

Elaborazione di segnali e immagini

VR443470

febbraio 2023

Indice

1 Fondamenti	5
1.1 Matematica preliminare	5
1.1.1 Numeri complessi	5
1.1.2 Funzioni complesse di variabile reale	6
1.1.3 Funzioni pari e dispari	7
1.1.4 Segnali periodici	8
1.2 Operazioni fondamentali	9
1.2.1 Somma	9
1.2.2 Shift (o traslazione)	10
1.2.3 Funzione box Π e impulso di Dirac	11
1.2.4 Funzione sinc	12
1.2.5 Funzione triangolo Λ	12
1.2.6 Funzione segno (sgn)	12
1.2.7 Funzione gradino	12
1.2.8 Treno di impulsi	13
1.2.9 Energia di un segnale	13
1.2.10 Potenza media di un segnale	14
1.3 Altre operazioni fondamentali	15
1.3.1 Rescaling (o riscalatura)	15
1.3.2 Cross-Correlazione	16
1.3.3 Esercizi d'esame	17
1.3.4 Cross-Correlazione Normalizzata	23
1.3.5 Convoluzione	29
2 Analisi di Fourier	30
2.1 Serie di Fourier	30
2.1.1 Proprietà della serie di Fourier	36
2.2 Trasformata di Fourier continua	37
2.2.1 Trasformata di Fourier	37
2.2.2 Trasformata di Fourier inversa	37
2.2.3 Proprietà della trasformata di Fourier	39
2.2.4 Trasformata di Fourier di una box	40
2.2.5 Trasformata di Fourier di un sinc	41
2.2.6 Trasformata di Fourier di un impulso	42
2.2.7 Trasformata di Fourier di un treno di impulsi	43
2.2.8 Sintesi	44
2.3 Trasformata di Fourier a tempo discreto	45
2.3.1 Campionamento	45
2.3.2 Trasformata di Fourier a tempo discreto	46
2.3.3 Teorema del campionamento	47
2.3.4 Considerazioni	47
2.4 Trasformata di Fourier discreta	48
2.5 Riassunto Trasformate	50
2.6 Domanda da esame	51

3 Elaborazione di immagini - Dominio spaziale	52
3.1 Strumento per l'elaborazione: istogramma	53
3.2 Domini	55
3.3 Operazioni puntuali	56
3.3.1 Identità	57
3.3.2 Negativo	57
3.3.3 Clamping	57
3.3.4 Stretching/Shrinking	57
3.3.5 Trasformazione logaritmica	58
3.3.6 Trasformazione esponenziale	58
3.3.7 Trasformazione di potenza	59
3.3.8 Binarizzazione	59
3.3.9 Binarizzazione attraverso il metodo di Otsu	60
3.3.10 Equalizzazione	61
3.4 Operazioni locali	65
3.4.1 Filtraggi spaziali: lineari e non lineari	66
3.5 Rumore nelle immagini	68
3.5.1 Rumore gaussiano additivo bianco	69
3.5.2 Rumore impulsivo	70
3.6 Altre operazioni locali: tipologie di filtraggio	71
3.6.1 Smoothing - Filtro media	72
3.6.2 Smoothing - Filtro mediano	73
3.6.3 Smoothing - Filtro gaussiano	74
3.6.4 Domanda da esame	75
3.6.5 Filtraggi di sharpening	76
3.6.6 Sharpening - Basic Highpass Spatial Filtering	78
3.6.7 Sharpening - Filtro Laplaciano	79
3.6.8 Sharpening - High Boost Filtering	79
4 Elaborazione di immagini - Rinforzo del dominio delle frequenze	80
4.1 Ripasso formule utili	80
4.2 Rumore nel dominio delle frequenze	82
4.3 Panoramica su filtri passa alto (<i>high-pass</i>) e passa basso (<i>low-pass</i>)	83
4.4 Filtri passa basso (<i>low-pass</i>)	84
4.4.1 Filtri passa basso ideale	84
4.4.2 Filtri passa basso di Butterworth	86
4.4.3 Filtri passa basso Gaussiano	88
4.4.4 Sintesi	88
4.5 Filtri passa alto (<i>high-pass</i>)	89
4.5.1 Filtri passa alto ideale	90
4.5.2 Filtri passa alto di Butterworth	90
4.5.3 Filtri passa alto Gaussiano	91
4.6 Filtri per enfatizzare le alte frequenze	92
4.7 Filtri passa banda ideale 1D e ferma banda ideale	93

5 Elaborazione di immagini - Estrazione di edge robusta	96
5.1 Estrazione di contorni (<i>edge</i>)	96
5.2 Criteri per <i>edge detector</i> ottimi	98
5.3 Il gradiente	99
5.4 Estrazione dell' <i>edge</i>	100
5.4.1 Il gradiente discreto	100
5.4.2 Operatore di Roberts	100
5.4.3 Operatore di Prewitt	100
5.4.4 Operatore di Sobel	101
5.4.5 Operatore di Kirsch	101
5.4.6 Operatore di Robinson	101
5.4.7 Confronto tra operatori	102
5.5 Derivata prima	104
5.5.1 Problemi	104
5.5.2 Effetti del rumore	104
5.5.3 Rilevazione di <i>edge</i> o <i>detection</i> robusta	104
5.5.4 Teorema di convoluzione	105
5.6 Derivata seconda	106
5.6.1 Operatori di derivata seconda	106
5.6.2 Operatore Laplaciano	106
5.7 Laplaciano di Gaussiana	107
5.8 Gradiente e Laplaciano di Gaussiana a confronto	110
5.9 Filtro di Canny	111
5.9.1 Definizione	111
5.9.2 Modello di <i>edge</i> per Canny	111
5.9.3 Algoritmo di Canny	111
5.9.4 Filtro di Rinforzo	112
5.9.5 Non-Maxima Suppression	113
5.9.6 Sogliatura ad isteresi	114
5.9.7 Edge linking	115
5.10 Prestazioni di un <i>edge detector</i>	116
5.11 Figura di merito per gli <i>edge</i>	116
5.12 Trasformata di Hough	117
5.12.1 Definizione	117
5.12.2 Trasformata per le rette	119
5.12.3 Trasformata per i cerchi	121

1 Fondamenti

1.1 Matematica preliminare

1.1.1 Numeri complessi

Un numero complesso c appartiene all'insieme dei complessi \mathbb{C} e la sua forma è del tipo:

$$c = \Re + j\Im$$

con \Re, \Im variabili $\in \mathbb{R}$ e j chiamata *unità immaginaria* rappresentata come $j = \sqrt{-1}$. Inoltre, \Re rappresenta la *parte reale* e \Im la *parte immaginaria*. Il coniugato di c è

$$\tilde{c} = \Re - j\Im$$

I numeri complessi, dal punto di vista geometrico, possono essere visti come punti su un piano (chiamato *piano complesso*) e descritti da coordinate (R, I) . Nel piano complesso, le ascisse (x) sono rappresentate dalla parte reale, mentre le ordinate (y) dalla parte immaginaria.

Spesso è utile rappresentare i numeri complessi in coordinate polari formate nel seguente modo (*modulo, angolo*). Questa forma viene denominata *forma polare* di un numero complesso:

$$c = \Re + j\Im = |c|(\cos \theta + j \sin \theta)$$

dove:

$$|c| = \sqrt{\Re^2 + \Im^2} \longrightarrow \text{chiamato } \textit{modulo} \text{ o } \textit{magnitudo}$$

invece, *theta* rappresenta:

$$\theta \cong \arctan \left(\frac{\Im}{\Re} \right) \longrightarrow \text{chiamato } \textit{angolo}, \textit{fase} \text{ o } \textit{argomento in radianti}$$

Grazie alla formula di Eulero:

$$e^{j\theta} = \cos \theta + j \sin \theta$$

è possibile riscrivere la forma polare di un numero complesso in maniera alternativa, ossia:

$$c = \Re + j\Im = |c| (\cos \theta + j \sin \theta) = |c| e^{j\theta}$$

La **somma** e la **moltiplicazione** di due numeri complessi diventa:

$$c_1 = R_1 + jI_1 \quad c_2 = R_2 + jI_2$$

$$\text{Somma: } c_1 + c_2 = (R_1 + R_2) + j(I_1 + I_2)$$

$$\text{Moltiplicazione con Eulero: } c_1 \cdot c_2 = (R_1 R_2 - I_1 I_2) + j(R_1 I_2 + I_1 R_2) \longrightarrow = |c_1||c_2| e^{j(\theta_1 + \theta_2)}$$

1.1.2 Funzioni complesse di variabile reale

Dato $t \in \mathbb{R}$, una funzione f complessa di variabile reale è $f : D_1 \subseteq \mathbb{R} \rightarrow D_2 \subseteq \mathbb{C}$. Viene introdotto questo concetto poiché il **fasore** è un eSEMPIO fondamentALE. Le **caratteristiche** di questa funzione:

- È una funzione complessa che modella la posizione di un punto che ruota attorno all'origine con raggio determinato $|c|$ e velocità angolare costante $\theta(t)$.
- Se la funzione fosse nei numeri reali, sarebbe più dispendioso in termini di numero di funzioni da utilizzare.

L'**obiettivo** dei fasori è quello di *passare dal dominio del tempo* (o spazio) a *quello dell'analisi frequenziale*.

La particolarità è che nel tempo il fasore riesce a variare un numero complesso (in forma polare) mantenendo il modulo $|c|$ fisso:

$$|c|e^{j\theta} \rightarrow |c|e^{j\theta(t)}$$

dove $\theta(t)$ indica la **velocità angolare**. Quest'ultima può essere calcolata tramite:

$$\theta(t) \longrightarrow \frac{2\pi}{T_0}t + \phi$$

dove T_0 indica il *tempo* impiegato per eseguire 2π radianti.

Soltamente si utilizza il fasore con le seguenti supposizioni:

- Coordinate rappresentate con (R, I)
- Impostata una distanza unitaria fissa dall'origine $|c| = 1$
- Velocità angolare costante pari a $2\pi/\text{sec.}$, ossia $\theta(t) = 2\pi t, T_0 = 1\text{sec.}$
- Con $t = 0$ si ha $\theta = 0$
- Viene mantenuto $\phi = 0$

1.1.3 Funzioni pari e dispari

Una funzione $f : \mathbb{R} \rightarrow \mathbb{R}$ è **pari** se e solo se:

$$f(t) = f(-t)$$

Invece, una funzione $f : \mathbb{R} \rightarrow \mathbb{R}$ è **dispari** se e solo se:

$$f(t) = -f(-t)$$

1.1.4 Segnali periodici

Un segnale f è **periodico** di periodo T o T -periodico se:

$$\exists T_0 \in R^+ : f(t + T_0) = f(t), \quad \forall t \in D_1$$

e T_0 è il minor numero per cui la condizione di ripetizione si verifica.

Dato un periodo T_0 con la lettera μ_0 si indica la **frequenza fondamentale**:

$$\mu_0 = \frac{1}{T_0}$$

Fissato $T_0 > 0$ i **segnali trigonometrici** di minimo periodo T_0 sono:

$$f(t) = \cos(2\pi\mu_0 t) \quad f(t) = \sin(2\pi\mu_0 t)$$

dove μ è una frequenza generale, mentre $\mu_0 = \frac{1}{T_0}$ è la **frequenza fondamentale**. Invece, spesso la **velocità angolare** o **pulsazione** viene rappresentata come:

$$2\pi\mu_0 = \frac{2\pi}{T_0} = \omega_0$$

Inoltre, fissato un $\theta \in \mathbb{R}$ chiamato **fase** si osserva che anche le funzioni:

$$f(t) = \cos(2\pi\mu_0 t + \theta) \quad f(t) = \sin(2\pi\mu_0 t + \theta)$$

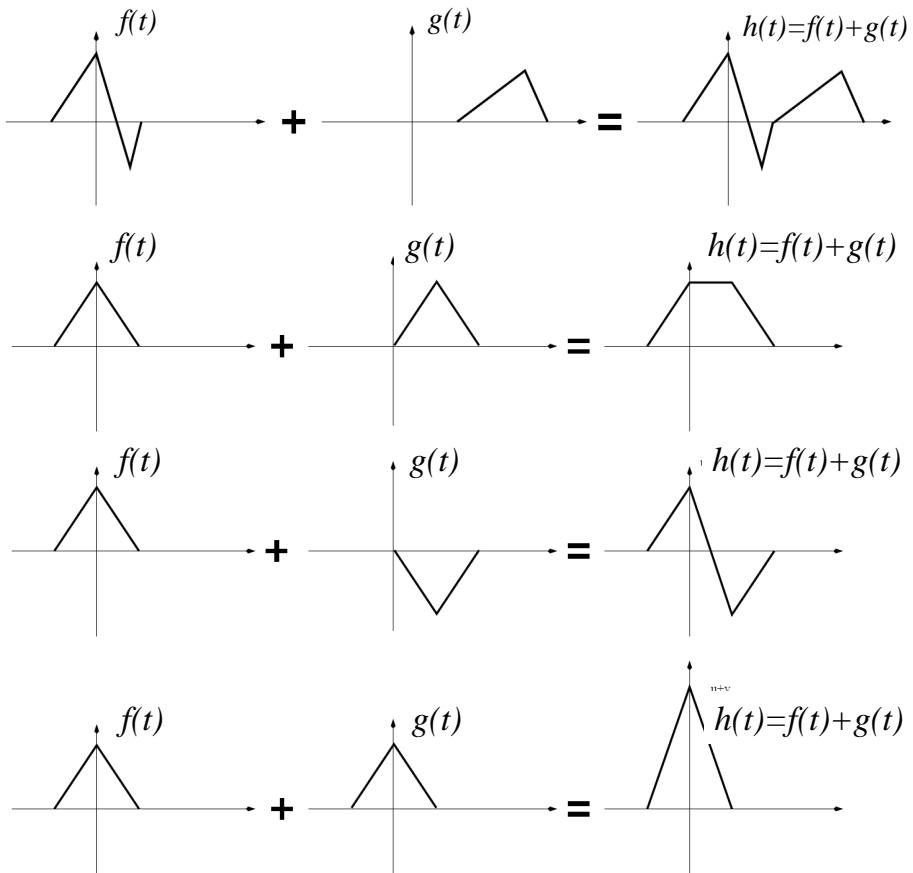
hanno il medesimo periodo T .

Infine, la fase θ permette di eseguire operazione di *shift*.

1.2 Operazioni fondamentali

1.2.1 Somma

La *somma* di due segnali è facile quando essi non interferiscono, ovvero quando **non** sono contemporaneamente $\neq 0$. Alcuni esempi qui di seguito.



1.2.2 Shift (o traslazione)

Lo **shift** (o traslazione) è il cambio di posizione di un segnale. Può essere effettuato:

- **Traslazione a destra** con la funzione $f(t - \tau)$
- **Traslazione a sinistra** con la funzione $f(t + \tau)$

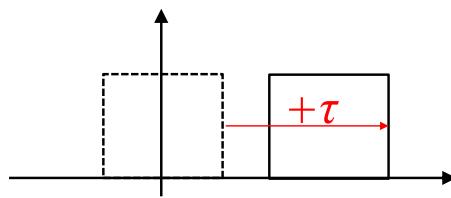


Figura 1: Shift a destra

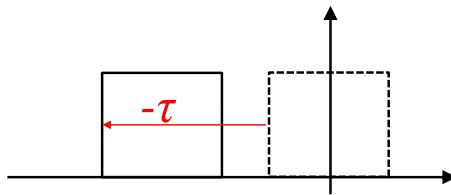


Figura 2: Shift a sinistra

1.2.3 Funzione box II e impulso di Dirac

La funzione ***box*** è definita nel seguente modo:

$$A\Pi\left(\frac{x}{b}\right) \quad x \in \left[-\frac{b}{2}, \frac{b}{2}\right]$$

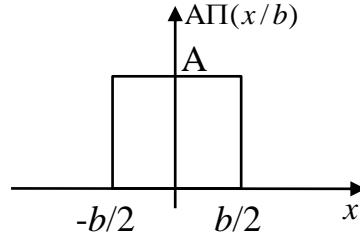


Figura 3: Box generica

La funzione $\delta(x)$ è chiamata ***impulso unitario*** o ***impulso di Dirac*** perché è definita nel seguente modo:

$$\delta(x) = \begin{cases} \infty & \text{se } x = 0 \\ 0 & \text{se } x \neq 0 \end{cases} \quad \int_{-\infty}^{\infty} \delta(x) dx = 1$$

Quindi è un impulso che tende all'infinito solamente quando la x è nell'origine, ma il suo integrale è uguale a 1. Alcune **proprietà** dell'impulso:

1. $\delta(x - x_0) = 0 \quad \forall x \neq x_0$
2. Data una funzione generica f (**setacciamento**): $\int_{-\infty}^{\infty} f(x)\delta(x - x_0) dt = f(x_0)$
3. $\delta(x - x_0) = \delta(x_0 - x)$
4. $\delta(ax) = \frac{1}{|a|}\delta(x) \quad \forall x \in \mathbb{R}, \text{ fissato } a \in \mathbb{R} - \{0\}$

1.2.4 Funzione sinc

La funzione **sinc** è definita nel seguente modo:

$$\text{sinc}(t) = \frac{\sin(\pi t)}{\pi t}$$

Ha due **caratteristiche** importanti: (1) l'intersezione con l'asse delle x avviene sempre nei numeri interi positivi e negativi (quindi 1 e -1, 2 e -2, ecc.); (2) il limite $\lim_{t \rightarrow \pm\infty} \text{sinc}(t) = 0$.

Questa funzione è **importante per l'analisi nel dominio del tempo (o frequenza)**.

1.2.5 Funzione triangolo Λ

La funzione **triangolo** è definita nel seguente modo:

$$\Lambda(x) = \begin{cases} 1 - |x|, & |x| < 1 \\ 0 & \text{altrimenti} \end{cases}$$

Questa funzione è **importante per l'analisi spettrale e per le operazioni di convoluzione**.

