle direzioni diagonali:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

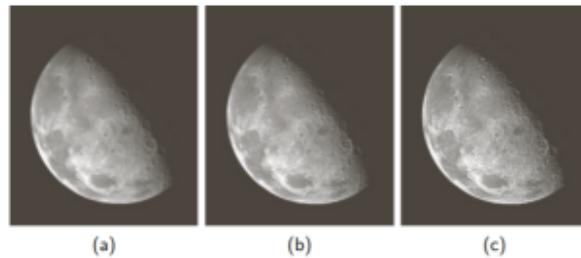


Figure 3.18: (a) Immagine originale, (b) con laplaciano, (c) con laplaciano con diagonali.

3.5.6 Filtro Differenziale

Sfrutta la proprietà di attraversamento dello zero della derivata seconda nelle varie direzioni, coordinate e diagonali, che ha come effetto quello di evidenziare direzionalmente i bordi di un'immagine lungo di esse.

Dalle approssimazioni:

$$\frac{\partial^2 f}{\partial x^2}(x, y) \sim f(x+1, y) + f(x-1, y) - 2f(x, y),$$

$$\frac{\partial^2 f}{\partial y^2}(x, y) \sim f(x, y+1) + f(x, y-1) - 2f(x, y),$$

$$\frac{\partial^2 f}{\partial x, y}(x, y) \sim \frac{1}{4}[f(x-1, y-1) - f(x-1, y+1) - f(x+1, y-1) + f(x+1, y+1)]$$

Si ottengono le maschere:

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 1 & -2 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{4} & 0 & -\frac{1}{4} \\ 0 & 0 & 0 \\ -\frac{1}{4} & 0 & \frac{1}{4} \end{bmatrix}$$

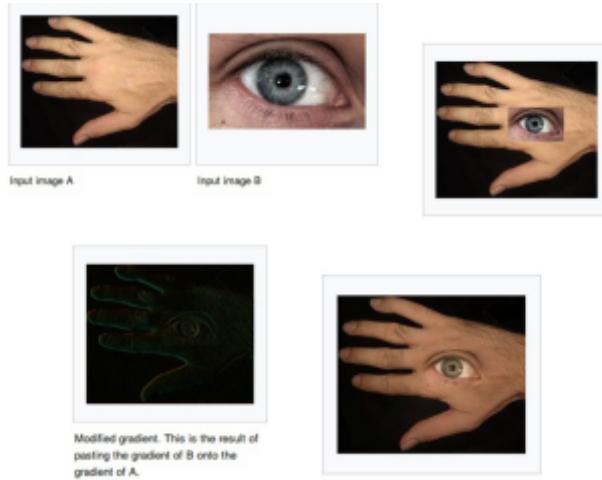


Figure 3.19: Possiamo utilizzare il filtro differenziale per effettuare il Digital Compositing.

3.6 Operazioni aritmetiche e Image Enhancement

Date due immagini digitali $f(x, y)$ e $g(x, y)$ della stessa dimensione, si possono definire le seguenti operazioni:

- **Sommatoria di immagini:** date due immagini è possibile definire la loro somma come:

$$s(x, y) = f(x, y) + g(x, y)$$

Questa operazione è utile per effettuare l'enhancement con un'operazione di sommatoria.

Siano $g_i(x, y)$ K immagini corrotte da rumore additivo η_i (eta), ovvero

$$g_i(x, y) = f(x, y) + \eta_i(x, y), \quad i = \dots K,$$

dove $f(x, y)$ è l'immagine priva di rumore.

Se $\bar{g}(x, y)$ è la loro media, ovvero

$$\bar{g}(x, y) = \frac{1}{K} \sum_{i=1}^K g_i(x, y)$$

allora sotto opportune ipotesi, risultata:

$$E[\bar{g}(x, y)] = f(x, y) \quad \text{e} \quad \sigma_{\bar{g}(x, y)}^2 = \frac{1}{K} \sigma_{\eta(x, y)}^2$$

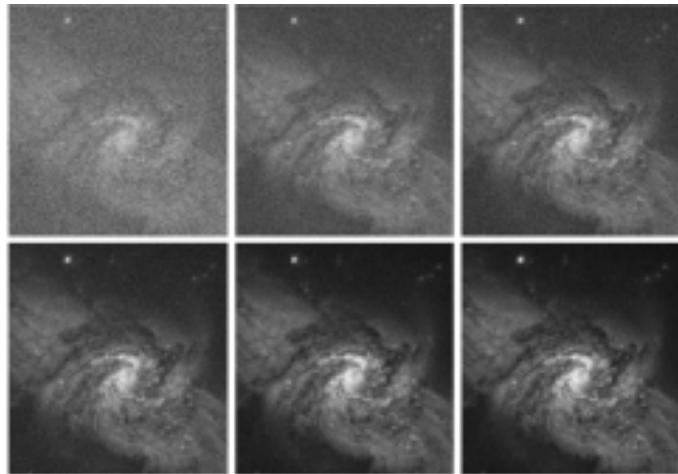
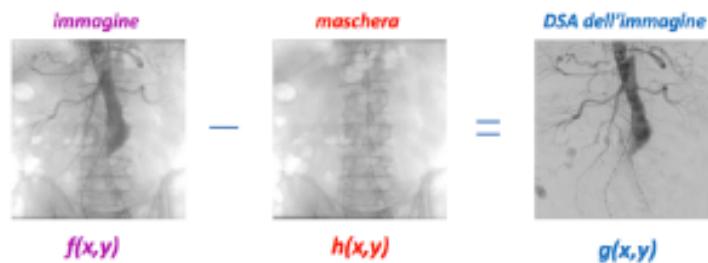


Figure 3.20: a) Immagine originale affetta da rumore gaussiano b)-f) Risultato della media di 5, 10, 20, 50 e 100 immagini rumorose.

- **Sottrazione di immagini**

$$g(x, y) = f(x, y) - h(x, y)$$

Si utilizza, ad esempio, quando si vogliono evidenziare le differenze fra due immagini, dove $h(x, y)$, detta maschera, è un'immagine della regione di interesse. Un esempio è la **Digital Subtraction Angiography (DSA)**.



- Moltiplicazione e divisione di immagini

Ricontrollare ed ampliare!

$$g(x, y) = f(x, y) \cdot h(x, y)$$

Esempi:

- Correzione dell'ombreggiatura(shading)
- Selezione di una ROI (Region of Interest)



Chapter 4

Image Segmentation

La segmentazione di un'immagine nell'elaborazione digitale è un processo di partizione dell'immagine in regioni significative.