1.2.6 Funzione segno (sgn)

La funzione **segno** è definita nel seguente modo:

$$\text{sgn}(x) = \begin{cases} -1, & x < 0 \\ +1, & x > 0 \\ 0 & x = 0 \end{cases}$$

Questa funzione ribalta segnali sopra o sotto l'asse delle x .

1.2.7 Funzione gradino

La funzione **gradino** è definita nel seguente modo:

$$u(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$$

Questa funzione rappresenta un **segnale** che si attiva a partire dal tempo specificato e rimane attivo indefinitamente. Attenzione! Non si confonda questo segnale con il segno.

1.2.8 Treno di impulsi

Il **treno di impulsi** $S_{\Delta T}(x)$ è la somma di un numero infinito di impulsi periodici discreti distanziati di una quantità ΔT :

$$S_{\Delta T}(x) = \sum_{n=-\infty}^{\infty} \delta(x - n\Delta T) \quad n \in \mathbb{Z}$$

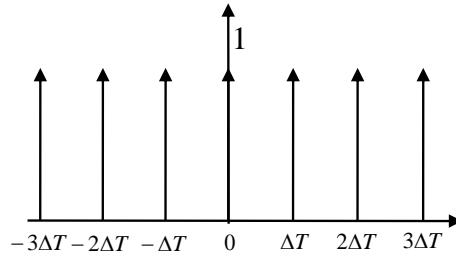


Figura 4: Treno di impulsi

1.2.9 Energia di un segnale

L'**energia di un segnale** è definita nel seguente modo:

$$E_f = \begin{cases} \int_{-\infty}^{+\infty} f^2(t) dt & \text{se } f \in \mathbb{R} \\ \int_{-\infty}^{+\infty} |f(t)|^2 dt & \text{con } |f(t)|^2 = \tilde{f}(t)f(t), \quad f \in \mathbb{C} \end{cases}$$

Un segnale si dice **ad energia finita** (o **di energia**) se l'integrale che rappresenta l'energia converge ed è diverso da 0. Quindi:

- ☞ **Condizione sufficiente** all'esistenza della sua trasformata di Fourier. Le funzioni trigonometriche non sono di energia ma hanno comunque la Trasformata di Fourier.
- ☞ **Condizione necessaria** per essere un segnale ad energia finita, all'infinito ($+\infty$ e $-\infty$) l'**ampiezza** va a zero.

Alcuni esempi:

- ★ **Segnali di energia.** Impulsi rettangolari, oscillazioni smorzate (sinc);
- ★ **Segnali non di energia.** Funzioni trigonometriche sin e cos.

L'**unità di misura** è il *joule*.

1.2.10 Potenza media di un segnale

La *potenza media di un segnale* è definita nel seguente modo:

$$P_f = \begin{cases} \lim_{T \rightarrow +\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{+\frac{T}{2}} f^2(t) dt & \text{se } f \in \mathbb{R} \\ \lim_{T \rightarrow +\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{+\frac{T}{2}} |f(t)|^2 dt \quad \text{con } |f(t)|^2 = \tilde{f}(t)f(t), & f \in \mathbb{C} \end{cases}$$

Un segnale si dice **a potenza finita** (o **di potenza**) se l'integrale che rappresenta la potenza converge ed è diverso da 0. L'**unità di misura** è il *watt*.

Infine, un segnale ad energia finita ha la potenza che tende a zero (per cui un segnale non può appartenere ad entrambe le categorie). Invece, esistono segnali che non sono né di energia, né di potenza finita.

1.3 Altre operazioni fondamentali

1.3.1 Rescaling (o riscalatura)

La funzione di *rescaling* è definita nel seguente modo:

$$\forall f(t) : D_1 \in \mathbb{R}, \quad \omega \neq 0$$

Simile allo *shift*, il *rescaling* ha una definizione generica e due varianti:

- **Definizione generica** con la funzione semplice $f(t)$ (immagine 5).
- **Ritardo lineare del segnale di un fattore ω** con la funzione $f(\omega t)$, $0 < \omega < 1$ (immagine 6).
- **Accelero lineare del segnale di un fattore ω** con la funzione $f(\omega t)$, $\omega > 1$ (immagine 7).

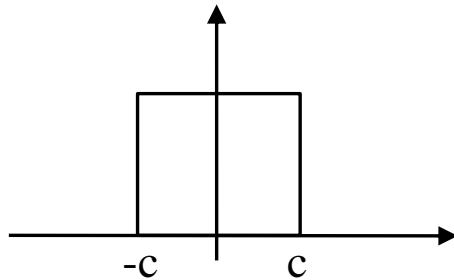


Figura 5: Definizione generica

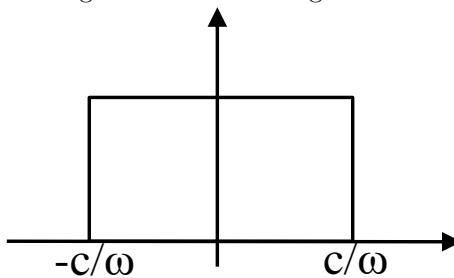


Figura 6: Ritardo lineare del segnale di un fattore ω

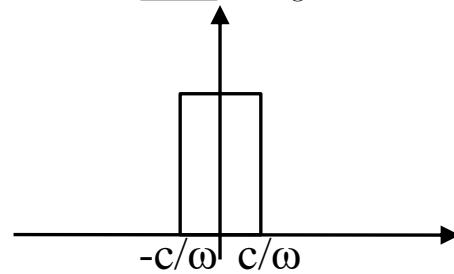


Figura 7: Accelero lineare del segnale di un fattore ω

1.3.2 Cross-Correlazione

Dati $f_1(\tau), f_2(\tau)$ segnali continui, $\tau \in \mathbb{R}$ il segnale di ***cross-correlazione*** viene definito come:

$$f_1 \otimes f_2(t) = \int_{-\infty}^{+\infty} \tilde{f}_1(\tau) f_2(\tau - t) d\tau$$

In cui $\tilde{f}_1(\tau)$ rappresenta un *complesso coniugato*. Nel caso in cui f_1 è reale, allora $\tilde{f}_1(\tau) \rightarrow f_1(\tau)$.

Infine, con $t = 0$ si ha l'***integrale di cross-correlazione***, il quale è definito se l'integrale converge (ovviamente se il segnale non è né di energia, né di potenza, la convergenza non esiste!).

1.3.3 Esercizi d'esame

Esercizio.

Il primo esercizio fornisce una funzione $f(t)$:

$$f(t) = \Pi\left(\frac{t-2}{4}\right) e^{-2t}$$

Le **richieste** dell'esercizio sono le seguenti:

- I Rappresentare graficamente il segnale;
- II Calcolare sia l'energia che la potenza media. Inoltre, dire se $f(t)$ è una funzione di energia o di potenza fornendo una motivazione valida. Infine, calcolare l'energia o la potenza nel caso in cui $f(t)$ sia solo composta da e^{-2t} ;
- III Scrivere l'espressione analitica rispetto $z(t) = -f(-t)$ e $v(t) = f(t+4)$

Risoluzione I.

Il **primo passo** è quello di scomporre la funzione così da avere una visione più chiara sulle operazioni da effettuare:

$$f(t) = \Pi\left(\frac{t-2}{4}\right) e^{-2t} \longrightarrow f(t) = \Pi\left(\frac{1}{4} \cdot (t-2)\right)$$

Come si può osservare, ci sono due operazioni da eseguire. Quindi, dopo l'esplicitazione si esegue la rappresentazione del segnale base $\Pi(t)$:

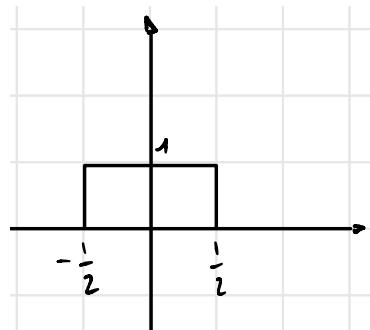


Figura 8: Rappresentazione della funzione $f(t)$, ovvero un box.

Adesso si esegue l'operazione di moltiplicazione per un fattore che in questo caso è $\frac{1}{4}$. Quindi si rappresenta la box $\Pi\left(\frac{1}{4} \cdot t\right)$:

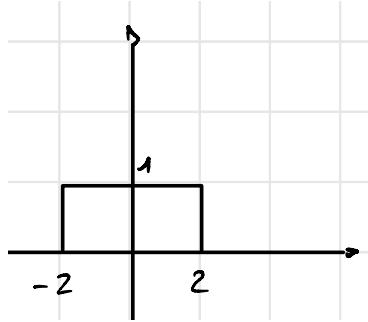


Figura 9: Box $\Pi\left(\frac{1}{4} \cdot t\right)$ allargata.

L'operazione che è stata effettuata è stata semplicemente considerare la box del tipo $\Pi\left(\frac{t}{4}\right)$. Ricordandosi le nozioni del corso di Sistemi, per definizione quindi la box è definita nell'intervallo $-2, +2$.

Infine, viene applicata l'ultima operazione, ovvero il -2 all'incognita t . Quindi, la funzione box diventerà $\Pi\left(\frac{1}{4}(t - 2)\right)$ e la sua rappresentazione grafica sarà uno shift a destra (ritardo):

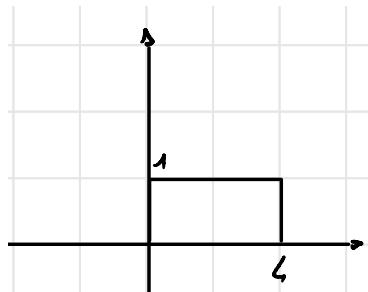


Figura 10: Box $\Pi\left(\frac{1}{4}(t - 2)\right)$ dopo lo shift a destra.

Il **primo punto si conclude** con la rappresentazione del segnale e^{-2t} e la sua combinazione con la box. Quindi:

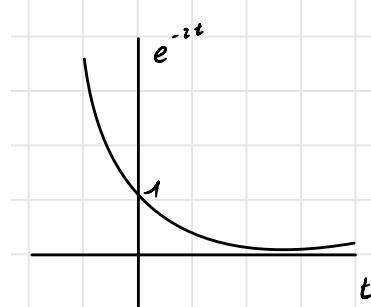


Figura 11: Rappresentazione della funzione e^{-2t}

E infine la sua concatenazione con la box, quindi una sorta di applicazione di un filtro:

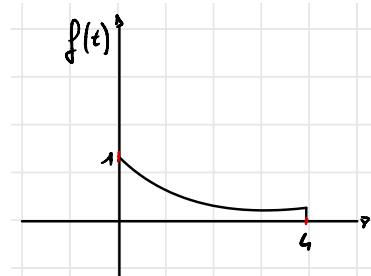


Figura 12: Rappresentazione finale della funzione $f(t) = \Pi\left(\frac{t-2}{4}\right) e^{-2t}$

Risoluzione II.

Guardando la figura 12 si può già intuire che tipo di segnale sia. Infatti, dato che è limitato e non si estende all'infinito, per definizione è un **segnale finito**, quindi **di energia e non di potenza**. Per dimostrare questa affermazione, si eseguono i calcoli:

$$\text{Definizione di energia: } E_f = \int_{-\infty}^{\infty} |f(t)|^2 dt = \int_{-\infty}^{\infty} f(t)^2 dt$$

$$\text{Definizione di potenza: } P_f = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f^2(t) dt$$

Dopo le definizioni, si esegue l'effettivo calcolo con i valori numerici:

Energia finita

$$E_f = \int_0^4 e^{-4t} dt = \left. \frac{e^{-4t}}{-4} \right|_0^4 = \frac{-e^{-16} + 1}{4} = \frac{1}{4} \neq 0$$

Potenza finita

$$P_f = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^4 e^{-4t} dt = \lim_{T \rightarrow \infty} \frac{1}{T} \cdot \frac{1}{4} = 0$$

Come si osserva dai risultati, è un segnale di energia finita poiché è un valore noto, invece non è un segnale di potenza poiché il risultato è zero e non rispetta la definizione.

Al contrario, se la funzione fosse composta solamente dall'esponenziale, il calcolo dell'energia e della potenza sarebbe:

$$\text{Energia: } E_f = \int_{-\infty}^{\infty} e^{-4t} dt = \left. \frac{e^{-4t}}{-4} \right|_{-\infty}^{\infty} = \lim_{T \rightarrow \infty} \frac{e^{-4t} - e^{4t}}{-4} = \infty$$

$$\text{Potenza: } P_f = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} e^{-4t} dt = \lim_{T \rightarrow \infty} \frac{1}{T} \cdot \left. \frac{e^{-4t}}{-4} \right|_{-\frac{T}{2}}^{\frac{T}{2}} = \lim_{T \rightarrow \infty} \frac{e^{-2T} - e^{2T}}{-4T} = \infty$$

Come si evince dai calcoli, il segnale non è né di energia né di potenza perché entrambi i risultati sono uguali a infinito.

Risoluzione III.

Considerando la funzione $z(t)$, si osserva che è la copia simmetrica rispetto all'origine di $f(t)$. Invece, la funzione $v(t)$ è identica alla funzione $f(t)$ ma "shiftata" a sinistra di 4:

$$f(t) = -f(-t) \quad v(t) = f(t+4)$$

Esercizio 2.

Il secondo esercizio fornisce una funzione $f(t)$:

$$f(t) = \operatorname{sgn} \left(a \cdot \cos \left(\frac{2\pi}{T_0} t \right) \right)$$

Con $T_0 = 2$. Le **richieste** dell'esercizio sono le seguenti:

- I Rappresentare graficamente il segnale;
- II Calcolare sia l'energia che la potenza media. Inoltre, dire se $f(t)$ è una funzione di energia o di potenza fornendo una motivazione valida.

Risoluzione I.

Viene rappresentato il segnale della funzione segno sng:

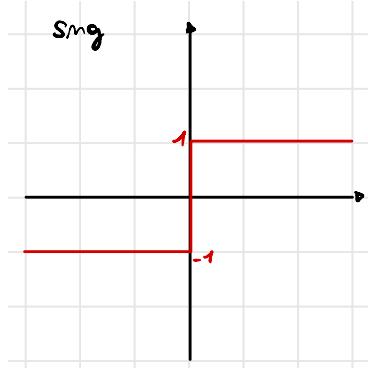


Figura 13: Funzione segno sng.

Si esplicitando le operazioni della funzione:

$$f(t) = \operatorname{sgn} \left(a \cdot \cos \left(\frac{2\pi}{T_0} t \right) \right) = \cos \left(\frac{1}{T_0} \cdot 2\pi t \right)$$

E si rappresenta inizialmente la funzione $\cos(2\pi)$ con $T_0 = 1$:

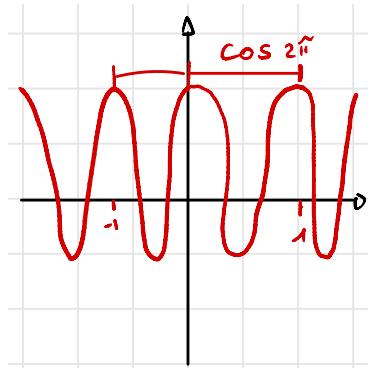


Figura 14: Funzione coseno $\cos(2\pi)$.

Si conclude la rappresentazione grafica aumentando T_0 in maniera molto semplice:

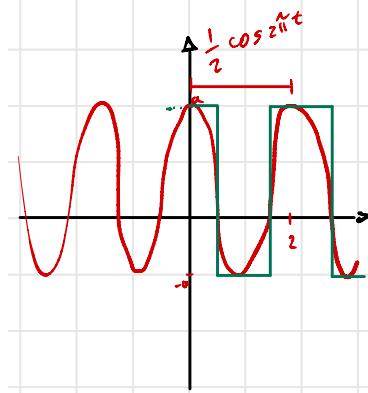


Figura 15: Funzione coseno $\cos(2\pi)$ moltiplicata per $\frac{1}{T_0} = \frac{1}{2}$.

Risoluzione II.

Si conclude l'esercizio calcolando l'energia o la potenza del segnale. Per farlo, dato che non è definito in un intervallo ma continua all'infinito, si calcolano i rispettivi integrali in un intervallo arbitrario n e poi lo si estende all'infinito:

$$\begin{aligned} E_f &= \int_{-\infty}^{\infty} f^2(t) dt = \lim_{n \rightarrow \infty} \int_{-n \cdot \frac{T_0}{2}}^{n \cdot \frac{T_0}{2}} f^2(t) dt = \lim_{n \rightarrow \infty} n \cdot \int_{-\frac{T_0}{2}}^{\frac{T_0}{2}} f^2(t) dt = \infty \\ P_f &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f^2(t) dt = \lim_{n \rightarrow \infty} \frac{1}{n T_0} \int_{-n \frac{T_0}{2}}^{n \frac{T_0}{2}} f^2(t) dt = \lim_{n \rightarrow \infty} \frac{1}{n T_0} \cdot n \cdot \int_{-\frac{T_0}{2}}^{\frac{T_0}{2}} f^2(t) dt = \\ &= \frac{1}{T_0} \cdot T_0 = \frac{1}{2} \cdot 2 = 1 \longrightarrow \neq 0 \end{aligned}$$

È evidente che il segnale è di potenza. Come si evince dalla figura 15, i tratti di colore verde indicano il rettangolo formato dal segnale. Calcolando l'area del rettangolo, si ottiene esattamente il valore di T_0 . Infatti, la base del rettangolo (verticale) è 2, mentre l'altezza (orizzontale) è 1.

1.3.4 Cross-Correlazione Normalizzata

Ha l'**obiettivo** di trattare segnali con range di valori diversi e consente di eseguire **confronti uno-a-molti** (*one-to-many*):

$$f_1 \bar{\otimes} f_2 (t) = \frac{\int_{-\infty}^{+\infty} \tilde{f}_1 (\tau) f_2 (\tau - t) d\tau}{\sqrt{E_{f_1} E_{f_2}}}$$

In cui E_f indica l'**energia** del segnale f . Ci sono due caratteristiche importanti:

- $f_1 \bar{\otimes} f_2 (t) \in [-1, 1]$
- $|f_1 \bar{\otimes} f_2 (t)| = 1 \iff f_1 (\tau) = \alpha f_2 (\tau - t)$

Inoltre, si parla di **autocorrelazione** (normalizzata e non) quando $f_1 = f_2$. Utile per i segnali stocastici.

Nel **caso di segnali discreti**, dati $x_1(k), x_2(k)$:

$$x_1 \otimes x_2 (n) = \sum_{k=-\infty}^{+\infty} \tilde{x}_1 (k) x_2 (k - n) \quad k \in \mathbb{Z}$$

Sotto l'ipotesi di convergenza della serie, cioè la serie deve convergere.

Nel caso in cui $x_1(k)$ e $x_2(k)$ sono limitati di lunghezza M ed N rispettivamente, allora la **cross correlazione è di lunghezza $M + N - 1$** .

Cross-Correlazione 1D

Data la definizione:

$$x_1 \otimes x_2(n) = \sum_{k=-\infty}^{+\infty} x_1(k) x_2(k-n)$$

Esistono diverse casistiche:

- $n = 0$ si confronta tra x_1 e x_2 nei loro domini temporali originali.
- $n > 0$ sposta x_2 a destra poiché c'è l'anticipo di x_2
- $n < 0$ sposta x_2 a sinistra poiché c'è ritardo di x_2

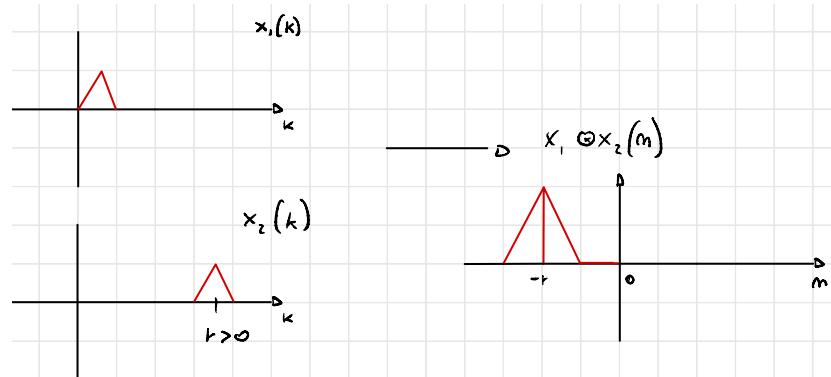


Figura 16: Esempio di cross-correlazione normalizzata 1D.

Il triangolo x_2 va verso sinistra e il lasso di tempo che x_2 non combacia con x_1 , viene rappresentato come una linea orizzontale sull'asse delle n nel piano cartesiano di destra.

Cross-Correlazione 2D

Data la definizione:

$$x_1 \otimes x_2 (m, n) = \sum_{u=-\infty}^{+\infty} \sum_{v=-\infty}^{+\infty} x_1(u, v) x_2(u - m, v - n) \quad u, v, m, n \in \mathbb{Z}$$

Nel 2D x_1 e x_2 possono essere pensate come **immagini infinite**.

Di solito x_1 e x_2 sono **immagini finite** (segnali digitali ad intervallo limitato), e gli estremi di sommatoria sono quindi finiti.

Il primo segnale x_1 viene chiamato **template**, o **matrice kernel**, mentre x_2 genericamente **immagine** (di solito, la matrice kernel x_1 ha una dimensionalità minore di quella dell'immagine).

Nel caso $x_1 = x_2$ si ha **autocorrelazione 2D**.

Cross-Correlazione normalizzata 2D

Si definisce come:

$$x_1 \otimes x_2 (m, n) = \frac{\sum_{u=-\infty}^{+\infty} \sum_{v=-\infty}^{+\infty} [x_1(u, v)] [x_2(u - m, v - n)]}{\sqrt{\sum_{u=-\infty}^{+\infty} \sum_{v=-\infty}^{+\infty} [x_1(u, v)]^2 \sum_{u=-\infty}^{+\infty} \sum_{v=-\infty}^{+\infty} [x_2(u, v)]^2}}$$

In altre parole, fissato il punto di applicazione n, m , si sottrae la media ad ogni punto nell'interno di applicazione dalla matrice kernel. Successivamente, si divide per il prodotto della varianza dei due segnali, estraendo a radice alla fine.

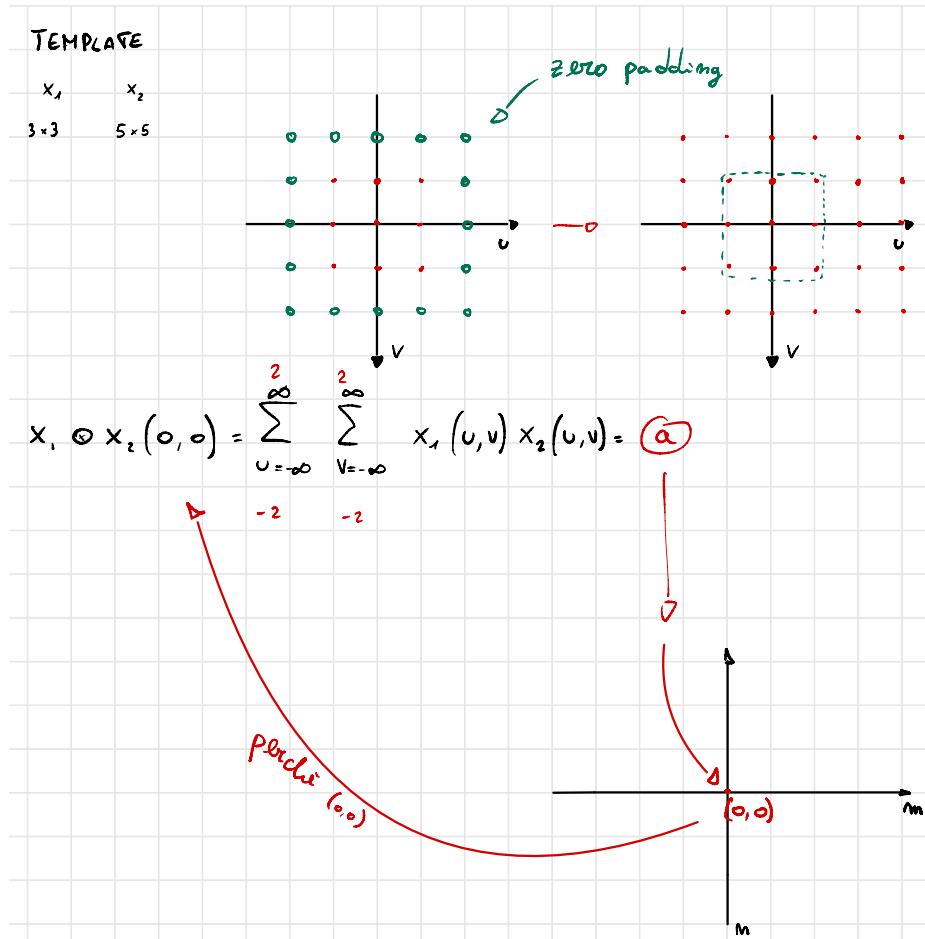


Figura 17: Esempio di Cross-Correlazione normalizzata 2D.

Esercizio Cross-Correlazione 2D

Dati le due immagini x_1 di dimensione 5×5 e x_2 di dimensione 3×3 , si calcola la cross-correlazione 2D. Quindi, si effettua la rappresentazione grafica.

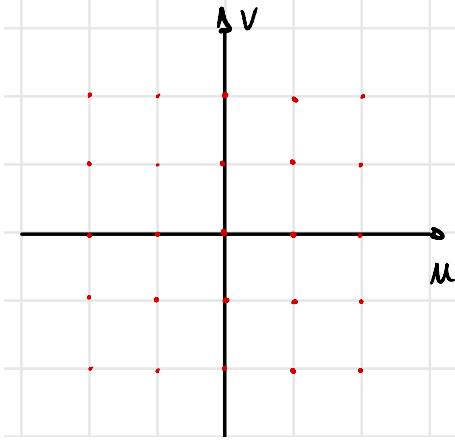


Figura 18: Piano cartesiano di x_2 di dimensione 5×5 .

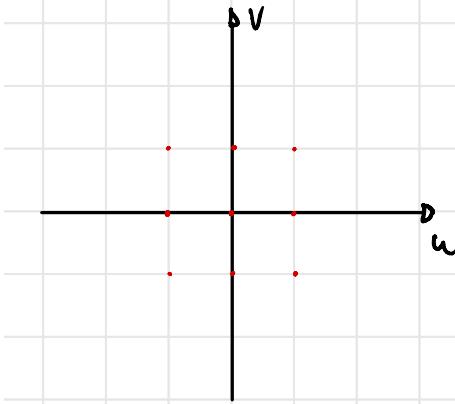


Figura 19: Piano cartesiano di x_1 di dimensione 3×3 .

E vengono fornite dall'esercizio le due matrici:

$$x_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad x_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Esse indicano i valori nei punti corrispondenti. L'**obiettivo dell'esercizio** è trovare:

- L'argomento massimo della cross-correlazione ($\arg \max x_1 \otimes x_2 (m, n)$);
- Il massimo della cross-correlazione ($\max x_1 \otimes x_2 (m, n)$).

L'argomento massimo è con i valori $m = 1$ e $n = -1$ poiché così facendo la diagonale incontra tutti i valori positivi e che formano il massimo. Infatti, prendendo in considerazione la matrice x_2 5×5 e osservando l'operazione di cross-correlazione 2D:

$$\sum_u \sum_v x_1(u, v) \cdot x_2(u - m, v - n)$$

$$\xrightarrow{\text{sostituzione termini noti } (m,n)} \sum_u \sum_v x_1(u, v) \cdot x_2(u - 1, v - (-1))$$

Risulta evidente come si debba spostare a destra, rispetto l'origine, la matrice x_2 di un solo valore¹ e sotto, rispetto sempre l'origine, di un valore negativo². Così facendo, la diagonale della matrice x_2 corrisponderà esattamente a tutti i valori 1 della matrice x_1 .

¹Shift a destra poiché $u - 1$ nell'equazione rappresenta un ritardo.

²Spostamento sotto l'asse delle ascisse poiché è un valore positivo $v + 1$.

1.3.5 Convoluzione

La **convoluzione** è un parente stretto della cross-correlazione, ma è leggermente diverso. È definito nel seguente modo:

$$f_1 * f_2(t) = \int_{-\infty}^{+\infty} f_1(\tau) f_2(t - \tau) d\tau$$

Con $t \in \mathbb{R}$. Si ricordi che se i **segnali** non sono né di energia né di potenza, l'integrale converge.

Nel caso in cui i **segnali** siano **discreti**, dati $x_1(n)$, $x_2(n)$:

$$x_1 * x_2(n) = \sum_{k=-\infty}^{+\infty} x_1(k) x_2(n - k)$$

Con $k \in \mathbb{Z}$.

Nel caso in cui $x_1(n)$ e $x_2(n)$ sono limitati di lunghezza M ed N rispettivamente, allora la **convoluzione** è di lunghezza $M + N - 1$.

Convoluzione 2D

Nel caso delle immagini, quindi del 2D, x_1 ed x_2 sono solitamente **segnali digitali ad intervallo limitato**, e la convoluzione diventa dunque:

$$x_1 * x_2(m, n) = \sum_{u=-\infty}^{+\infty} \sum_{v=-\infty}^{+\infty} x_1(u, v) x_2(m - u, n - v) \quad u, v, m, n \in \mathbb{Z}$$

Solitamente il primo segnale x_1 viene chiamato **filtro**, o **matrice kernel**, mentre x_2 genericamente **immagine** (solitamente la matrice kernel ha una dimensione inferiore di quella dell'immagine).

2 Analisi di Fourier

2.1 Serie di Fourier

Una funzione, chiamata **funzione di sintesi**, $f : \mathbb{R} \rightarrow \mathbb{R}$ di variabile continua t , periodica di periodo T , si esprime come:

$$f(t) = \sum_{n=-\infty}^{+\infty} c_n \underbrace{e^{j \frac{2\pi n}{T} t}}_{\text{fasore}} \quad n \in \mathbb{Z}$$

Dove c_n è un numero complesso. Invece, una **funzione di analisi** è espressa come:

$$c_n \in \mathbb{C} = \frac{1}{T} \int_{-\frac{T}{2}}^{+\frac{T}{2}} f(t) \underbrace{e^{-j \frac{2\pi n}{T} t}}_{\text{fasore}} dt \quad n \in \mathbb{Z}$$

N.B. si ricorda che $e^{j \frac{2\pi n}{T} t}$ è un **fasore rotante** di velocità angolare $\frac{2\pi n}{T} t$.

La **funzione di sintesi** quindi non è altro che una somma di infiniti termini. Ciascuno è composto dalla moltiplicazione tra un numero complesso ed un fasore, il quale *produce un altro fasore*. Esprimendo c_n come numero complesso in forma polare:

$$c_n e^{j \frac{2\pi n}{T} t} = |c_n| e^{j \theta_n} e^{j \frac{2\pi n}{T} t} = |c_n| e^{j(\frac{2\pi n}{T} t + \theta_n)}$$

Si può notare come questa conversione corrisponda ad **estendere** il fasore $e^{j \frac{2\pi n}{T} t}$ ad una lunghezza $|c_n|$ facendolo partire con un **angolo di partenza** uguale a θ_n (chiamato **angolo di fase**).

Altra osservazione: se c_n appartiene all'insieme \mathbb{R} , significa che θ_n non compare. Questo comporta un cambiamento nella lunghezza dell' n -esimo fasore pari a $|c_n|$:

$$c_n = |c_n| e^{j \theta_n}$$

Esempio 1

Il primo esempio di serie di Fourier si applica per il segnale trigonometrico:

$$f(t) = \cos(2\pi t) \quad \text{con } T = 1$$

Applicando la **funzione di analisi** e saltando i passaggi perché complessi, si ottengono i seguenti valori:

$$c_{-1} = \frac{1}{2} \quad c_0 = 0 \quad c_1 = \frac{1}{2} \quad c_{i \leq -2, i \geq 2} = 0$$

E sostituendo nella **funzione di sintesi**:

$$\cos(2\pi t) = \frac{1}{2}e^{-j2\pi t} + \frac{1}{2}e^{j2\pi t} = \frac{e^{j2\pi t} + e^{-j2\pi t}}{2}$$

Ci sono **tre osservazioni** da fare:

I. $\frac{2\pi}{T} = f_0$;

II. $c_n = |c_n|e^{j\theta_n}$;

III. In questo caso, $c_n \in \mathbb{R}$ quindi l'angolo di fase non è presente.

Le parti dell'equazione sono le seguenti:

$$\cos(2\pi t) = \frac{1}{2}e^{-j2\pi t} + \frac{1}{2}e^{j2\pi t}$$

☞ $\cos(2\pi t)$ → La funzione trigonometrica da studiare

☞ $\frac{1}{2}e^{-j2\pi t}$ → Fasore di modulo 0.5 e velocità angolare $-2\pi t$

☞ $\frac{1}{2}e^{j2\pi t}$ → Fasore di modulo 0.5 e velocità angolare $2\pi t$

I coefficienti $c_{n=-1}$ e $c_{n=1}$ sono relativi ai **moduli o ampiezze dei fasori** complessi di frequenza $f_0 \cdot n$ con $n = -1, 1$ e ricordando che:

$$\exp\left(j\left(\frac{2\pi n}{T}t\right)\right) = \exp(j(f_0 nt))$$

Che si possono annotare con le variabili f_{-1} e f_1 per $f_0 \cdot n$ con $n = -1, 1$ e analogamente per gli altri $n \in \mathbb{Z}$.

Inoltre, è possibile disegnare lo **spettro di ampiezza** che **mostra i moduli dei fasori costruiti con la trasformata di Fourier**, in particolare la funzione di sintesi.

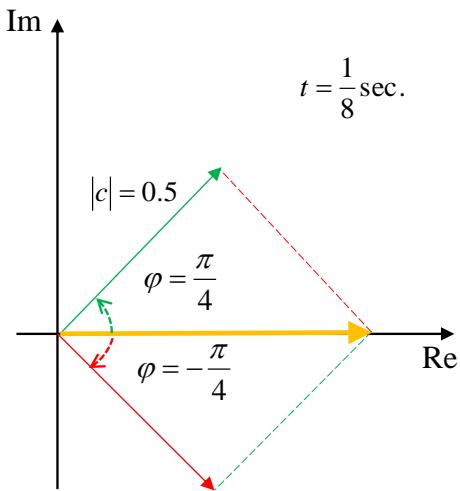


Figura 20: Grafico rappresentante i due fasori. La freccia verde rappresenta il valore assunto da $\cos(2\pi t)$ per $t = \frac{1}{8}$.

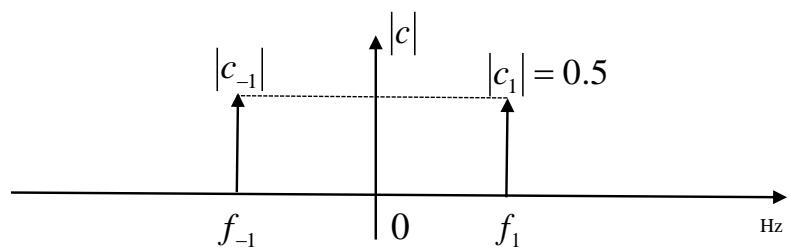


Figura 21: Grafico che rappresenta lo spettro di ampiezza.