Definizione 29. *E' il processo con il quale si classificano i pixel dell'immagine che hanno caratteristiche comuni: pertanto ciascun pixel in una regione è simile agli altri della stessa regione per una qualche proprietà o caratteristica (colore, intensità o texture).*

Matematicamente, il processo di segmentazione di un'immagine $f(x, y)$ in regioni R_1, \dots, R_n deve soddisfare le seguenti condizioni:

- $\cup_{i=1}^n R_i = f(x, y)$
- Ogni regione R_i con $i = 1, \dots, n$ è connessa;
- $R_i \cap R_j = \emptyset$ con $i \neq j$;
- Se $P(\cdot)$ è un predicato che indica la conformità di tutti i pixel di una regione R_i ad un particolare modello della regione stessa, allora devono valere:
 - $P(R_i) = \text{vero } \forall i = 1 \dots n$
 - $P(R_i \cup R_j) = \text{falso } \forall R_i, R_j$ regioni adiacenti

La segmentazione dell'immagine ha molte applicazioni:

Riorganizzare la visualizzazione delle immagini!

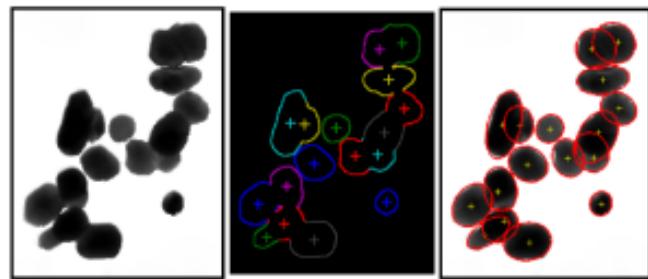


Figure 4.1: Biologia

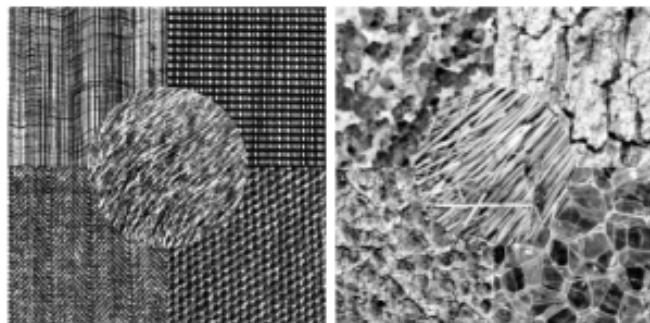


Figure 4.2: Texture

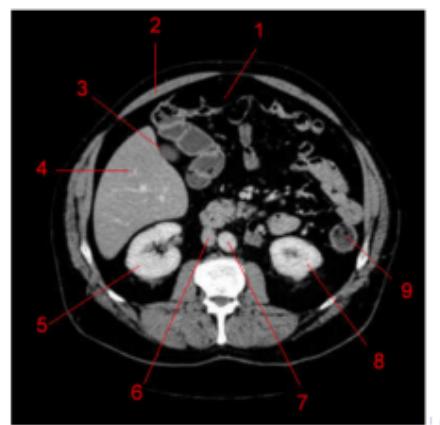


Figure 4.3: Medicina

Le fondamentali metodologie di segmentazione sono:

- **Thresholding Segmentation:** segmentazione mediante sogliatura;

- **Edge Segmentation:** edge based segmentation - localizzazione dei bordi;
- **Region Based Segmentation:** operatori morfologici;
- **Clustering Based Segmentation:** raggruppamento di oggetti sulla base della loro distanza reciproca;
- **Matching Techniques:** tecniche basate sulla conoscenza a priori.

4.1 Thresholding Based Segmentation

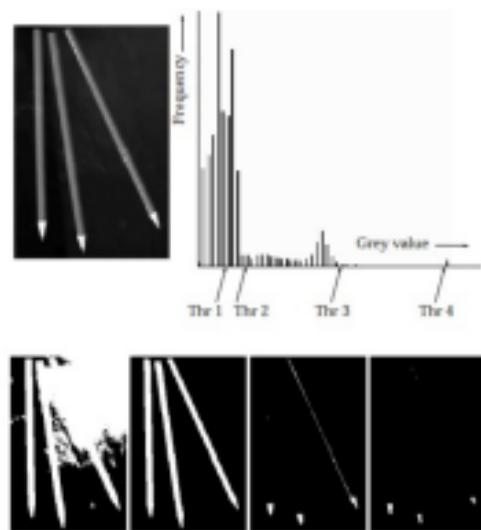
Singoli pixel dell'immagine sono catalogati come "pixel oggetto" se il loro valore è maggiore di una certa soglia e come "pixel di sfondo" se il valore è sotto la soglia. L'immagine binaria in uscita ha valore pari a 1 per ogni pixel dell'oggetto e pari a 0 per lo sfondo.

Problema: individuare il valore ottimo che minimizza l'errore dovuto all'over-segmentation (troppi oggetti) e all'undersegmentation (pochi oggetti).

Soluzione: sogliatura locale → scelta di un valore di soglia diverso, adatto ad ogni regione esaminata.

4.1.1 Thresholding manuale

- Procedura supervisionata → dipendente dall'operatore
- Procedura NON automatica



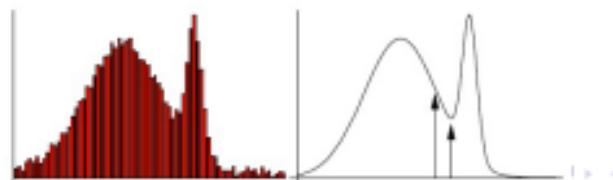
4.1.2 Thresholding automatico basilare

Separazione dei picchi negli istogrammi che rappresentano oggetti diversi nell'immagine.



Figura: Immagine sintetica vs immagine reale

Problema: Cosa accade se le code si sovrappongono?



4.1.3 Metodo di Otsu (1979)

Idea: trovare la soglia ottimale massimizzando la **varianza interclasse** o, equivalentemente, minimizzando la **varianza intraclasse**.

Vantaggi:

- procedura non parametrica;
- senza supervisione;
- automatica;
- facile da implementare e costo computazionale basso nel caso di immagine con istogramma bimodale.

Limiti:

- è un metodo di sogliatura globale → non tiene conto delle piccole variazioni di livelli di grigio;
- si basa solamente sull'istogramma dell'immagine (e quindi sul livello di grigio dei pixel) e prevede che sia bimodale.

4.1.3.1 Descrizione

Si ipotizzi che un'immagine digitale di $M \times N$ pixel, abbia L differenti livelli di grigio e si indichi con η_i il numero di pixel di intensità i . L'istogramma normalizzato ha componenti $p_i = \eta_i / (M \times N)$ da cui si ha:

$$\sum_{i=0}^{L-1} p_i = 1, \quad p_i \geq 0$$

Ora si ipotizzi di selezionare una soglia $T(k) = k$, $0 < k < L - 1$ e di utilizzarla come soglia sull'immagine di input ottenendo due classi C_0 e C_1 . La probabilità $\omega_0(k)$ che un pixel venga assegnato alla classe C_0 è data dalla probabilità cumulativa

$$\omega_0(k) = \sum_{i=0}^k p_i,$$

Analogamente per la classe C_1 si ha:

$$\omega_1(k) = \sum_{i=k+1}^{L-1} p_i = 1 - \omega_0(k)$$

Il valore medio (valore atteso) di intensità dei pixel appartenenti a C_0 è

$$\mu_0(k) = \frac{1}{\omega_0(k)} \sum_{i=0}^k i p_i$$

Analogamente, il valore medio di intensità della seconda classe è

$$\mu_1(k) = \frac{1}{\omega_1(k)} \sum_{i=k+1}^{L-1} i p_i$$

La media e la probabilità cumulativa fino al livello k sono date da:

$$\mu(k) = \sum_{i=0}^k i p_i, \quad \omega(k) = \sum_{i=0}^k p_i$$

mentre la media globale, cioè l'intensità media dell'intera immagine, è

$$\mu_T = \mu(L) = \sum_{i=0}^{L-1} i p_i$$

devono valere le seguenti relazioni:

- $\omega_0 \mu_0 + \omega_1 \mu_1 = \mu_T$,



- $\omega_0 + \omega_1 = 1$

La varianza delle due classi è data da:

$$\sigma_0^2 = \frac{1}{\omega_0} \sum_{i=0}^k (i - \mu_0)^2 p_i, \quad \sigma_1^2 = \frac{1}{\omega_1} \sum_{i=0}^k (i - \mu_1)^2 p_i$$

La **varianza interclasse** è definita come:

$$\sigma_W^2 = \omega_0 \sigma_0^2 + \omega_1 \sigma_1^2$$

ovvero la somma pesata delle varianze delle due classi.

Mentre la **varianza interclasse** è data da

$$\sigma_B^2 = \omega_0(\mu_0 - \mu_T)^2 + \omega_1(\mu_1 - \mu_T)^2 = \omega_0 \omega_1 (\mu_1 - \mu_0)^2 = \frac{[\mu_T \omega_k - \mu(k)]^2}{\omega(k)(1 - \omega(k))}$$

Da notare che σ_B^2 cresce all'aumentare della distanza tra i due valori medi μ_0 e μ_1 , mettendo in evidenza che la varianza interclasse è la misura della separabilità tra le classi.

La soglia ottimale è data dal valore k^* che massimizza $\sigma_B^2(k)$, ovvero

$$\sigma_B^2(k) = \max_{0 \leq k \leq L-1} \sigma_B^2(k)$$

Una volta che è stato trovato k^* l'immagine viene segmentata nel modo seguente:

$$g(x, y) = \begin{cases} 1 & \text{se } f(x, y) > k^* \\ 0 & \text{se } f(x, y) \leq k^* \end{cases}$$

L'algoritmo può essere riassunto come segue:

1. calcolare l'istogramma normalizzato dell'immagine denotando le componenti p_i , $i = 0, 1, 2, \dots, L - 1$;
2. calcolare la probabilità cumulativa $\omega(k)$;
3. calcolare la media cumulativa $\mu(k)$;
4. calcolare l'intensità media globale μ_T ;
5. calcolare la varianza interclasse σ_B^2 ;



6. calcolare la soglia di Otsu k^* , cioè il valore di k che rende massima la varianza interclasse;

Esiste un'estensione del metodo di Otsu che prende il nome di metodo di **Phansalkar**

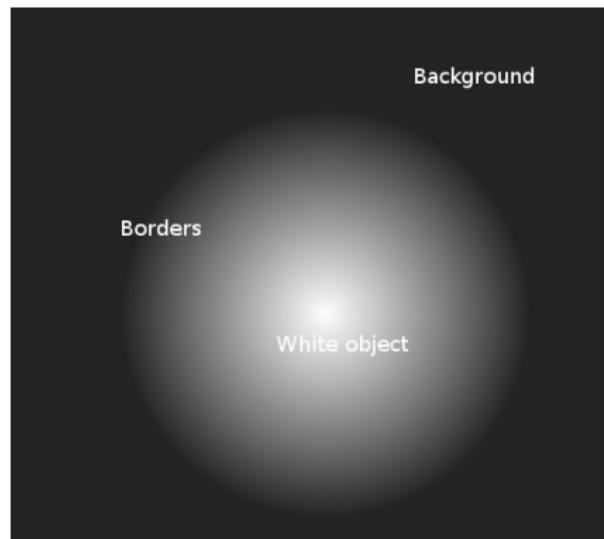


Figure 4.4: Immagine originale e immagine sogliata con l'algoritmo Otsu

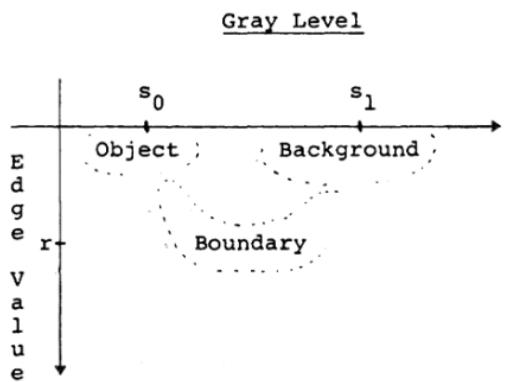
4.1.4 Edge-Gray Levels Histogram method (1977)

Da integrare!

Testato per la prima volta su immagini termografiche (FLIR Forward-Looking Infra Red). Sfrutta un vettore multidimensionale di criteri piuttosto che un solo parametro caratterizzante come il livello di grigio.



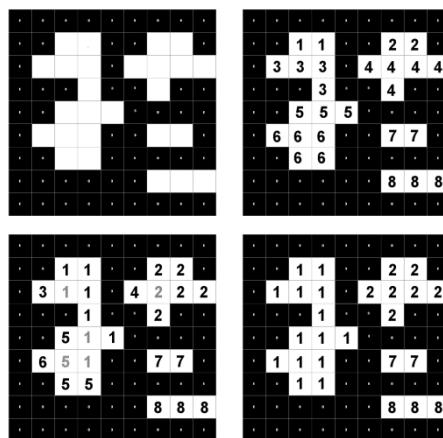
È un metodo di segmentazione delle immagini in uno spazio bidimensionale di parametri, in cui i criteri considerati per ciascun pixel sono da una parte il livello di grigio (grey level), dall'altra il valore ai bordi (edge value), rappresentativo nel mondo digitale del modulo approssimato del valore puntuale del gradiente.