Esempio 2

Il secondo esempio di serie di Fourier è il segnale trigonometrico:

$$f(t) = \sin(2\pi t) \quad \text{con } T = 1$$

Applicando la **funzione di analisi** e saltando i passaggi perché complessi, si ottengono i seguenti valori:

$$c_{-1} = -\frac{1}{2j} \quad c_0 = 0 \quad c_1 = \frac{1}{2j} \quad c_{i \leq -2, i \geq 2} = 0$$

Dove questa volta $c_n \in \mathbb{C}$ ed in particolare:

$$\pm \frac{1}{2j} = \pm \frac{1}{2j} \cdot \frac{j}{j} = \pm \frac{1}{2} \cdot \frac{j}{j^2} = j \cdot \mp \frac{1}{2}$$

Si passa alla forma di esponenziale complesso:

$$\begin{aligned} j \cdot \frac{1}{2} &= 0 + j \cdot \frac{1}{2} \\ |c| &= \sqrt{0^2 + \left(\frac{1}{2}\right)^2} = \frac{1}{2} \\ \theta &= \arctan\left(\frac{0.5}{0}\right) \rightarrow \frac{\pi}{2} \\ \frac{1}{2}e^{j \cdot \frac{\pi}{2}} &= c_{-1} \end{aligned}$$

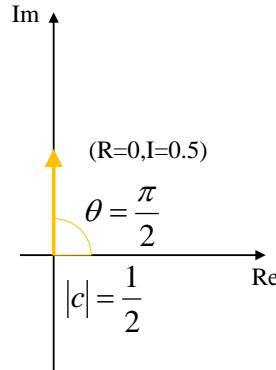


Figura 22: Grafico di c_{-1} .

Analogamente:

$$\begin{aligned}j \cdot -\frac{1}{2} &= 0 + j \cdot \left(-\frac{1}{2}\right) \\|c| &= \sqrt{0^2 + \left(-\frac{1}{2}\right)^2} = \frac{1}{2} \\\theta &= \arctan\left(-\frac{0.5}{0}\right) \rightarrow -\frac{\pi}{2} \\\frac{1}{2}e^{j \cdot \left(-\frac{\pi}{2}\right)} &= c_1\end{aligned}$$

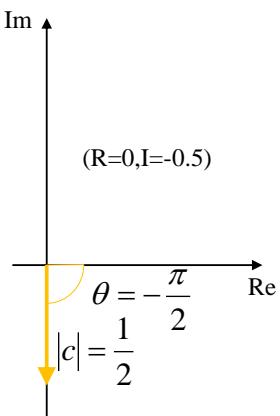


Figura 23: Grafico di c_1 .

Applicando l'**equazione di sintesi** e sostituendo i termini noti:

$$\begin{aligned}\sin(2\pi t) &= \sum_{n=-\infty}^{+\infty} c_n e^{j \frac{2\pi n}{T} t} = c_{-1} e^{j -2\pi t} + c_1 e^{j 2\pi t} \\ \text{sostituzione dei termini noti } c_{-1}, c_1 &\Rightarrow \frac{1}{2} e^{j \frac{\pi}{2}} e^{j \cdot (-2\pi t)} + \frac{1}{2} e^{j -\frac{\pi}{2}} e^{j 2\pi t} \\ \text{forma finale} &\Rightarrow \frac{1}{2} \exp\left(j\left(-2\pi t + \frac{\pi}{2}\right)\right) + \exp\left(j\left(2\pi t - \frac{\pi}{2}\right)\right)\end{aligned}$$

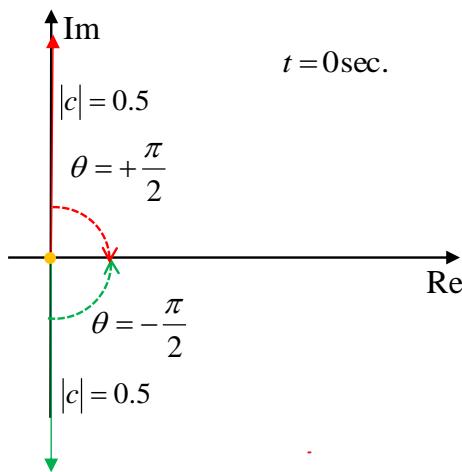


Figura 24: Grafico finale.

Infine, si disegna lo **spettro di ampiezza** e lo **spettro di fase**, quest'ultimo è un **grafico in cui si riportano gli angoli di fase della funzione**.

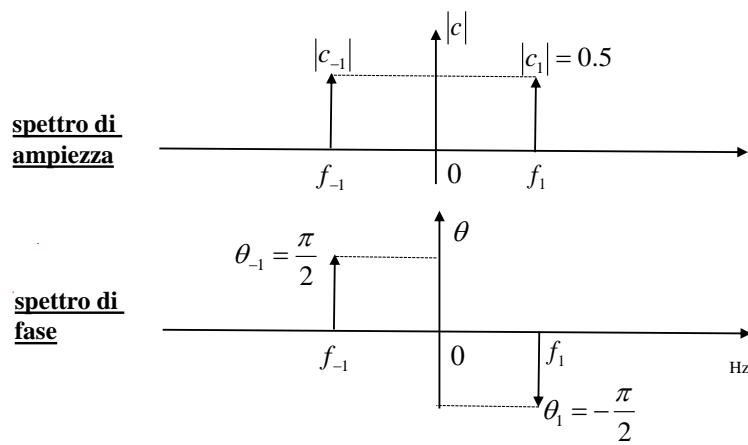


Figura 25: Spettro di ampiezza e di fase della funzione $\sin(2\pi t)$.

2.1.1 Proprietà della serie di Fourier

Lo **spettro di ampiezza e di fase** sono funzioni nel dominio delle frequenze che formano lo **spettro di Fourier**. Lo spettro di Fourier per i segnali periodici gode delle **seguenti proprietà**:

- Lo **spettro di ampiezza** è *simmetrico* rispetto all'asse y ;
- Lo **spettro di fase** è *antisimmetrico* rispetto all'asse y ;
- Se i coefficienti c_n sono reali, **non esiste lo spettro di fase**;
- Entrambe gli spettri sono funzioni a pettine, definite su frequenze multiple rispetto a quella fondamentale:

$$\left\{ \frac{2\pi n}{T} \right\}_{n \in \mathbb{Z}} = \{f_0 \cdot n\}_{n \in \mathbb{Z}} \equiv \{f_n\}_{n \in \mathbb{Z}}$$

2.2 Trasformata di Fourier continua

2.2.1 Trasformata di Fourier

Sia $f(t)$ un **segnale reale continuo** $f : \mathbb{R} \rightarrow \mathbb{R}$ periodico o non, si definisce la **Trasformata di Fourier** (TdF) $\mathcal{F}(f(t)) = F(\mu)$ il segnale $\mathcal{F} : \mathbb{R} \rightarrow \mathbb{C}$:

$$\mathcal{F}(f(t)) = F(\mu) = \int_{-\infty}^{+\infty} f(t) e^{-j2\pi\mu t} dt$$

L'**unità frequenziale** μ è l'angolo di $\frac{n}{T}$ della serie di Fourier (per esempio, con $n = 1$, $T = 1$ sec. $\rightarrow \mu = 1$ sec. $^{-1}$ = 1 Hz).

La Trasformata di Fourier esiste se $f(t)$ è un **segnale di energia**. Condizione sufficiente e non necessaria perché altri segnali ammettono la TdF.

2.2.2 Trasformata di Fourier inversa

Sia $F(\mu)$ la trasformata di Fourier di un segnale $f : \mathbb{R} \rightarrow \mathbb{R}$. Si definisce la **trasformata di Fourier inversa** il segnale $\mathcal{F}^{-1}(F(\mu)) = f(t)$:

$$\mathcal{F}^{-1} = (F(\mu)) = f(t) = \int_{-\infty}^{+\infty} F(\mu) e^{j2\pi\mu t} d\mu$$

La trasformata di Fourier restituisce, per una data frequenza μ , un coefficiente di “presenza” $F(\mu)$. Infatti, la sua inversa permette di ricostruire f a partire da F .

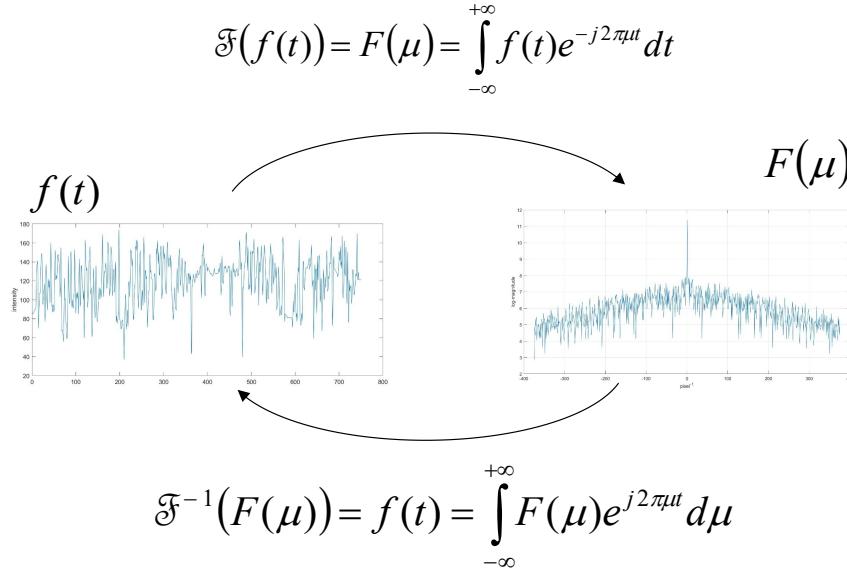


Figura 26: (Anti)Trasformata di Fourier su un segnale $f(t)$.

Nel caso in cui il segnale $f(t)$ non è reale, la trasformata è complessa:

- t rappresenta il **tempo** (in secondi), allora μ rappresenta gli **Hertz**, cioè $\frac{\text{numero cicli}}{\text{secondi}}$;
- t rappresenta lo **spazio** (in metri), allora μ rappresenta la **frequenza spaziale**, cioè $\frac{\text{numero cicli}}{\text{metri}}$

Mentre nella serie di Fourier le funzioni rappresentate negli spettri di ampiezza e di fase erano a “pettine” (paragrafo 2.1.1), in questo caso le funzioni sono solitamente continue, nello spettro di ampiezza, o continue a tratti:

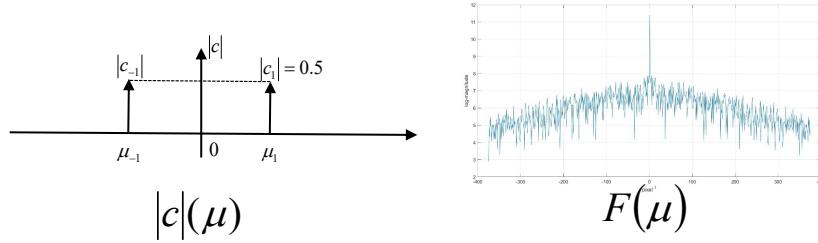


Figura 27: Esempio di spettro di ampiezza.

2.2.3 Proprietà della trasformata di Fourier

☞ Linearità

$$a_1 f_1(t) + a_2 f_2(t) \xrightarrow{\mathcal{F}} a_1 F_1(\mu) + a_2 F_2(\mu)$$

☞ Scalatura temporale

$$z(t) = f(at) \xrightarrow{\mathcal{F}} Z(\mu) = \frac{1}{a} F\left(\frac{\mu}{a}\right)$$

☞ Dualità

$$\begin{aligned} f(t) &\xrightarrow{\mathcal{F}} F(\mu) \\ F(t) &\xrightarrow{\mathcal{F}^-} f(-\mu) \end{aligned}$$

N.B. derivando la forma analitica per una trasformata, la sua antitrasformata ne produce un'altra con segno opposto.

☞ Time shift

$$\begin{aligned} \mathcal{F}(f(t - t_0)) &= \int_{-\infty}^{+\infty} f(t - t_0) e^{-j2\pi\mu t} dt \\ &= \int_{-\infty}^{+\infty} f(u) e^{-j2\pi\mu(u+t_0)} du \\ &= \int_{-\infty}^{+\infty} f(u) e^{-j2\pi\mu u} e^{-j2\pi\mu t_0} du \\ &= e^{-j2\pi\mu t_0} \int_{-\infty}^{+\infty} f(u) e^{-j2\pi\mu u} du \\ &= F(\mu) \underbrace{e^{-j2\pi\mu t_0}}_{\text{fase}} \end{aligned}$$

2.2.4 Trasformata di Fourier di una box

La trasformata di Fourier di una box (paragrafo 1.2.3) è la seguente:

$$\mathcal{F}(f(t)) = \int_{-\infty}^{+\infty} A \Pi\left(\frac{t}{w}\right) e^{-j2\pi\mu t} dt = F(\mu)$$

Il risultato corrisponde alla funzione sinc:

$$f(\mu) = Aw \cdot \text{sinc}(\mu w)$$

Dove la funzione sinc è uguale a:

$$\text{sinc} = \frac{\sin(\pi\mu w)}{\pi\mu w}$$

Per ripassare la funzione sinc, si rimanda al paragrafo 1.2.4. Tuttavia, si ricorda che la sua forma generale è del tipo:

$$\text{sinc}(m) = \frac{\sin(\pi m)}{\pi m}$$

E risultata uguale a:

- $\text{sinc}(0) = 1$
- $\text{sinc}(m) = 0 \quad \forall m \in \mathbb{Z}$

Prima di concludere, si ricorda che:

- ☞ All'aumentare della larghezza della box, la funzione sinc tenderà a stringersi;
- ☞ La box è **limita**, invece la sinc è **infinita** a destra e sinistra, anche se il termine al denominatore attenua il valore della funzione comportando un limite a 0.
- ☞ In sintesi, la TdF di una box è:

$$\Pi\left(\frac{t}{w}\right) \xrightarrow{\mathcal{F}} w \cdot \text{sinc}(\mu w)$$

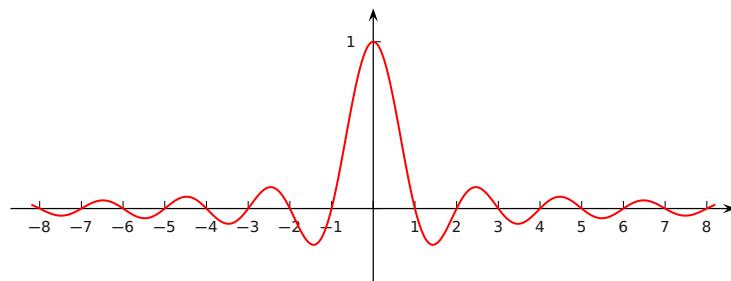


Figura 28: Grafico della funzione sinc.

2.2.5 Trasformata di Fourier di un sinc

La trasformata di Fourier di un segnale sinc (segnale rappresentato in figura 28) è la seguente:

$$\mathcal{F}(f(t)) = \int_{-\infty}^{+\infty} \text{sinc}(tw) e^{-j2\pi\mu t} dt = F(\mu)$$

Dato che la TdF di una box è:

$$\Pi\left(\frac{t}{w}\right) \xrightarrow{\mathcal{F}} w \cdot \text{sinc}(\mu w)$$

Al contrario, si ottiene la trasformata di Fourier di un sinc:

$$\text{sinc}(tw) \xrightarrow{\mathcal{F}} \frac{1}{w} \Pi\left(-\frac{\mu}{w}\right) = \frac{1}{w} \cdot \Pi\left(\frac{\mu}{w}\right)$$

2.2.6 Trasformata di Fourier di un impulso

La trasformata di Fourier di un impulso³ è la seguente:

$$\mathcal{F}(f(t)) = F(\mu) = \int_{-\infty}^{+\infty} \delta(t) e^{-j2\pi\mu t} dt$$

Il risultato della trasformata di Fourier di un impulso è molto semplice grazie alle sue proprietà. Infatti, il risultato è uguale a:

$$\int_{-\infty}^{+\infty} \delta(t) e^{-j2\pi\mu t} dt = \int_{-\infty}^{+\infty} \delta(0) e^{-j2\pi\mu 0} dt = 1$$

La proprietà che consente di ottenere il risultato uguale a 1 è la seguente:

$$\delta(t) = \begin{cases} \infty & \text{se } t = 0 \\ 0 & \text{se } t \neq 0 \end{cases} \quad \rightarrow \quad \int_{-\infty}^{+\infty} \delta(t) dt = 1$$

N.B. In questo caso è rappresentabile solo lo spettro di ampiezza!

Analogamente, con un impulso centrato in t_0 , quindi non nell'origine:

$$\mathcal{F}(f(t)) = F(\mu) = \int_{-\infty}^{+\infty} \delta(t - t_0) e^{-j2\pi\mu t} dt = e^{-j2\pi\mu t_0}$$

Il risultato è stato ottenuto grazie alla proprietà di setacciamento (definita a pagina 11). Tuttavia, in questo caso i valori non sono più reali ma complessi.

³Definizione di impulso al paragrafo 1.2.3.

2.2.7 Trasformata di Fourier di un treno di impulsi

Data la definizione di treno di impulsi (funzione definita nel paragrafo 1.2.8):

$$S_{\Delta T}(t) = \sum_{n=-\infty}^{+\infty} \delta(t - n\Delta T) \quad n \in \mathbb{Z}$$

Si ottiene la sua relativa trasformata di Fourier:

$$\mathcal{F}(S_{\Delta T}(t)) = \int_{-\infty}^{+\infty} S_{\Delta T}(t) e^{-j2\pi\mu t} dt = F(\mu)$$

Tralasciando i vari calcoli numerici per arrivare al risultato, si può scrivere la trasformata di Fourier in maniera più semplice:

$$S_{\Delta T}(t) \xrightarrow{\mathcal{F}} \sum_{n=-\infty}^{+\infty} \frac{1}{\Delta T} \delta\left(\mu - \frac{n}{\Delta T}\right)$$

2.2.8 Sintesi

Qui di seguito si lascia un riassunto rapido delle trasformate di Fourier continue dei segnali più importanti.