4.2 Clustering Based Segmentation

4.2.1 Fuzzy Selection

La Fuzzy selection seleziona tutti i pixel di un oggetto fra loro adiacenti e che abbiano un valore compreso in un certo intervallo di livelli di grigio. Si tratta, anche in questo caso, di una selezione multiparametrica. L'algoritmo che la implementa dovrà attraversare l'immagine etichettando i vertici in base alla connettività ed ai valori relativi dei vicini. Al termine della scansione tutti i polimini, fra loro disgiunti, saranno etichettati con label differenti.



Note:

- **Polimini:** è definito come *un insieme di quadrati che abbiano a due a due un lato in comune*. Volgarmente solo le forme del tetris.

4.3 Region Based Segmentation

4.3.1 Operatori morfologici

La parola morfologia denota comunemente lo studio della forma. Gli operatori morfologici matematici effettuano elaborazioni dipendentemente dalla forma di un oggetto, estraendo dall'immagine componenti utili alla rappresentazione e descrizione della forma di una regione (contorni, scheletro, ecc...). La struttura dell'immagine viene "sondata" con un insieme di forma definibile dall'utente (elemento strutturante). Ora analizzeremo i principali operatori morfologici.

4.3.1.1 Erosione

Sia I l'immagine da analizzare ed SE un elemento strutturante.

L'erosione tra I e SE è definita come:

$$I \ominus SE = \{z \mid (SE)_z \subset I\} \quad (4.1)$$

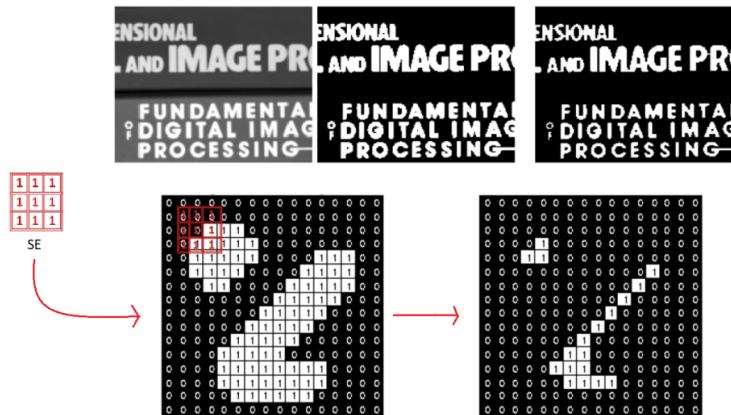
dove $(SE)_z$ indica l'elemento strutturante centrato nel punto z . Quando vado a fare l'erosione quindi vado ad **eliminare z**.

Equivalentemente si usa

$$I \ominus SE = \{z \mid (SE)_z \cap I^c = \emptyset\}$$

Quinti l'erosione fa sì che il pixel, in cui SE è centrato, diventi 0 se nell'intorno di quel pixel c'è almeno uno zero;





Nel caso in cui stiamo lavorando con un’immagine a scala di grigi l’erosione prende il minimo valore nell’elemento strutturante:

$$I \ominus SE = \min\{I(z) \mid z \in SE\}$$

Note:

- Nella (4.1) si intende che l’intero kernel deve appartenere all’immagine, quindi se ho un kernel di tutti 1 e lo centro in un punto con vicinato di tutti 0, questo non appartiene all’immagine I .

4.3.1.2 Dilatazione

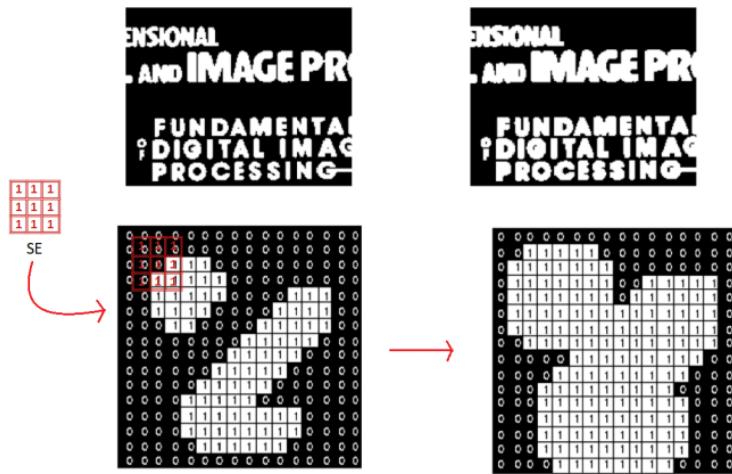
La dilatazione, detta anche addizione di Minkowsky, tra I ed SE è definita come:

$$I \oplus SE = \{z \mid (SE)_z \cap I \neq \emptyset\}$$

La formulazione appena riportata è equivalente a:

$$I \oplus SE = \{(z \mid (SE)_z \cup I) \subset I\}$$

Quindi la dilatazione fa sì che il pixel, in cui SE è centrato, diventi 1 se nell’intorno di quel pixel c’è almeno un 1.



Nel caso in cui stiamo lavorando con un'immagine a scala di grigi la dilatazione prende il massimo valore nell'elemento strutturante:

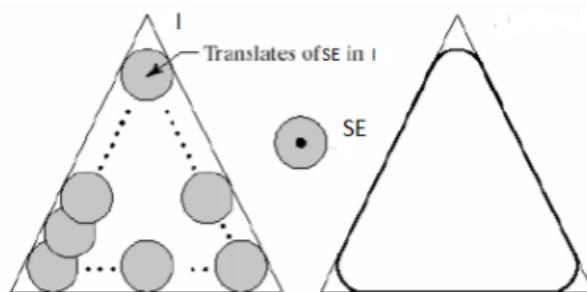
$$I \oplus SE = \max\{I(z) \mid z \in SE\}$$

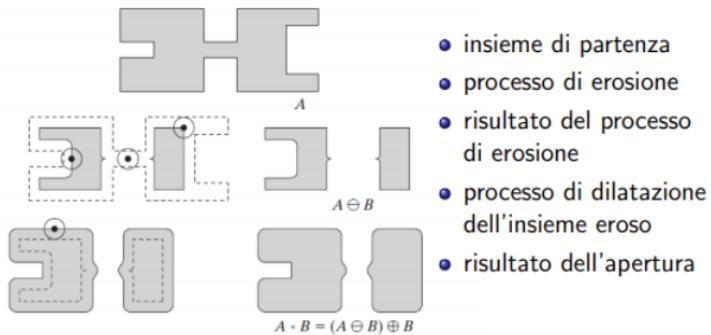
4.3.1.3 Apertura

L'apertura è una erosione seguita da una dilatazione, dove si considera sempre lo stesso elemento strutturante e non è commutativa. È definita come:

$$I \circ SE = (I \ominus SE) \oplus SE$$

L'effetto dell'apertura è di preservare il più possibile regioni di forma simile all'elemento strutturante, e di eliminare quelle differenti.



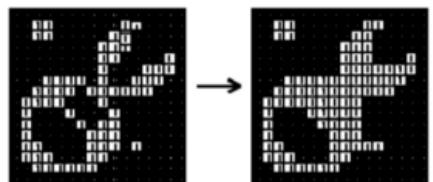
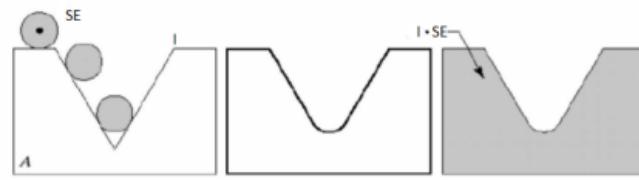


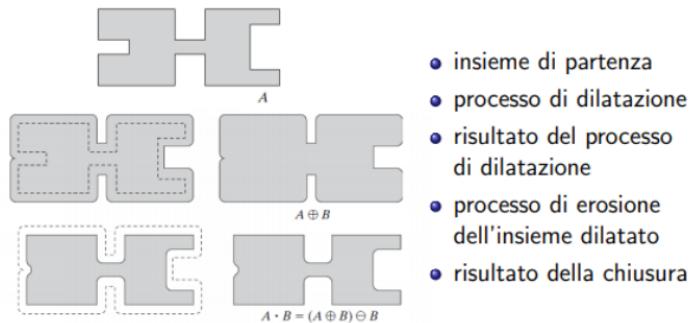
4.3.1.4 Chiusura

La chiusura è una dilatazione seguita da una erosione, dove si considera sempre lo stesso elemento strutturante e non è commutativa. È definita come:

$$I \bullet SE = (I \oplus SE) \ominus SE$$

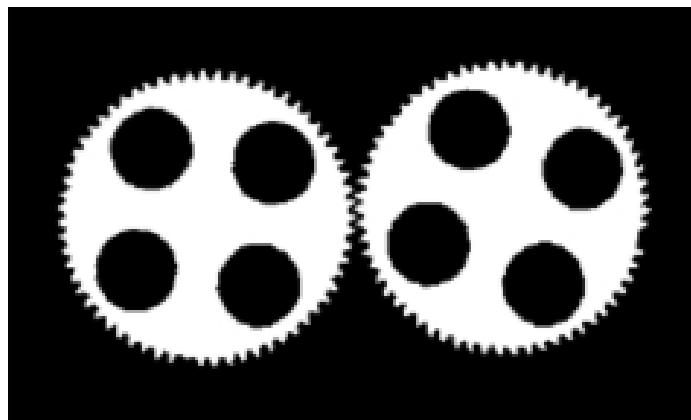
L'effetto della chiusura è di chiudere gli eventuali buchi interni tenendo conto della forma dell'elemento strutturante.





4.3.1.5 Applicazioni

E' importante osservare come la convoluzione dell'elemento strutturante modifichi l'immagine di partenza solo alla fine dell'analisi di tutti i punti del dominio. Un esempio di applicazione degli operatori morfologici è il problema posto da Sternberg nel 1985, al fine di controllare, a partire da una foto binaria, l'integrità dei denti degli ingranaggi prodotti in un'azienda di orologi.



Come primo passo è necessario riempire i "buchi" degli ingranaggi nell'immagine originale. Si nega l'immagine di partenza B ottenendo B^c e si effettua un'erosione con un SE grande come i buchi centrali degli ingranaggi, ottenendo l'immagine (b).



Figure 4.5: L'immagine a sinistra (a) è l'immagine originale B . Quella a destra (b) è ottenuta come $B1 = B^c \ominus \text{hole_ring}$

Si prende un SE di dimensioni maggiori rispetto ai buchi degli ingranaggi di partenza e si opera una dilatazione: il risultato è mostrato in (c). Quindi si opera un'operazione logica di OR fra B e $B2$ al fine di ottenere $B3$, figura (d)



L'immagine a sinistra (c) è ottenuta con $B2 = B1 \oplus \text{hole_mask}$. L'immagine a destra (d) è ottenuta con $B3 = B \text{ OR } B2$

A questo punto si può considerare un SE che abbia una dimensione tale da includere soltanto i denti degli ingranaggi al suo interno, $B7$ di figura (e). Tramite un'operazione di AND è facile ottenere un'immagine binaria contenente solo i denti degli ingranaggi da analizzare (f).

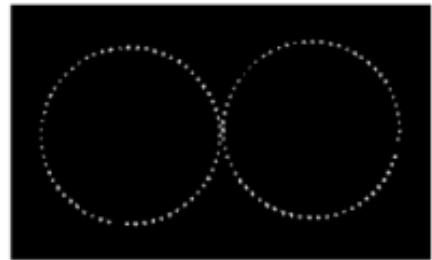
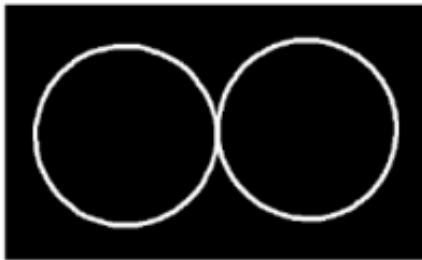
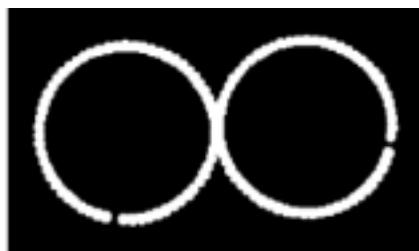


Figure 4.6: A sinistra l'immagine (e) $B7$ ottenuta dall'applicazione di un kernel che evidenzia solo i denti degli ingranaggi. A destra (f) l'immagine ottenuta con $Bs = B \text{ AND } B7$

Se si sceglie un SE circolare con diametro pari alla metà della distanza fra due denti successivi e se si applica una dilatazione, si otterranno “buchi” in corrispondenza delle strutture dentate uscenti come in figura (g). Infine, un altro elemento strutturante di diametro superiore a quello della distanza interdentale viene utilizzato per evidenziare i denti mancanti (h).