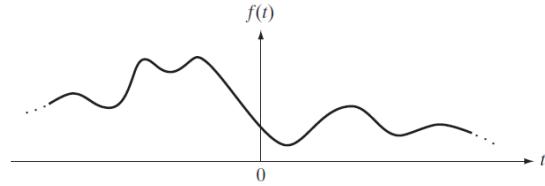
Segnale	Trasformata di Fourier	
Box:	$A\Pi\left(\frac{t}{w}\right)$	$\xrightarrow{\mathcal{F}}$ $Aw \cdot \text{sinc}(\mu w)$
Sinc:	$\text{sinc}(tw)$	$\xrightarrow{\mathcal{F}}$ $\frac{1}{w} \cdot \Pi\left(-\frac{\mu}{w}\right) = \frac{1}{w} \cdot \Pi\left(\frac{\mu}{w}\right)$
Impulso:	$\delta(t)$	$\xrightarrow{\mathcal{F}}$ $\begin{cases} 1 & \text{se valori reali} \\ e^{-j2\pi\mu t_0} & \text{se valori complessi} \end{cases}$
Treno di impulsi:	$S_{\Delta T}(t)$	$\xrightarrow{\mathcal{F}}$ $\sum_{n=-\infty}^{+\infty} \frac{1}{\Delta T} \cdot \delta\left(\mu - \frac{n}{\Delta T}\right)$

Tabella 1: Trasformate di Fourier continue.

2.3 Trasformata di Fourier a tempo discreto

2.3.1 Campionamento

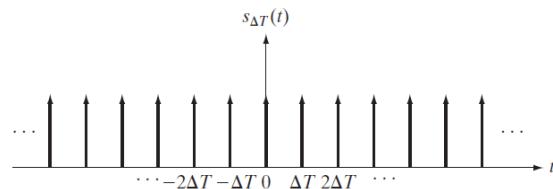
Sia $f(t)$ un segnale reale continuo definito $f :]-\infty, +\infty[\in \mathbb{R} \rightarrow \mathbb{R}$ (attenzione al dominio non limitato), anche non periodico:



Questo tipo di segnale, per essere elaborato al computer deve essere **campionato** ad intervalli discreti. Per farlo, si prenda in considerazione il treno di impulsi:

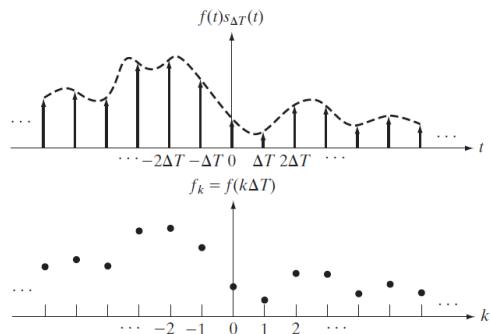
$$S_{\Delta T}(t) = \sum_{n=-\infty}^{+\infty} \delta(t - n\Delta T)$$

Con periodo ΔT , ossia con **frequenza di campionamento** pari a: $\mu_S = \frac{1}{\Delta T}$



Si assume che il treno di impulsi sia un **segnale discreto**. Matematicamente parlando, **campionare un segnale** significa moltiplicarlo per un treno di impulsi:

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



2.3.2 Trasformata di Fourier a tempo discreto

Sia $F(\mu)$ la trasformata di Fourier di un segnale $f(t) : \mathbb{R} \rightarrow \mathbb{R}$. Si considera $\tilde{f}(t)$ e si calcola la trasformata di Fourier $\tilde{F}(\mu)$ (entrambi sono a tempo discreto). Grazie alla convoluzione si ottiene:

$$\tilde{F}(\mu) = \mathcal{F}\left\{\tilde{f}(t)\right\} = F(\mu) * S_{\Delta T}(\mu)$$

Si ricorda che:

$$S_{\Delta T}(\mu) = \frac{1}{\Delta T} \sum_{n=-\infty}^{+\infty} \delta\left(\mu - \frac{n}{\Delta T}\right)$$

E risolvendo la convoluzione, si ottiene che la **trasformata di Fourier a tempo discreto** corrisponde a:

$$F(\mu) * S_{\Delta T}(\mu) = \int_{-\infty}^{+\infty} F(\tau) \cdot S_{\Delta T}(\mu - \tau) d\tau = \frac{1}{\Delta T} \sum_{n=-\infty}^{+\infty} F\left(\mu - \frac{n}{\Delta T}\right)$$

Analizzando la formula si evidenziano alcuni termini:

- $F(\mu)$ è la trasformata di Fourier della funzione originale $f(t)$;
- $F\left(\mu - \frac{n}{\Delta T}\right)$ è la trasformata di Fourier della funzione originale $f(t)$ shiftato a destra di una quantità pari a $\frac{n}{\Delta T}$;
- $\frac{1}{\Delta T} \sum_{n=-\infty}^{+\infty} F\left(\mu - \frac{n}{\Delta T}\right)$ sono infinite copie dello spettro $F(\mu)$, ripetute ovviamente ogni $\frac{1}{\Delta T}$.

Inoltre, è un **segnale periodico** (nelle frequenze) di periodo $\frac{1}{\Delta T}$, ovvero si ripete ogni $\frac{1}{\Delta T}$ Hz.

La sua scalatura nell'ampiezza è pari a $\frac{1}{\Delta T}$ e rappresenta la T.d.F. a tempo discreto.

2.3.3 Teorema del campionamento

Un segna~~le reale continuo $f(t)$~~ , limitato in banda, può essere ricostruito senza errori completamente dai suoi campioni se essi sono acquisiti con un tempo di campionamento ΔT tale per cui:

$$\frac{1}{\Delta T} = \mu_S \geq 2\mu_{\max}$$

Ovvero se nel tempo si adotta una frequenza di campionamento $\frac{1}{\Delta T}$ almeno doppia rispetto alla frequenza massima del segnale μ_{\max} .

In altre parole, il teorema del campionamento afferma che tutte le proprietà di un segnale possono essere espresse usando dei campioni.

Attenzione! L'espressione $\frac{1}{\Delta T}$ viene chiamata *frequenza di Nyquist* e per frequenze minori si crea aliasing, fenomeno che impedisce la ricostruzione senza errori.

2.3.4 Considerazioni

Dal punto di vista teorico la trasformata di Fourier a tempo discreto consente di ricostruire il segnale. Tuttavia, è impossibile implementarla in un computer poiché tende, come limiti, all'infinito e ci vorrebbe un numero infinito di campioni e di segnali di tipo sinc.

2.4 Trasformata di Fourier discreta

La trasformata di Fourier di un segnale reale continuo $f(t)$ di dominio illimitato e non periodico, campionato nel tempo con periodo di campionamento ΔT , è una funzione continua, periodica (di periodo $\frac{1}{\Delta T}$) anch'essa di dominio illimitato:

$$\tilde{f}(t) = f(t) \cdot \sum_{n=-\infty}^{+\infty} \delta(t - n\Delta T) \xrightarrow{\text{F}} \tilde{F}(\mu) = \frac{1}{\Delta T} \sum_{n=-\infty}^{+\infty} F\left(\mu - \frac{n}{\Delta T}\right)$$

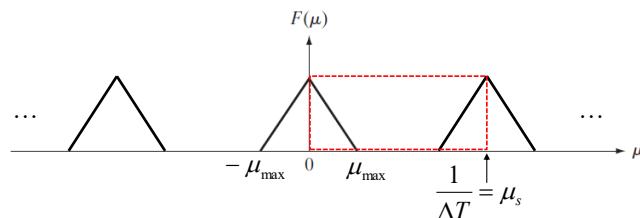
Il **problema** di questa formulazione è data l'espressione analitica dello spettro, la quale suppone che si è a **conoscenza della T.d.F. teorica F del segnale di partenza**. Questo, spesso, è molto difficile.

Si ricava dunque una *forma più semplice* da manipolare. Essa consente di **costruire una rappresentazione spettrale a partire dai campioni della funzione originale $f(t)$** :

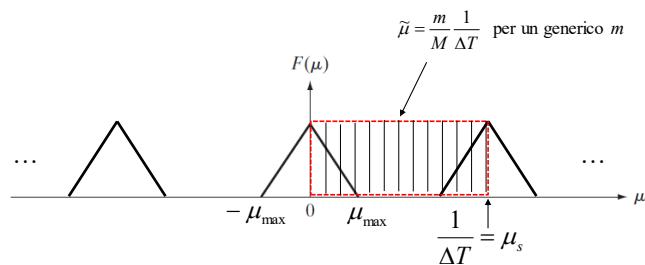
$$\tilde{F}(\mu) = \sum_{n=-\infty}^{+\infty} f_n e^{-j2\pi\mu n \Delta T}$$

Al contrario, la prima formulazione era più chiara per comprendere il fatto che la T.d.F. di una funzione campionata genera delle repliche dello spettro originale $F(\mu)$.

L'equazione alternativa deve essere modificata, eseguendo un campionamento per il dominio spettrale, per poterla implementare su un computer. Per farlo, si prende in considerazione solo l'intervallo frequenziale da 0 a $\frac{1}{\Delta T} = \mu_s$.



Inoltre, si prendono in considerazione M campioni tramite l'operazione di campionamento.



In cui:

$$\tilde{\mu} = \frac{m}{M} \cdot \frac{1}{\Delta T} \quad \text{con } m = 0, \dots, M - 1 \text{ e dove } \frac{m}{M} \in \left[0, 1 - \frac{1}{M}\right]$$

La m indica il **range di variazione** degli indici dei campioni frequenziali. Calcolando la trasformata di Fourier a tempo discreto sui campioni M , si giunge alla **forma finale della trasformata di Fourier discreta**:

$$\tilde{F}(\tilde{\mu}) = \tilde{F}\left(\frac{m}{M} \cdot \frac{1}{\Delta T}\right) = \sum_{n=0}^{M-1} f_n e^{-j2\pi \frac{m}{M} n} \quad \text{con } m = 0, \dots, M - 1$$

La **trasformata di Fourier discreta inversa**, ovvero l'antitrasformata:

$$\tilde{f}(n\Delta T) = f(n\Delta T) = f_n = \frac{1}{M} \sum_{m=0}^{M-1} F_m e^{j2\pi \frac{m}{M} n}$$

2.5 Riassunto Trasformate

Qui di seguito vengono rappresentate le trasformate più importanti.

Funzione	Serie di Fourier
Funzione di sintesi:	$f(t) \sum_{n=-\infty}^{+\infty} c_n \underbrace{e^{j \frac{2\pi n}{T} t}}_{\text{fasore}}$ $n \in \mathbb{Z}$
Funzione di analisi:	$c_n \in \mathbb{C} \quad \frac{1}{T} \int_{-\frac{T}{2}}^{+\frac{T}{2}} f(t) \underbrace{e^{-j \frac{2\pi n}{T} t}}_{\text{fasore}} dt \quad n \in \mathbb{Z}$
Segnale	Trasformata di Fourier continua
Box:	$A \Pi\left(\frac{t}{w}\right) \xrightarrow{\mathcal{F}} Aw \cdot \text{sinc}(\mu w)$
Sinc:	$\text{sinc}(tw) \xrightarrow{\mathcal{F}} \frac{1}{w} \cdot \Pi\left(-\frac{\mu}{w}\right) = \frac{1}{w} \cdot \Pi\left(\frac{\mu}{w}\right)$
Impulso:	$\delta(t) \xrightarrow{\mathcal{F}} \begin{cases} 1 & \text{se valori reali} \\ e^{-j2\pi\mu t_0} & \text{se valori complessi} \end{cases}$
Treno di impulsi:	$S_{\Delta T}(t) \xrightarrow{\mathcal{F}} \sum_{n=-\infty}^{+\infty} \frac{1}{\Delta T} \cdot \delta\left(\mu - \frac{n}{\Delta T}\right)$
Funzione	Trasformata di Fourier a tempo discreto
$F(\mu) * S_{\Delta T}(\mu)$	$\frac{1}{\Delta T} \sum_{n=-\infty}^{+\infty} F\left(\mu - \frac{n}{\Delta T}\right)$
Funzione	Trasformata di Fourier discreta
$\tilde{F}(\tilde{\mu})$	$\tilde{F}\left(\frac{m}{M} \cdot \frac{1}{\Delta T}\right) = \sum_{n=0}^{M-1} f_n e^{-j2\pi \frac{m}{M} n}$
	$m = 0, \dots, M-1$
Funzione	Trasformata di Fourier discreta inversa
$\tilde{f}(n\Delta T) = f(n\Delta T) = f_n$	$\frac{1}{M} \sum_{m=0}^{M-1} F_m e^{j2\pi \frac{m}{M} n}$

Tabella 2: Trasformate di Fourier.

2.6 Domanda da esame

All'esame è possibile che sia richiesto come domanda: "quali sono le trasformate di Fourier studiate durante il corso?"

La risposta, anche se banale, è la seguente: le trasformate di Fourier studiate durante il corso sono 4:

- I. Serie di Fourier (paragrafo 2.1)
- II. Trasformata di Fourier continua (paragrafo 2.2)
- III. Trasformata di Fourier a tempo discreto (paragrafo 2.3)
- IV. Trasformata di Fourier discreta (paragrafo 2.4)

3 Elaborazione di immagini - Dominio spaziale

L'**elaborazione delle immagini** consiste nel prendere come input un'immagine (segnalet) e restituirne un'altra (sempre segnale) come output.

Il **rinforzo (enhancement) di immagini** è un tipo di elaborazione delle immagini. Il suo **obbiettivo** è elaborare un'immagine in modo che il risultato sia più adatto alle esigenze soggettive dell'utente. La definizione è *problem-oriented* poiché, per esempio, per visualizzare lo spettro di ampiezza di un'immagine, è necessario eseguire un'operazione di rinforzo (*log-transformation*); oppure, per migliorare la visibilità dei dettagli di un'immagine, si effettua un'altra operazione di rinforzo (*sharpening*).

La **qualità di un'immagine** è la combinazione pesata di tutti gli attributi significativi di un'immagine. Infatti, *non* esiste una ricetta univoca per determinare quando un'immagine sia di qualità poiché è un'opinione soggettiva. Tuttavia, è più facile dire quando un'immagine *non* è di qualità. In genere, un'**immagine non è di qualità** quando non viene interpretata facilmente da un operatore umano.

A differenza del rinforzo, il **restauro (restoration)** è un processo di ricostruzione dell'immagine a partire da un modello di degradazione noto.

3.1 Strumento per l'elaborazione: istogramma

I pixel di un'immagine sono una “popolazione” sulla quale è possibile calcolare tutte le quantità statistiche descrittive che vengono usate normalmente come media, mediana, varianza, deviazione standard, quartili, percentili, etc. Uno **strumento fondamentale** per l'elaborazione delle immagini è l'**istogramma**, il quale può essere visto come una funzione continua o discreta.

Infatti, per ogni livello di grigio (in un'immagine solo a livelli di grigi) vengono riportati il numero di pixel. Per un'immagine $I [M, N]$ si identifica con M, N il **numero di pixel righe per colonne** e con la funzione $H(r)$ il **numero di pixel di valore r** , quest'ultimo è definito nell'intervallo $0 \leq r \leq L - 1$ con $r, L \in \mathbb{N}$ dove L indica i livelli di grigio. Inoltre:

$$\sum_{r=0}^{L-1} H(r) = M \cdot N$$

Grazie all'istogramma, è possibile comprendere immediatamente le caratteristiche dell'immagine (come in figura).

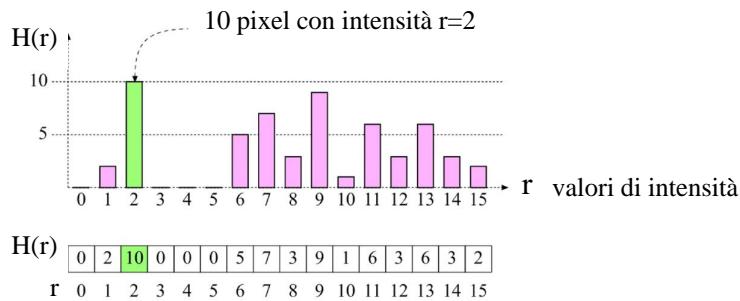


Figura 29: Esempio di istogramma.

Un istogramma può essere anche visto come una distribuzione di probabilità:

$$p_h(r) = \frac{H(r)}{M \cdot N} \quad \sum_r p_h(r) = 1$$

Uno **svantaggio** di questo strumento è che immagini diverse potrebbero avere istogrammi simili, questo perché l'istogramma non tiene conto della distribuzione spaziale dei pixel. Dunque, **utilizzando solo questo metodo è impossibile ricostruire un'immagine**.

Al contrario, un **vantaggio** dell'istogramma è la possibilità di identificare facilmente il **contrasto**: rapporto o differenza tra il valore più alto (punto più luminoso) e il valore più basso (punto più scuro) della luminosità (che corrisponde al livello di grigio per le immagini a livello di grigio).

Un'immagine viene definita:

- Valori più alti sulla destra:
 - **chiara**, caratteristica dell'immagine;

- **sovraesposta**, caratteristica di come è stata acquisita l'immagine.
- Valori più alti sulla sinistra:
 - **scura**, caratteristica dell'immagine;
 - **sottoesposta**, caratteristica di come è stata acquisita l'immagine.

3.2 Domini

L'elaborazione delle immagini può avvenire nel **dominio spaziale** o nel **dominio frequenziale** (dopo aver applicato la T.d.F. discreta 2D). Nel **dominio spaziale**, l'elaborazione delle immagine può essere espressa come:

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

In cui:

- f è l'immagine di ingresso (input) da elaborare;
- g è l'immagine d'uscita (output) elaborata;
- T è un operatore su f definito in un intorno di (x, y) .

L'**operatore** definito in un intorno di (x, y) può essere di tre tipi:

- **Puntuale:** $[f(x, y)] = f(x, y)$, l'intorno coincide con il pixel stesso;
- **Locale:** $[f(x, y)]$ rappresenta una regione, per esempio quadrata, attorno al pixel di locazione (x, y) ;
- **Globale:** $[f(x, y)]$ rappresenta l'intera immagine f .

3.3 Operazioni puntuale

Si dice **operatore puntuale**, un operatore che ha preso in input il valore di un pixel e ne restituisce uno cambiato, il quale dipende esclusivamente dal valore del pixel in ingresso.

L'**obbiettivo** è quello di variare il contrasto. Infatti, eseguendo questa operazione, si evidenziano le differenze strutturali dell'oggetto rappresentato. Per farlo, basta cambiare il valore di un pixel per renderlo più scuro o più chiaro.