Riscrivere caption!



g) $B9 = B8 \oplus \text{Up_spacing}$

h) $\text{RESULT} = ((B8 * B9) \oplus \text{defect_cue}) OR B9$

Chapter 5

Indici di Similarità

5.1 Mean Square Error (MSE)

Al fine di valutare matematicamente la differenza in termini di intensità fra due immagini A e B, viene introdotto il concetto di **Errore Quadratico Medio (MSE)**.

$$MSE_{AB} = \sum_{i=1}^N \sum_{j=1}^M \frac{|I_A(i, j) - I_B(i, j)|^2}{NM}$$

dove I_A e I_B sono le intensità dei livelli di grigio delle due immagini di dimensioni $M \times N$ di cui si calcola la differenza. Più il valore di questo indice è basso, minore sarà la differenza tra le immagini sia numericamente in termini di bit, sia in termini di qualità visiva.

5.2 Peak Signal to Noise Ratio (PSNR)

Per quantificare l'entità del suddetto errore secondo un termine di paragone, viene introdotto un ulteriore indice di qualità delle immagini. Si definisce come **Peak Signal to Noise Ratio (PSNR)**, il rapporto tra la massima potenza ammisibile di un segnale e l'MSE.

$$PSNR_{AB} = 10 \log\left(\frac{255^2}{MSE_{AB}}\right)$$

Può anche essere espresso come:

Ricontrollare se la formula seguente è vera !

$$PSNR_{AB} = 20 \log\left(\frac{255}{\sqrt{MSE}}\right)$$

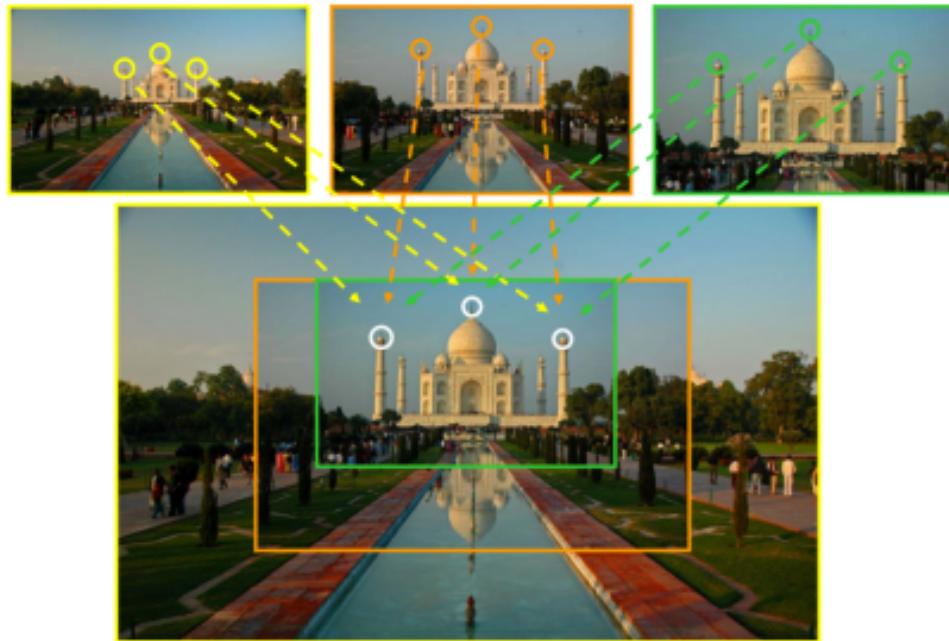


Chapter 6

Image Registration

Definizione 30. La registratura d'immagini (*image registration*) è quel processo che permette la trasformazione di differenti insiemi di dati, presenti in diversi insiemi di coordinate, in un sistema dove ogni coordinata spaziale corrisponde.

La registratura è necessaria per poter confrontare o integrare i dati ottenuti da diverse misure. Si prende una delle immagini come sorgente (source) e ci si riferisce alla seconda immagine come bersaglio (target).

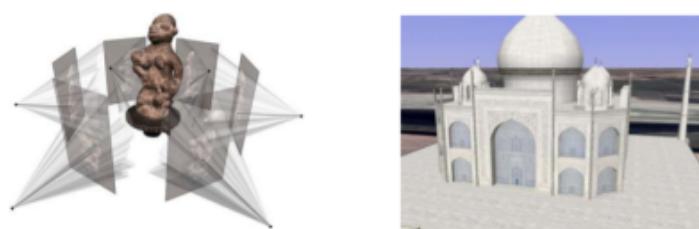


Alcune possibili applicazioni di questa tecnica sono:

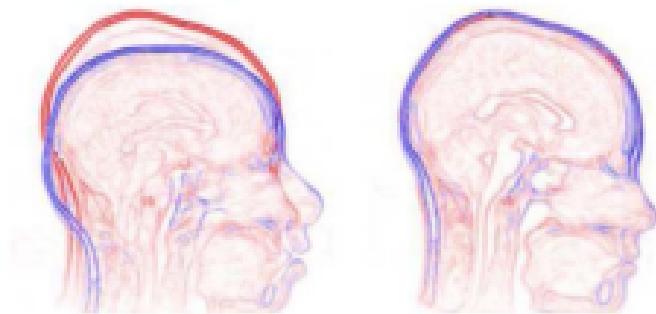
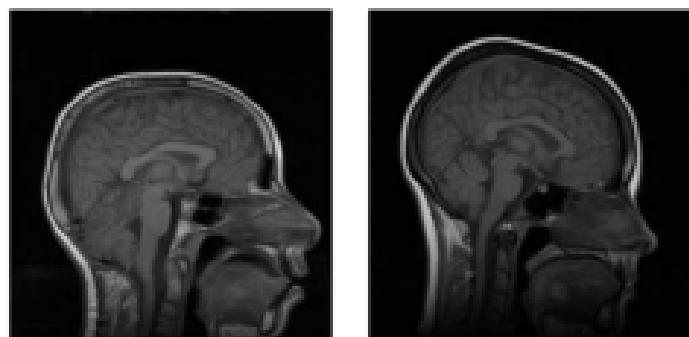
- Rilevamento del movimento con telecamera non stazionaria



- Ricostruzione 3D



- Medical imaging



Una classificazione dei metodi di registratura si basa su nove criteri:

1. *Dimensione del dominio dell'immagine*: si possono considerare:

(a) solamente le dimensioni spaziali:

- 2D/2D
- 2D/3D
- 3D/3D

(b) immagini acquisite in tempi differenti, con dimensioni spaziali:

- 2D/2D
- 2D/3D
- 3D/3D

2. *Natura della registrazione*:

• **Estrinseca**: I metodi estrinseci si fondano su oggetti artificiali che vengono attaccati sul paziente. Si dividono in:

- Invasivi
 - * telaio stereotassico (stereotactic frame)
 - * marcatori a vite (screw mounted markers)
- Non Invasivi
 - * stampi, adattatori dentali
 - * marcatori su pelle

• **Intrinseca** I metodi intrinseci si basano sul solo contenuto dell'immagine del paziente. Si dividono in:

- **Landmark based**: basata sui punti fiduciali: i punti di possono essere anatomici, cioè punti precisi e localizzabili della morfologia dell'anatomia visibile, solitamente identificati in modo interattivo dall'utente.
- **Segmentation based**: può essere **rigid model based**, dove si individuano le stesse strutture anatomiche (principalmente superfici) estratte da entrambe le immagini da registrare e utilizzate come unico input per la procedura di allineamento; oppure deformable model based, dove una struttura estratta (principalmente superfici e curve) da un'immagine è deformata elasticamente per adattarsi alla seconda immagine
- **Voxel property based**: si basano solamente sui livelli di grigio dell'immagine (metodo globale).



- **Non basata su immagini**

3. *Tipo di Trasformazione:*

- **Rigida:** comprende solo traslazioni e rotazioni
- **Affine:** mappa linee parallele in linee parallele
- **Proiettiva:** mappa linee in linee
- **Elastica:** mappa linee in curve

4. *Dominio della Trasformazione:* Una trasformazione è chiamata globale se è applicata all'intera immagine, locale se è applicata ad un sottinsieme dell'immagine.

5. *Interazione:* Fra gli algoritmi di registrazione, si possono individuare tre livelli di interazione:

- **Automatica:** l'utente fornisce all'algoritmo solo i dati dell'immagine ed eventualmente informazioni sull'acquisizione dell'immagine.
- **Interattiva:** l'utente effettua personalmente la registrazione, assistito da software.
- **Semi-autimatica:** l'utente deve inizializzare l'algoritmo, ad esempio, segmentando i dati o guidando l'algoritmo a rifiutare o accettare le ipotesi di registrazione suggerite.

6. *Procedura di ottimizzazione:*

- **Parameters computed:** i parametri che compongono la trasformazione della registrazione vengono calcolati direttamente, cioè determinato in maniera esplicita dai dati disponibili
- **Parameters searched for:** i parametri che compongono la trasformazione della registrazione vengono ricercati, cioè determinati trovando un massimo di qualche funzione definita nello spazio dei parametri.

7. *Modalità:*

- **Monomodale:** le immagini da registrare appartengono alla stessa modalità (radiografia, CT, MR, PET, SPECT, US, raggi X o DSA, ecc..)

- **Multimodale:** le immagini da registrare derivano da due diverse modalità (TC-MR, TC-PET, TC-SPECT, DSA-MR, PET-MR, US-TC, raggi X-MR, ecc..).

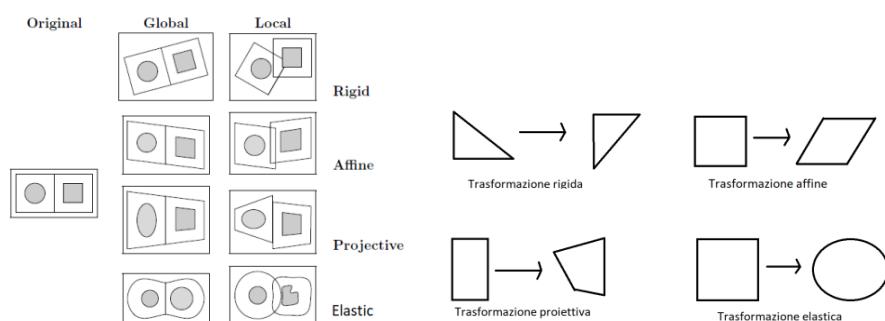
8. *Soggetto:*

- **Intrasubject:** tutte le immagini coinvolte nella registrazione sono acquisite da un singolo paziente;
- **Intersubject:** la registrazione viene effettuata utilizzando due immagini di diversi pazienti (o un paziente e un modello);
- **Atlas:** un'immagine viene acquisita da un singolo paziente e l'altra immagine è in qualche modo costruita da un database di informazioni su un'immagine ottenuta utilizzando l'imaging di molti soggetti.

9. *Oggetto:*

- **Testa:** cervello, occhio, denti.
- **Torace:** intero, cardiaco, seno.
- **Addome:** generale, rene, fegato.
- **Bacino e perineo**
- **Arti** generale, femorale, omero, mano
- **Colonna vertebrale e vertebre**

Allineare e sistemare le immagini !



Chapter 7

Approfondimenti

In questo capitolo sono contenuti alcuni approfondimenti che possono essere ignorati in quanto non necessari per sostenere l'esame.

Il successivo contenuto non è stato riletto e controllato !!

7.1 Compressione di immagini

La creazione di un'immagine digitale comporta la generazione di una enorme quantità di dati, tanto più per immagini ad alta risoluzione e con molte sfumature di colore

Questo può essere un inconveniente per varie operazioni quali:

- archiviazione dell'immagine
- processamento e trasferimento dell'immagine
- trasmissione in rete

Esempio

Per acquisire tramite uno scanner a 300 dots per inch (dpi) una pagina quantizzata con 2 livelli di grigio vengono generati più di 8.000.000 di bits (1.000.000 di bytes, ovvero 1 MB). Per rappresentare in formato digitale l'Encyclopædia Britannica sono necessari dunque oltre 25 GB. **NB:** le dimensioni di un Hard Disk si aggirano in media intorno agli 80 GB. Un'immagine di 512x480 pixel in **scala di grigi** occupa 512x480x1 bytes, ovvero 240 Kb, in quanto ogni pixel equivale ad un byte.





La stessa immagine, in **RGB**, occuperebbe 512x480x3 bytes, cioè 720 Kb (ogni pixel equivale a 3 bytes, uno per ogni canale).

Esempio - filmato

Passando ai filmati video, la situazione diventa ancora più complessa: un filmato a 25fps (frame per secondo) occupa 99Kb x 25, ovvero 2.4 Mb al secondo.

Un filmato in RGB addirittura 7.2 Mb al secondo.