Un operatore puntuale può essere rappresentato tramite una **funzione** che prende in input un valore r e lo modifica in un valore $s = T(r)$ con s, r appartenenti allo stesso campo di definizione (esempio tra 0 e 255). Più in generale viene definita come:

$$T : [0, L - 1] \subset \mathbb{R} \longrightarrow [0, L - 1] \subset \mathbb{R}$$

Dato che un operatore puntuale dipende solo dal singolo valore del pixel, esso è dunque descritto da una tabella di questo tipo:

r	0	1	2	3	4	5	6	...
s	$T(0)$	$T(1)$	$T(2)$	$T(3)$	$T(4)$	$T(5)$	$T(6)$...

3.3.1 Identità

È l'operazione più semplice e non fa nulla:

$$s = r$$

3.3.2 Negativo

Rende l'immagine più scura:

$$s = L - 1 - r$$

Nel caso dei livelli di grigio:

$$s = 255 - r$$

Viene **utilizzata** quando si hanno dettagli grigi immersi in zone nere che si vogliono evidenziare.

3.3.3 Clamping

Limita l'intensità ad un range definito $[a, b]$:

$$T(r) = \begin{cases} a & \text{se } r < a \\ r & \text{se } a \leq r \leq b \\ b & \text{se } r > b \end{cases}$$

Viene **utilizzata** nel caso in cui ci siano dei pixel di rumore molto chiari o molto scuri che *mascherano* l'immagine. Quindi, si pensi per esempio ad un'immagine con dei puntini bianchi al quale si applica il *clamping* per rimuoverli.

3.3.4 Stretching/Shrinking

Stira/comprime le intensità di un range $[r_{\min}, r_{\max}]$ ad un range definito $[a, b]$:

$$s = \left[\frac{r - r_{\min}}{r_{\max} - r_{\min}} \right] [b - a] + a$$

In cui:

- r_{\min} / r_{\max} sono il più piccolo/grande livello di grigio del range che si vuole trattare;
- a, b sono il minimo e il massimo “stretchati”.

Nota bene: l'operazione è seguita da *rounging* (arrotondamento) nel caso di dominio di valori nei naturali (come in 0-255). Inoltre, lo **stretching** non risolve il problema del rumore impulsivo (puntini neri o bianchi), neanche se mascherato con il clamping.

3.3.5 Trasformazione logaritmica

La forma generale è:

$$s = c \cdot \log(1 + r) \quad r \in [0, L - 1]$$

Con c che rappresenta la **costante di normalizzazione**:

$$c = \frac{L - 1}{\log(L)}$$

La quale assicura la mappatura in $[0, L - 1]$. Inoltre, l'aggiunta di 1 permette di evitare il calcolo di quantità $\in [0, 1]$ che generano valori minori di zero ed in particolare il calcolo di $\log(0) = -\infty$.

Viene **utilizzata** quando si vuole mappare fasce strette di valori dell'immagine originale in fasce più ampie, aumentandone così il range del contrasto, rendendo inoltre l'interpretazione umana più informativa.

3.3.6 Trasformazione esponenziale

Al contrario della trasformazione logaritmica, la trasformazione esponenziale consente di aumentare il range di una fascia determinata di livelli di grigi chiari:

$$s = (e^r)^{\frac{1}{c}} - 1 \quad r \in [0, L - 1]$$

Con la **costante di normalizzazione**:

$$c = \frac{L - 1}{\log(L)}$$

3.3.7 Trasformazione di potenza

La trasformazione di potenza può essere espressa come:

$$s = cr^\gamma \quad c, \gamma > 0 \in \mathbb{R}$$

La costante c è scelta **in dipendenza da** γ in modo da normalizzare i valori di s nell'intervallo $[0, 255]$.

- $\gamma < 1$, la trasformazione ha effetti analoghi alla trasformazione logaritmica (3.3.5), cioè espansione della dinamica per bassi valori di r , mentre compressione della dinamica per alti valori di r ;
- $\gamma > 1$, la trasformazione ha effetti opposti ai valori negativi di gamma.

Nella pratica, il termine di normalizzazione c è complicato da definire analiticamente, quindi si preferiscono due versioni di s :

- **Non normalizzata:** $\tilde{s} = r^\gamma$;
- **Normalizzata** $s = cr^\gamma$

Per passare dalla versione **non normalizzata** alla **versione normalizzata** si esegue lo *stretching*:

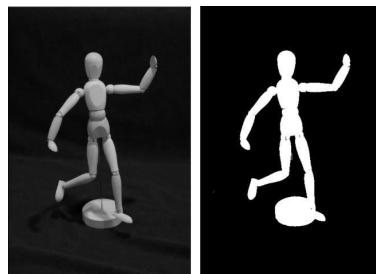
$$s = \left[\frac{\tilde{s} - \tilde{s}_{\min}}{\tilde{s}_{\max} - \tilde{s}_{\min}} - \tilde{s}_{\min} \right] [\max - \min]$$

Dove $\tilde{s}_{\min/\max}$ sono il più piccolo/grande livello di grigio e max e min sono il massimo e il minimo livello di grigio possibile (255, 0).

3.3.8 Binarizzazione

Produce un'immagine che ha solo due livelli: nero e bianco. Si **ottiene** scegliendo una soglia T , si imposta a colore nero tutti i pixel il cui valore è minore a T e si imposta a colore bianco tutti gli altri.

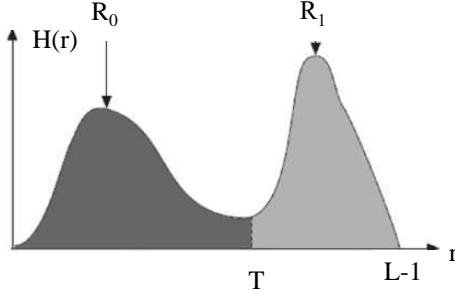
Si **utilizza** la binarizzazione per discriminare un oggetto dalla scena.



Soltanamente il suo utilizzo è prevalente nell'ambito delle immagini biomedicali e di videosorveglianza. La difficoltà maggiore di questa tecnica è il saper scegliere la soglia T più ragionevole.

3.3.9 Binarizzazione attraverso il metodo di Otsu

Questo metodo assume che ci siano due regioni da scegliere, come nella seguente figura:



Se l'immagine ha un **istogramma bimodale** la binarizzazione è efficace, altrimenti no. Le due regioni sono definite come:

$$\begin{aligned} p_{0 \rightarrow T}(r) &= \frac{H(r)}{\sum_{r=0}^T H(r)} \\ \sigma_{0 \rightarrow T}^2 &= \sum_{r=0}^T p_{0 \rightarrow T}(r) (r - \mu_{0 \rightarrow T})^2 \\ \mu_{0 \rightarrow T} &= \sum_{r=0}^T r \cdot p_{0 \rightarrow T}(r) \end{aligned}$$

La formula da minimizzare su T è la seguente:

$$\sigma_w^2(T) = W_0(T) \sigma_0^2(T) + W_1(T) \sigma_1^2(T)$$

Dove si considera la versione probabilistica dell'istogramma, ovvero la sua versione normalizzata, e si ha:

$$\begin{aligned} W_0(T) &= \sum_{r=0}^{T-1} p(r) \\ W_1(T) &= \sum_{r=T}^{L-1} p(r) \\ p(r) &= \frac{1}{M \cdot N} \sum_{r=0}^{L-1} H(r) \end{aligned}$$

Dove W_0, W_1 sono le probabilità che le due classi siano separate da T e σ_0^2 e σ_1^2 sono le varianze sui valori di istogramma assunti dalle due classi. **L'approccio "cicla" su tutti i possibili valori di T e restituisce:**

$$T_{best} = \arg_T \min (\sigma_w^2(T))$$

3.3.10 Equalizzazione

Un’immagine si dice **equalizzata** quando il contributo di ogni differente tonalità di grigio è simile. L’istogramma tende ad essere uniforme o appiattito.

L’**obbiettivo** è vedere l’istogramma come una distribuzione e di renderla il più simile a quella uniforme. Una **distribuzione uniforme** ha un’**entropia massima**⁴. Nelle immagini, ogni valore della distribuzione è un valore di grigio, per cui ogni valore di grigio appartiene all’entropia massima.

Si **utilizza** questo operatore poiché un istogramma piatto assicura a livello percettivo una risposta del cervello migliore (in termini di numero di dettagli che si riesce a riconoscere), per cui l’immagine diventa più “informativa” da osservare.

Se r_k è il k -esimo livello di grigio $k = 0, \dots, L - 1$ e $H(r_k)$ è il conteggio dato dall’istogramma dell’immagine di dimensione $M \times N$, allora si può definire:

$$p_r(r_k) = \frac{H(r_k)}{M \cdot N}$$

L’**equalizzazione dell’istogramma** si basa sulla seguente funzione T , con s_k che rappresenta il k -esimo valore di grigio in cui viene “mappato” r_k :

$$s_k = T(r_k) = (L - 1) \sum_{j=0}^k p_r(r_j) = \frac{\sum_{j=0}^k H(r_j)}{\frac{M \cdot N}{(L - 1)}}$$

Algoritmo

- I. Calcolare le L somme cumulative $\sum_{j=0}^k p_r(r_j)$ dei valori dell’istogramma visto come distribuzione con $k = 0, \dots, L - 1$;
- II. Moltiplicare i valori del passo precedente per il massimo di livelli di grigio $L - 1$;
- III. Normalizzazione dei valori calcolati al primo passo, dividendo per il numero totale di pixel $M \cdot N$ e arrotondamento;
- IV. Applicare il mapping T ottenuto.

⁴Secondo l’entropia, un sistema isolato si trasforma ed evolve nel tempo fino a raggiungere uno stato di equilibrio finale macroscopico in cui le differenze locali sono minime.

Esempio

Sia data un'immagine con $L = 8, 64 \times 64$ pixel ($M \cdot N = 4096$), con la seguente distribuzione d'intensità:

r_k	$H(r_k)$	$p_r(r_k) = \frac{H(r_k)}{M \cdot N}$
$r_0 = 0$	790	0.19
$r_1 = 1$	1023	0.25
$r_2 = 2$	850	0.21
$r_3 = 3$	656	0.16
$r_4 = 4$	329	0.08
$r_5 = 5$	245	0.06
$r_6 = 6$	122	0.03
$r_7 = 7$	81	0.02

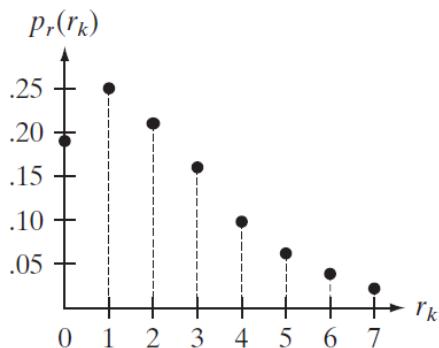


Figura 30: Rappresentazione della distribuzione di intensità.

Si applica la formula di equalizzazione:

$$\begin{aligned}
 s_0 &= T(r_0) = 7 \sum_{j=0}^0 p_r(r_j) = 7p_r(r_0) = 1.33 \\
 s_1 &= T(r_1) = 7 \sum_{j=0}^1 p_r(r_j) = 7p_r(r_0) + 7p_r(r_1) = 3.08
 \end{aligned}$$

Analogamente anche per gli altri valori si applica la formula e si trovano i seguenti valori:

$$\begin{aligned}s_2 &= 4.55 \\ s_3 &= 5.67 \\ s_4 &= 6.23 \\ s_5 &= 6.65 \\ s_6 &= 6.86 \\ s_7 &= 7.00\end{aligned}$$

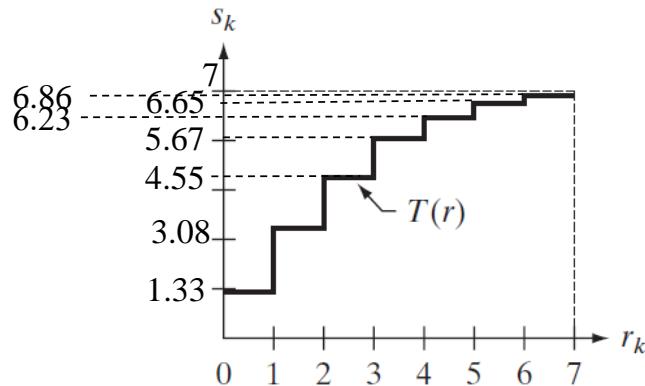


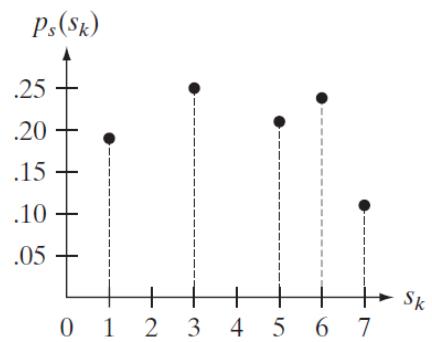
Figura 31: LUT (*Lookup Table*)

LUT (*Lookup Table*) è un termine utilizzato per descrivere una predeterminata lista di numeri che offre una “scorciatoia” per una specifica computazione. Nel contesto dei colori, una LUT trasforma i colori, ricevuti come input (camera), in un output desiderato (final footage).

L’immagine è quantizzata, quindi si effettua l’arrotondamento dei valori ottenendo l’intero più vicino:

$$\begin{aligned}s_0 &= 1.33 \longrightarrow 1 \\ s_1 &= 3.08 \longrightarrow 3 \\ s_2 &= 4.55 \longrightarrow 5 \\ s_3 &= 5.67 \longrightarrow 6 \\ s_4 &= 6.23 \longrightarrow 6 \\ s_5 &= 6.65 \longrightarrow 7 \\ s_6 &= 6.86 \longrightarrow 7 \\ s_7 &= 7.00 \longrightarrow 7\end{aligned}$$

Dopo l'arrotondamento, si ottiene una nuova immagine e il suo relativo istogramma.



3.4 Operazioni locali

Un'**operazione locale** restituisce un pixel che dipende da un limitato intorno del corrispondente punto in input. Tali operazioni vengono **utilizzati** per migliorare la qualità delle immagini o per estrarre delle informazioni dall'immagine.

Le operazioni locali sono come dei filtri spaziali dell'immagine. Il **filtraggio spaziale** è un'elaborazione T dell'immagine f dove un pixel di locazione (n, m) , di intensità $f(n, m)$, viene cambiato in $g(n, m)$ da una funzione dei pixel in un intorno di (n, m) , ossia:

$$g(n, m) = T([f(n, m)])$$

Dove le parentesi quadrate indicano che viene preso in considerazione un intorno di n, m . Ovviamente, il risultato dell'operazione, se applicato a tutti i pixel dell'immagine f , è una nuova immagine g .

Gli intorni presi maggiormente in considerazione sono di grandezza $K \times K$, con K dispari (per fare in modo di considerare uniformemente i pixel attorno al punto (n, m) di applicazione), di solito $3 \times 3, 5 \times 5, 7 \times 7$. I pixel al di fuori dell'intorno non prendono parte alla funzione.

Pseudocodice

- **Input:**

- Immagine f definita con un suo valore di pixel generico $f(n, m) \in [0 \dots L - 1] \subset \mathbb{N}$ con $(n, m) \in [1 \dots N] \times [1 \dots M] \subset \mathbb{N} \times \mathbb{N}$;
- Intorno di valori di pixel $[f(n, m)]$ ossia $f(n - u, m - v)$ definita come:

$$(u, v) \in \left[-\frac{K-1}{2} \dots \frac{K-1}{2} \right] \times \left[-\frac{K-1}{2} \dots \frac{K-1}{2} \right] \subset \mathbb{N} \times \mathbb{N} \quad K \in \{3, 5, 7, \dots\}$$

- **Output:**

- Nuova immagine $g(n, m) \in \mathbb{R}$ che attraverso operazioni puntuali può essere riportata in $g(n, m) \in [0 \dots L - 1] \subset \mathbb{N}$

- **Procedimento:**

```
for n = 1 ... N
    for m = 1 ... M
        g(n, m) = T([f(n, m)])
```

3.4.1 Filtraggi spaziali: lineari e non lineari

Le due principali categorie di operazioni locali sono lineari e non lineari:

- **Filtraggio lineare** se T è una combinazione lineare dei valori di pixel nel vicinato. Quindi, la convoluzione di un'immagine con un kernel è una somma di fattori ognuno dei quali è una moltiplicazione di un valore dell'immagine per un coefficiente del filtro:

$$g(n, m) = T([f(n, m)]) = h * f(n, m) = \sum_{u=-k}^{+k} \sum_{v=-k}^{+k} h(u, v) f(n - u, m - v)$$

$$k = \frac{K - 1}{2}$$

Un esempio di operazione lineare è la convoluzione;

- **Filtraggio non lineare** se T contiene operazioni non lineari sulle variabili indipendenti. Un esempio di operazioni non lineari sono la mediana dei pixel nell'intorno e il valore massimo dei pixel nel vicinato.

I filtraggi lineari **non** presentano un **problema ai bordi** nel momento in cui l'intorno è definito all'interno dell'immagine (cioè non cade fuori dall'immagine). Alcuni filtri lineari:

- **Cropping** è un filtro dove solo l'intorno cade all'interno dell'immagine, quindi il filtro viene applicato ad un'area ristretta e non a tutta l'immagine. Un esempio:

$$\text{Input} = \begin{bmatrix} 1 & 2 & 2 & 3 & 1 \\ 3 & 2 & 2 & 1 & 4 \\ 2 & 5 & 2 & 7 & 1 \\ 9 & 0 & 1 & 1 & 2 \\ 3 & 1 & 2 & 4 & 1 \end{bmatrix} \longrightarrow \text{Output} = \begin{bmatrix} * & * & * & * & * \\ * & 30 & 45 & 30 & * \\ * & 46 & 27 & 37 & * \\ * & 34 & 41 & 28 & * \\ * & * & * & * & * \end{bmatrix}$$

Le aree con un * non verranno calcolate;

- **Zero Padding** utilizzato per inserire degli zero che creano degli artefatti così da consentire il filtraggio. Un esempio:

$$\text{Input} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 2 & 3 & 1 & 0 \\ 0 & 3 & 2 & 2 & 1 & 4 & 0 \\ 0 & 2 & 5 & 2 & 7 & 1 & 0 \\ 0 & 9 & 0 & 1 & 1 & 2 & 0 \\ 0 & 3 & 1 & 2 & 4 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \longrightarrow \text{Output} = \begin{bmatrix} 11 & 19 & 17 & 22 & 11 \\ 25 & 30 & 45 & 30 & 31 \\ 25 & 46 & 27 & 37 & 19 \\ 35 & 34 & 41 & 28 & 29 \\ 16 & 27 & 12 & 18 & 10 \end{bmatrix}$$

Le aree in grigio non vengono calcolate.

- **Replicazione** utilizzata per creare artefatti, infatti l'immagine risultante non è realistica. Un esempio:

$$\text{Input} = \begin{bmatrix} 1 & 1 & 2 & 2 & 3 & 1 & 1 \\ 1 & 1 & 2 & 2 & 3 & 1 & 1 \\ 3 & 3 & 2 & 2 & 1 & 4 & 4 \\ 2 & 2 & 5 & 2 & 7 & 1 & 1 \\ 9 & 9 & 0 & 1 & 1 & 2 & 2 \\ 3 & 3 & 1 & 2 & 4 & 1 & 1 \\ 3 & 3 & 1 & 2 & 4 & 1 & 1 \end{bmatrix} \longrightarrow \text{Output} = \begin{bmatrix} 25 & 27 & 29 & 31 & 33 \\ 34 & 30 & 45 & 30 & 39 \\ 51 & 46 & 27 & 37 & 32 \\ 54 & 34 & 41 & 28 & 35 \\ 48 & 32 & 24 & 34 & 26 \end{bmatrix}$$

Le aree in grigio non vengono calcolate.

3.5 Rumore nelle immagini

Il **rumore nelle immagini** è un disturbo dell'immagine introdotto dal sistema di acquisizione (e.g. fotocamera) o dal mezzo di propagazione che ne degrada la qualità (e.g. Whatsapp). Il rumore è tipicamente *modellato* come **additivo e casuale**:

$$\tilde{f}(n, m) = f(n, m) + \varepsilon(n, m)$$

Dove f è l'**immagine** priva di rumore e ε è un processo aleatorio che genera delle quantità che seguono una distribuzione particolare, indipendentemente da dove il processo è collocato nell'immagine, ovvero indipendentemente da n, m .

Esistono due **tipi di rumore**: gaussiano additivo bianco (rumore generato da una distribuzione gaussiana) e impulsivo (rumore generato da una distribuzione bernoulliana).

La **quantità di rumore** può essere stimata attraverso la misura di *SNR* (*signal to noise ratio*), di cui esistono varie versioni. La più utilizzata è la **mean square, SNR_{ms}**:

$$SNR_{ms} = \frac{\sum_{n=1}^N \sum_{m=1}^M \tilde{f}(n, m)^2}{\sum_{n=1}^N \sum_{m=1}^M [\tilde{f}(n, m) - f(n, m)]^2}$$

Una forma alternativa della *SNR* può essere stimata grazie alla varianza σ_n^2 , o alla deviazione standard σ_n :

$$SNR = \frac{\sigma_s}{\sigma_n}$$

Dove σ_s è la deviazione standard del segnale e σ_n è la deviazione standard dell'immagine affetta da rumore. Per questo motivo si utilizzano immagini ad alto contrasto, poiché σ_s risulta maggiore!

3.5.1 Rumore gaussiano additivo bianco

Il **rumore gaussiano additivo bianco** è un processo stocastico, ovvero una variabile aleatoria che emette valori casuali nel tempo $\varepsilon(t)$ o nello spazio $\varepsilon(n, m)$ con le seguenti caratteristiche:

- Si somma al segnale pulito;
- Non è periodico nel tempo o nello spazio;
- I valori vengono prodotti con la seguente probabilità:

$$P(\varepsilon(n, m) = l) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{(l - 0)^2}{2\sigma^2}\right)$$

In cui 0 è uguale a μ , ovvero indica la media del rumore. Inoltre, data una distribuzione gaussiana μ, σ^2 il 98% di valori $l \in [\mu - 2.5\sigma, \mu + 2.5\sigma]$.

- I valori seguono una distribuzione gaussiana di media pari a zero, ed una particolare varianza σ^2 (o deviazione standard σ) dove più è alta la varianza, più distanti da zero saranno i numeri prodotti e sommati all'immagine pulita, più rumorosa l'immagine finale.

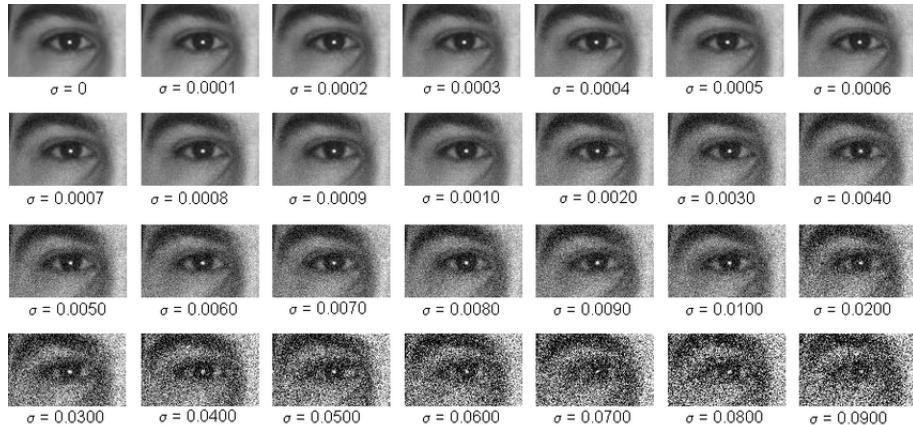


Figura 32: Esempio di rumore gaussiano a diversi valori di σ .

3.5.2 Rumore impulsivo

Il **rumore impulsivo** è causato da alterazione brusche nel segnale, viene parametrizzato da un fattore D (una percentuale) che è la densità con cui esso si localizza su pixel dell'immagine: maggiore il valore di intensità D , maggiore sarà il numero di pixel affetti.

Per esempio, il disturbo sale e pepe (*salt-and-pepper noise*) può essere utilizzato selezionando una percentuale D di pixel, ovvero $D\%$, in maniera uniforme nell'immagine e per ogni pixel si assegna un valore minimo o massimo con probabilità uniforme pari a $p = 0.5$.

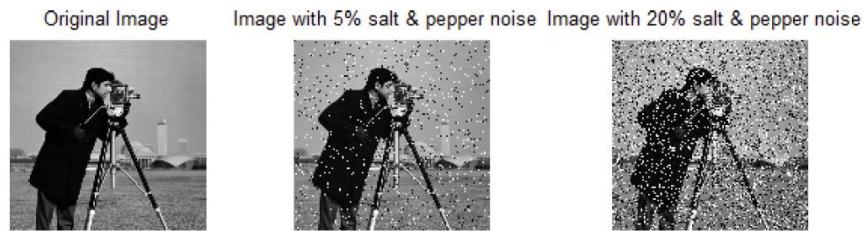


Figura 33: Esempio di applicazione dell'effetto *salt-and-pepper noise*.

3.6 Altre operazioni locali: tipologie di filtraggio

Esistono altre 3 tipologie principali di filtraggio:

- **Smoothing**, utilizzato per aumentare il *SNR* (pagina 68), ovvero per rimuovere il rumore.
Per esempio, il filtro di media, mediano, gaussiano;
- **Sharpening**, utilizzato per aumentare il grado di dettaglio delle immagini.
Per esempio, il filtro laplaciano;
- **Estrazione di caratteristiche**, utilizzato per estrarre rappresentazioni alternative alle immagini di partenza, che ne evidenzino aspetti particolari (edge, microstrutture, oggetti).
Per esempio, il filtro prewitt, sobel, canny.

3.6.1 Smoothing - Filtro media

Il **filtro media** è utilizzato per **rimuovere il rumore gaussiano**. È un filtraggio T lineare, si attua attraverso la convoluzione dell'immagine con la maschera media la quale ha le seguenti caratteristiche:

- Dimensioni $K \times K$ con K dispari;
- I suoi coefficienti sono tutti uguali e pari a $\frac{1}{K^2}$;
- Il suo funzionamento è il seguente: dato un intorno di applicazione $[(n, m)]$, esso calcola la media dei valori vii compresi $[\tilde{f}(n, m)]$, e la sostituisce al posto del valore $\tilde{f}(n, m)$:

$$g(n, m) = T([\tilde{f}(n, m)]) = E([f(n, m)])$$

Dove E è l'operatore di media, perché T essenzialmente è l'operatore di media

Si osservi che la somma dei valori del kernel è 1:

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

Questo significa che il filtraggio in una locazione (x, y) è una **combinazione lineare convessa**. In altre parole, la somma dei livelli di grigio dell'immagine originale f e di quella processata g sono uguali (a meno di padding!).

Maggiore è l'ampiezza K della maschera, **più severo è l'effetto della media** sulla struttura dell'immagine.

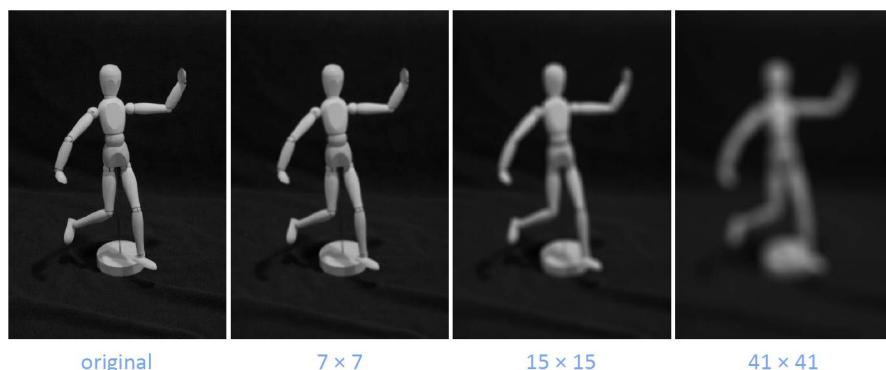


Figura 34: Esempio di filtro media all'aumentare di K , ovvero della grandezza.

3.6.2 Smoothing - Filtro mediano

Il **filtro mediano** è utilizzato per **rimuovere il rumore impulsivo**. È un filtraggio T non lineare, che si realizza attraverso un algoritmo. Data la matrice dell'immagine:

$$\begin{bmatrix} 240 & 245 & 0 \\ 247 & 0 & 244 \\ 251 & 246 & 250 \end{bmatrix}$$

1. Si calcola la media di tutti i valori della matrice:

$$\frac{240 + 245 + 0 + 247 + 0 + 244 + 251 + 246 + 250}{9} \approx 191$$

2. Si inseriscono in un vettore riga i valori in ordine crescente:

$$[0 \ 0 \ 240 \ 244 \ 245 \ 246 \ 247 \ 250 \ 251]$$

3. Si calcola la mediana del vettore:

$$9 \div 2 = 4.5 \longrightarrow [0 \ 0 \ 240 \ 244 \ \underline{245} \ 246 \ 247 \ 250 \ 251]$$



Figura 35: Esempio di applicazione di filtro mediano.

3.6.3 Smoothing - Filtro gaussiano

Il **filtro gaussiano** è quello di rendere più “lisica (*smooth*)” l’immagine, in modo simile al filtraggio di media, e come parametro l’operatore prende il valore σ che rappresenta la **forza** (maggiore è il valore, più è forte lo *smoothing*). La **differenza** sostanziale tra il filtraggio di media e il filtraggio gaussiano è che quest’ultimo è una media pesata, dove i pesi più vicini al centro della maschera hanno valori più alti. Così facendo si ha:

- **Vantaggio**

- Il filtraggio gaussiano effettua uno *smoothing* più lieve, **preservando i contorni meglio** di quanto faccia il filtraggio media. Quindi, la struttura viene preservata meglio.

- **Svantaggio**

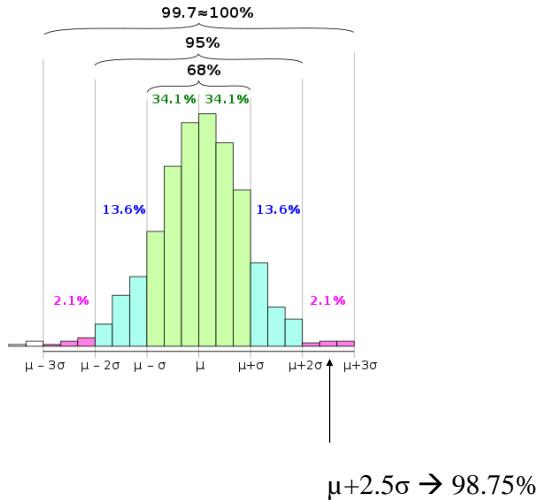
- Il rumore viene rimosso in maniera inferiore e questo provoca l'**impossibilità di applicare la formula di annullamento del rumore** visto per il filtro media.

Questo filtro può essere **implementato in maniera efficiente** in quanto la maschera è separabile, ovvero è possibile eseguirlo facendo un filtraggio prima su tutte le N righe dell’immagine come se fossero funzioni $1D$, e poi su tutte le M colonne:

$$\begin{aligned} I_G(i, j) &= \sum_{h=-\frac{m}{2}}^{\frac{m}{2}} \sum_{k=-\frac{m}{2}}^{\frac{m}{2}} G(h, k) I(i + h, j + k) \\ &= \sum_{h=-\frac{m}{2}}^{\frac{m}{2}} \sum_{k=-\frac{m}{2}}^{\frac{m}{2}} \exp\left(-\frac{h^2 + k^2}{2\sigma^2}\right) I(i + h, j + k) \\ &= \sum_{h=-\frac{m}{2}}^{\frac{m}{2}} \exp\left(-\frac{h^2}{2\sigma^2}\right) \sum_{k=-\frac{m}{2}}^{\frac{m}{2}} \exp\left(-\frac{k^2}{2\sigma^2}\right) I(i + h, j + k) \end{aligned}$$

Dove le variabili i, j, h, k, m sono indici per comprendere la **separabilità**. Quest’ultima consente di progettare manualmente un filtro gaussiano come segue nella prossima pagina.

Si definiscono i parametri σ e W . Quindi, si fissi σ e si trovi la dimensione della maschera W sapendo che W deve essere tale da contenere un'elevata percentuale di probabilità (uguale all'area della densità gaussiana, come in figura). In particolare, la statistica dice che con $W = 5\sigma$ si copre il 98.75% dell'area della densità gaussiana. In altre parole, se si vuole $\sigma = 1$ allora $W = 5 \cdot 1 = 5$; se si vuole $\sigma = 0.6$ allora $W = 5 \cdot 0.6 = 3$.



3.6.4 Domanda da esame

Domanda

Il livello di noise gaussiano presente in un'immagine, se esso noto (e.g. $\sigma_1 = 0.01$) può guidare la scelta del parametro σ_2 della maschera di filtro gaussiano?

Risposta

No, perché il rumore gaussiano lavora su tutti i valori di grigio, mentre il filtro gaussiano sulle coordinate e non c'è correlazione. Invece, il filtro media è quello ideale.

3.6.5 Filtraggi di sharpening

I **filtraggi di sharpening** servono per evidenziare i dettagli o come post processing dopo filtraggi di *smoothing* (questo perché i filtraggi di *smoothing* eliminano i dettagli). Per lo stesso motivo, i filtraggi di sharpening possono incrementare il rumore (un'immagine di rumore è un'immagine ad alta frequenza). Esistono due categorie: ***basic highpass spatial filtering*** e ***high boost filtering***.

I filtri di sharpening sono detti anche **filtri di derivata**, poiché calcolano numericamente nell'intorno in cui sono definiti la derivata locale (prima o seconda) dell'immagine.

Rispetto a x

Derivata asimmetrica

$$\begin{aligned} I_x(x, y) &= \frac{\partial I}{\partial x}(x, y) = \lim_{h \rightarrow 0} \frac{I(x+h, y) - I(x, y)}{h} \\ I_x[x, y] &= \frac{\partial I}{\partial x}[x, y] = I[x+1, y] - I[x, y] \end{aligned}$$

Derivata simmetrica

$$\begin{aligned} I_x(x, y) &= \frac{\partial I}{\partial x}(x, y) = \lim_{h \rightarrow 0} \frac{I(x+h, y) - I(x-h, y)}{2h} \\ I_x[x, y] &= \frac{\partial I}{\partial x}[x, y] = \frac{1}{2}(I[x+1, y] - I[x-1, y]) \end{aligned}$$

Filtro differenziale asimmetrico

$$\partial_x = [-1 \quad 1]$$

Filtro differenziale simmetrico

$$\partial_x = \frac{1}{2} [-1 \quad 0 \quad 1]$$

Rispetto a y

Derivata asimmetrica

$$I_y(x, y) = \frac{\partial I}{\partial y}(x, y) = \lim_{h \rightarrow 0} \frac{I(x, y + h) - I(x, y)}{h}$$

$$I_y[x, y] = \frac{\partial I}{\partial y}[x, y] = I[x, y + 1] - I[x, y]$$

Derivata simmetrica

$$I_y(x, y) = \frac{\partial I}{\partial y}(x, y) = \lim_{h \rightarrow 0} \frac{I(x, y + h) - I(x, y - h)}{2h}$$

$$I_y[x, y] = \frac{\partial I}{\partial y}[x, y] = \frac{1}{2} (I[x, y + 1] - I[x, y - 1])$$

Filtro differenziale asimmetrico

$$\partial_y = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Filtro differenziale simmetrico

$$\partial_y = \frac{1}{2} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

Proprietà della derivata prima:

- Nulla in regioni di intensità costante;
- Non nulla in presenza di variazioni di intensità.

Proprietà della derivata seconda:

- Nulla in regioni di intensità costante;
- Nulla in presenza di variazioni costanti di intensità (rampe);
- Non nulla in presenza di variazioni non costanti (all'inizio e alla fine di rampe).