7.1.1 Compressione

In quest'ottica diventa fondamentale ridurre la quantità di dati, mantenendo però le informazioni essenziali. La tecnica è quella di eliminare le eventuali ridondanze, lasciando soltanto le informazioni principali e necessarie. Le tecniche che hanno lo scopo di ridurre la quantità di dati necessari a rappresentare l'immagine, vanno sotto il nome di compressione. La compressione può essere vista come una trasformazione matematica, applicata all'immagine di partenza, che restituisce un'immagine priva di dati ridondanti. La compressione delle immagini ha importanti applicazioni, fra cui videoconferenze, trasmissione di immagini satellitari, trasmissioni FAX, controllo remoto di veicoli militari, etc.

7.1.1.1 informazioni e dati

I termini **dato** e informazione non indicano la stessa cosa.

Informazione: è una parte di conoscenza. Acquisendo un'informazione si accresce la conoscenza e si riduce il livello di incertezza.

Dato: attraverso il dato viene trasmessa l'informazione, presupponendone un'interpretazione.

Un dato, in sè e per sè, non comporta interpretazione e pertanto non apporta alcuna conoscenza. **Un dato corredato di un opportuno significato costituisce un'informazione**

7.2 Ridondanza

La **ridondanza dei dati** è un punto centrale nella compressione di un'immagine digitale. Non si tratta di un concetto astratto, bensì di una vera e propria entità matematica quantificabile. Più precisamente, la stessa informazione può essere rappresentata da diverse quantità di dati, ad esempio pari a n_1 e n_2 : allora la ridondanza relativa dei dati è definita come:

$$R_D = 1 - \frac{1}{C_R}$$

Esempio Un rapporto di compressione $C_R = \frac{n_1}{n_2} = 10$ indica che a 10 dati (ad esempio bits) del primo insieme corrisponde 1 dato dell'insieme compresso. La ridondanza relativa è allora $R_D = 1 - \frac{1}{C_R} = 0.9$ il che indica che il **90% dei dati presenti nel primo insieme sono ridondanti**.

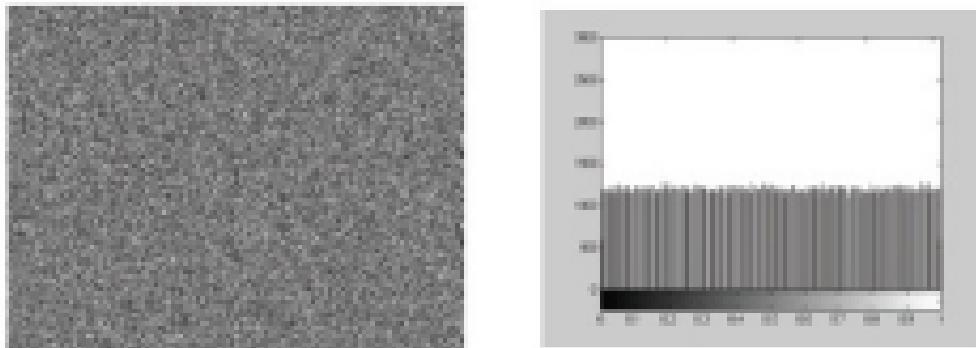
Nel campo della compressione di immagini si possono individuare tre tipi di ridondanza:

- la ridondanza nella codifica
- la ridondanza interpixel
- la ridondanza psicovisuale

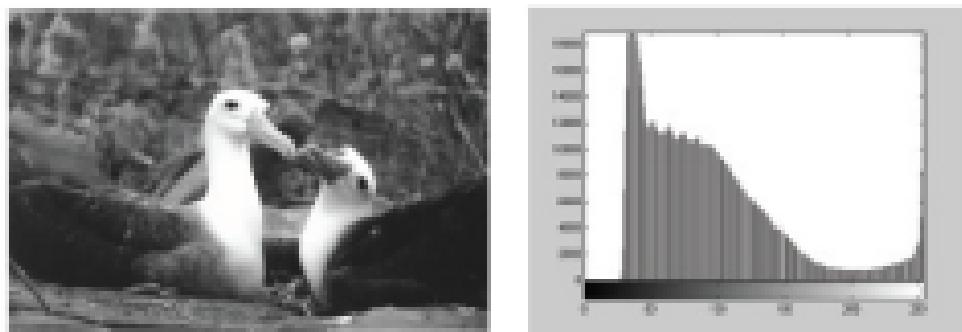
7.2.0.1 Ridondanza nella codifica

La **ridondanza nella codifica** deriva dalla scelta del codice (binario) adottato per rappresentare il colore oppure il livello di grigio assunto da un pixel. Può essere causata da un numero eccessivo di bits per pixel, oppure dall'ipotesi, spesso non vera, che tutti i valori che un pixel può assumere sono equiprobabili. **Esempio ridondanza nella codifica** immagine costituita da punti casuali





e una figura naturale



7.2.0.2 Ridondanza interpixel

La ridondanza interpixel (detta anche talvolta ridondanza spaziale, o geometrica, o interframe) si deve al fatto che in genere c'è una correlazione fra i valori assunti da pixel vicini: è molto probabile che pixel vicini assumano valori di intensità o di colori piuttosto simili.

7.2.0.3 Ridondanza psicovisuale

La **ridondanza psicovisuale** nasce dal fatto che l'occhio umano non percepisce con la stessa sensibilità tutte le informazioni visive: alcune informazioni hanno minore importanza di altre. Ad esempio, nel riconoscimento di un oggetto risultano più importanti i contorni che non il corpo dell'oggetto. **Esempio** La prima immagine presenta 256 livelli di grigio, mentre la seconda soltanto 64. L'occhio umano non è in grado di percepire una differenza marcata, ma la seconda immagine utilizza 192 colori in meno rispetto alla prima.



Figure 7.1: Immagine a profondità 8 (256 livelli di grigio) e immagine a profondità 6 (64 livelli)

7.2.0.4 Ridondanza

- La ridondanza è molto utile nel processo di compressione in quanto consente di eliminare una certa quantità di dati.
- E' bene però tenere presente che, nell'eliminare ridondanza interpixel o psicovisuale, è possibile avere anche una perdita di qualità.
- Per questo motivo è utile avere dei criteri con cui valutare la natura e l'entità delle perdite durante la compressione (criteri oggettivi, che hanno una formulazione matematica, e criteri soggettivi, quali ad esempio i sondaggi).

7.3 Il processo di compressione

Nel processo di compressione, l'immagine originale viene compressa attraverso un **coder**, il quale elimina in genere la ridondanza spaziale. I dati ottenuti vengono trasmessi poi ad un **decoder** che ricostruisce l'immagine di partenza, riaggiungendo alcuni dati ridondanti che risultano significativi.