3.6.6 Sharpening - Basic Highpass Spatial Filtering

Il filtraggio di sharpening chiamato **basic highpass spatial filtering** è un **filtraggio lineare** con il laplaciano, con la maschera H caratterizzata da coefficienti di un segno (e.g. positivo) vicino al centro e di segno opposto (e.g. negativo) nella periferia esterna.

Una tipica maschera di filtraggio, chiamata laplaciana:

$$\frac{1}{9} \times \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

Altre maschere laplaciane:

$$\begin{array}{ll} \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} & \begin{bmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{bmatrix} \\ \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} & \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix} \end{array}$$

Alcune caratteristiche:

- La **somma dei coefficienti è zero**. Questo indica che quando il filtro passa su regioni con livelli di grigio quasi stabili, l'output della maschera è zero o molto piccolo;
- L'**uscita è alta** quando il valore centrale differisce dai valori periferici;
- L'**immagine di output non assomiglierà** a quella originale;
- L'**immagine di output mostra tutti i dettagli**;
- Sono inclusi alcuni **ridimensionamenti e/o clipping**, necessari per compensare eventuali livelli di grigio negativi dopo il filtraggio.

Il filtraggio lineare utilizza il *basic highpass spatial filtering* per **creare un'immagine realistica**, simile a quella di partenza, con gli edge amplificati:

$$g(n, m) = T([f(n, m)]) = f(n, m) + c \cdot h * f(n, m)$$

In cui h indica la **maschera laplaciana**, c è una **costante** pari a uno nel caso in cui il pixel centrale della maschera laplaciana sia positivo, altrimenti -1 .

3.6.7 Sharpening - Filtro Laplaciano

La funzione **laplaciana** prende in ingresso un parametro α il cui significato è legato all'importanza che si vuole dare agli edge verticali e orizzontali ($\alpha = 0$), diagonali ($\alpha = 1$), tutti gli edge ($\alpha = 0.5$), attraverso questa formula:

$$h = \frac{1}{\alpha + 1} \begin{bmatrix} -\alpha & \alpha - 1 & -\alpha \\ \alpha - 1 & \alpha + 5 & \alpha - 1 \\ -\alpha & \alpha - 1 & -\alpha \end{bmatrix}$$

3.6.8 Sharpening - High Boost Filtering

L'immagine filtrata dallo *sharpening* si ottiene sottraendo l'immagine filtrata con *smoothing* dall'immagine originale:

Immagine filtrata dallo sharpening = Originale – Im. filtrata con smoothing

Se la costante A rappresenta un **fattore di amplificazione degli edge**, allora il filtro **high boost filtering** è definito come:

$$\text{High-boost} = A \cdot \text{Originale} + \text{Im. filt. dallo sharpening}$$

A differenza degli altri filtri, questo dà maggiore libertà al progettista. Infatti, il blur può avvenire attraverso una maschera di supporto arbitrariamente grande.

4 Elaborazione di immagini - Rinforzo del dominio delle frequenze

4.1 Ripasso formule utili

Si ricorda la trasformata di Fourier discreta a una dimensione (1D) che si rappresenta con la seguente equazione:

$$\tilde{F}\left(\frac{m}{M} \frac{1}{\Delta T}\right) = F\left(\frac{m}{M} \frac{1}{\Delta T}\right) = F_m = \sum_{n=0}^{M-1} f_n e^{-j2\pi \frac{m}{M} n}$$

Mentre la trasformata di Fourier discreta inversa (o antitrasformata), corrisponde a:

$$\tilde{f}(n\Delta T) = f(n\Delta T) = f_n = \frac{1}{M} \sum_{m=0}^{M-1} F_m e^{j2\pi \frac{m}{M} n}$$

È possibile fare alcune osservazioni riguardo le trasformate di Fourier in una e due dimensioni. Partendo dalla definizione:

$$F_m = \sum_{n=0}^{M-1} f_n e^{-j2\pi \frac{m}{M} n}$$

Si intende la trasformata di Fourier discreta come la moltiplicazione del segnale per delle funzioni sinusoidali 1D. Più il segnale "assomiglia" alla sinusoide di frequenza specifica con cui si moltiplica, più il valore di ampiezza della trasformata di Fourier discreta è alto per quella frequenza.

Si esegue un cambio di variabile, quindi μ diventa la **variabile delle frequenze** e x quella dello **spazio**. Entrambe rappresentano una quantità campionata e 1D è:

$$F_m = \sum_{n=0}^{M-1} f_n e^{-j2\pi \frac{m}{M} n} \longrightarrow F_u = F(u) = \sum_{x=0}^{M-1} f(x) e^{-j2\pi \frac{u}{M} x}$$

Si passa alla seconda dimensione aggiungendo y e v per l'**analisi verticale**:

$$F(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi \left(\frac{u}{M} x + \frac{v}{N} y \right)}$$

In cui $F(u, v)$ rappresenta la **risposta** di I all'immagine (complessa) di base $(u, v) e^{\dots}$, invece l'esponenziale rappresenta l'**immagine** (complessa) base che è funzione di (u, v) .

Si elencano un paio di **proprietà della trasformata di Fourier discreta 2D**, utili per questo capitolo:

- **Traslazione:**

$$\text{Segnale: } g(x, y) = f(x - x_0, y - y_0)$$

$$\text{T.d.F.: } G(u, v) = F(u, v) e^{-j2\pi(\frac{u}{M}x_0 + \frac{v}{N}y_0)}$$

$$\text{Segnale: } g(x, y) = f(x, y) e^{j2\pi(\frac{u}{M}x_0 + \frac{v}{N}y_0)}$$

$$\text{T.d.F.: } G(u, v) = f(u - u_0, v - v_0)$$

- **Rotazione:** la trasformata di Fourier di un'immagine a cui è stata applicata una rotazione θ , porterà ad un'immagine di trasformata ruotata di angolo θ :

$$\mathcal{F}(f_\theta)(u, v) = \mathcal{F}(f)_\theta(u, v)$$

Si richiama anche il **teorema di convoluzione**, ovvero la convoluzione nel tempo tra due segnali corrisponde alla moltiplicazione nel dominio delle frequenze. Ovviamente per cambiare dominio è necessario applicare Fourier:

$$\mathcal{F}(f * h) = \mathcal{F}(f) \cdot \mathcal{F}(h)$$

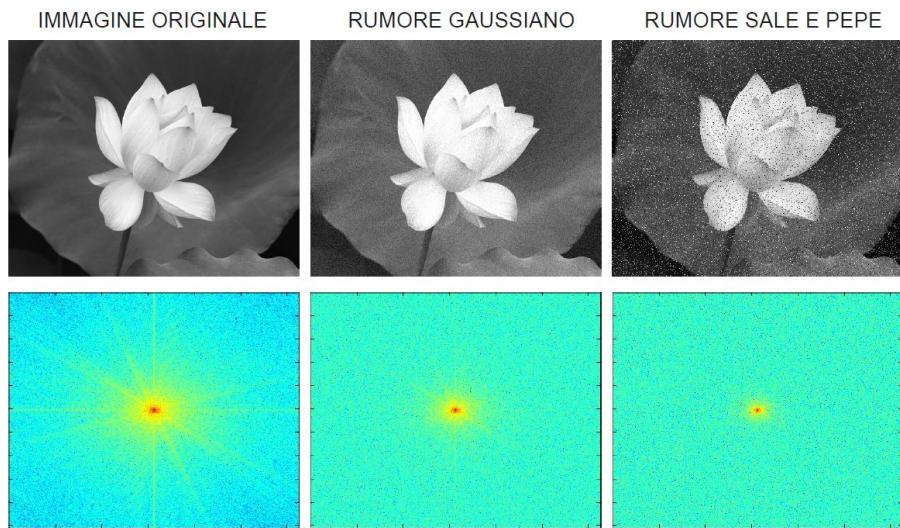
$$\mathcal{F}(f \cdot h) = \mathcal{F}(f) * \mathcal{F}(h)$$

Questo teorema è la base del filtraggio in frequenza.

4.2 Rumore nel dominio delle frequenze

Il **rumore** porta ad un aumento delle alte frequenze. Infatti, eseguendo una riduzione delle alte frequenze, si riduce sia il rumore (aspetto *positivo*) sia il livello di dettaglio (aspetto *negativo*).

Qui di seguito, vengono lasciate tre immagini per eseguire un confronto. L'immagine originale presenta certe frequenze, ma applicando il rumore gaussiano, o sale e pepe (*salt-and-pepper noise*), viene ridotto sensibilmente sia il rumore che il livello di dettaglio.



4.3 Panoramica su filtri passa alto (*high-pass*) e passa basso (*low-pass*)

I filtri passo alto (*high-pass*) e passa basso (*low-pass*) sono identici a quelli visti durante il corso di Sistemi. L'obbiettivo di questo paragrafo è introdurre qualche definizione e dare una panoramica generale dei due filtri approfonditi nei prossimi paragrafi.

In un'immagine, le **frequenze basse** rappresentano informazioni con variazioni di intensità “lente”. Per esempio, le gradazioni di colore su un muro illuminato, o la nuvolosità nel cielo.

Al contrario, le informazioni ad **alte frequenze** indicano informazioni con variazioni di intensità “repentine”. Per esempio, spigoli, angoli e rumore.

Per lavorare sulle frequenze esistono due filtri importanti:

- **Filtro passa basso** (*low-pass filter*) rimuove dall'immagine le informazioni ad alte frequenze e mantiene quelle a basse frequenze.
- **Filtro passa alto** (*high-pass filter*) rimuove dall'immagine le informazioni a basse frequenze e mantiene quelle ad alte frequenze.

Dato che ogni filtro è il contrario dell'altro, da ciascuno è possibile derivare il suo opposto:

$$H_{PA} = 1 - H_{PB}$$

In cui PA indica passa alto e PB passa basso.

4.4 Filtri passa basso (*low-pass*)

4.4.1 Filtri passa basso ideale

Il **filtro passa basso ideale** viene usato per ottenere: lo sfocamento e lo *smoothing*. Solitamente, per ottenere questi risultati si attenuano le alte frequenze, ma facendo così si ottiene anche una riduzione inevitabile del rumore.

Matematicamente parlando, un filtro passa basso ideale è una funzione di trasferimento (uguale alla sua trasformata nel dominio delle frequenze) di una box.

Viene detta **ideale** perché come si vede in figura, una transizione così netta in corrispondenza alla *frequenza di taglio* non è analogicamente realizzabile. In parole poche, è ideale poiché nell'elettronica non può fisicamente avvenire un calo di energia così repentino.

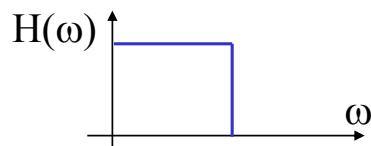


Figura 36: Funzione di trasferimento di un filtro passa basso ideale.

Dato che i filtri studiati sono solamente digitali, teoricamente potremmo tralasciare questo taglio repentino. Purtroppo, esso provoca un effetto visivo indesiderato chiamato **ringing**.

Il **ringing** (o effetto di Gibbs) è dovuto al fatto di eseguire un filtraggio passa basso ideale (in frequenza) equivalente ad eseguire una convoluzione con l'operatore sinc (nello spazio). Ne consegue che la risposta all'impulso del passa basso ideale è ancora un sinc e l'**immagine visivamente risulta increspata vicino ai bordi taglienti**⁵.

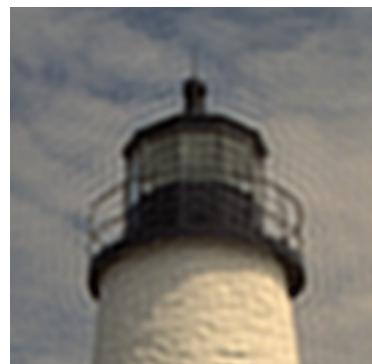
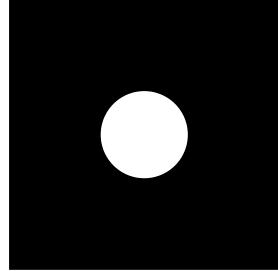


Figura 37: Effetto ringing su un'immagine.

⁵Fonte: *Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University*

Segnale H :



Le formule da applicare per il filtro basso ideale sono:

$$H(u, v) = \begin{cases} 1 & \text{se } D(u, v) \leq D_0 \\ 0 & \text{se } D(u, v) > D_0 \end{cases}$$

In cui D_0 è uguale a μ_{thresh} che indica la soglia e $D(u, v)$ è il raggio del cerchio:

$$D(u, v) = \sqrt{\left(u - \frac{M}{2}\right)^2 + \left(v - \frac{N}{2}\right)^2}$$

Da notare che solo le frequenze nel cerchio di raggio D_0 vengono mantenute.

4.4.2 Filtri passa basso di Butterworth

Il **filtro passa basso di Butterworth** è un filtro con attenuazione dolce in prossimità della frequenza di taglio. La **proprietà caratterizzante** di questo

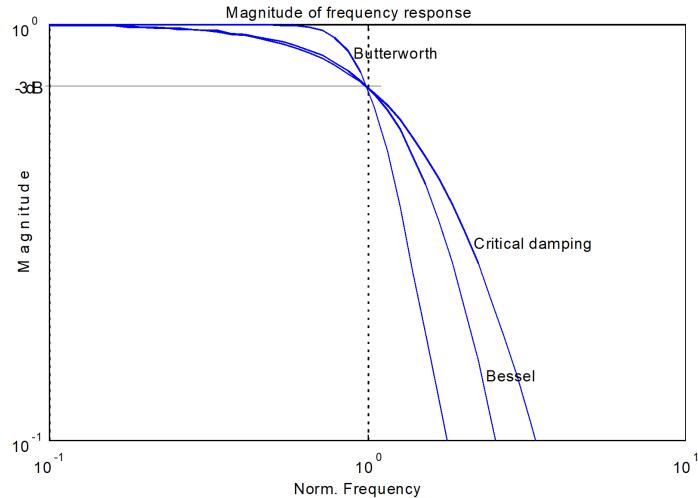


Figura 38: Filtro passa basso di Butterworth.

filtro è la **risposta molto ripida nella banda passante**.

L'ordine è n e la frequenza di taglio $D_0 = \mu_{\text{thresh}}$:

$$H(u, v) = \frac{1}{1 + \left(\frac{D(u, v)}{D_0}\right)^{2n}}$$

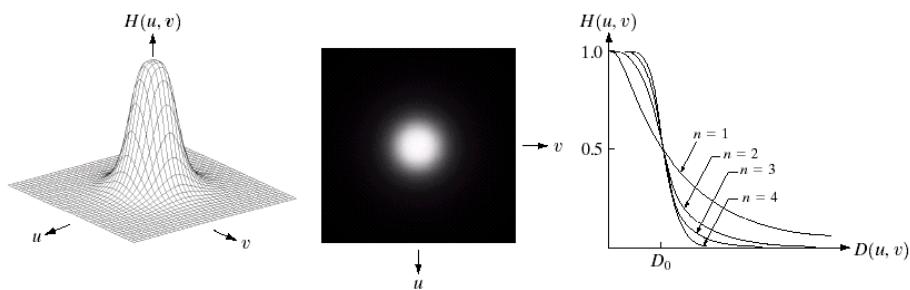


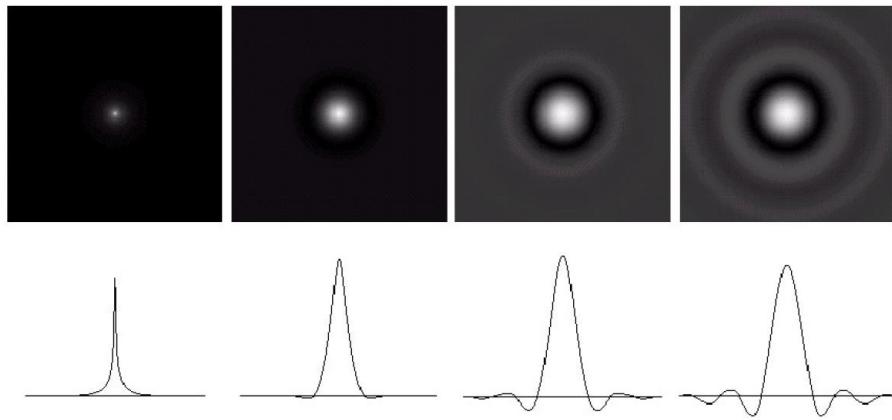
Figura 39: Applicazione dell'ordine e della frequenza di taglio.

Le formule da applicare:

$$H(u, v) = \frac{1}{1 + \left[\frac{D(u, v)}{D_0} \right]^{2n}}$$

$$D(u, v) = \sqrt{\left(u - \frac{M}{2} \right)^2 + \left(v - \frac{N}{2} \right)^2}$$

È interessante notare che facendo tendere n all'infinito, si ottiene il filtro passa basso ideale:



4.4.3 Filtri passa basso Gaussiano

La trasformata di Fourier di una funzione Gaussiana è anch'essa Gaussiana:

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

$$f(t) = \sqrt{2\pi}\sigma Ae^{-2\pi^2\sigma^2t^2}$$

Oppure allo stesso identico modo:

$$f(t) = Ae^{-\frac{t}{2\sigma^2}}$$

$$F(u) = \sqrt{2\pi}\sigma Ae^{-2\pi^2\sigma^2u^2}$$

Le equazioni di questa operazione sono:

$$\begin{aligned} H(u, v) &= e^{-\frac{D^2(u, v)}{2D_0^2}} \\ D(u, v) &= \sqrt{\left(u - \frac{M}{2}\right)^2 + \left(v - \frac{N}{2}\right)^2} \end{aligned}$$

In cui la costante D_0 può essere sostituita con σ , ovvero l'effettiva deviazione standard della distribuzione Gaussiana. In questo caso, la frequenza di taglio D_0 corrisponde alla deviazione standard σ . In altre parole, quando $D(u, v) = \sigma$, allora l'intensità di taglio è 0.607 e il filtraggio crea un attenuamento di quella frequenza pari al 60.7%.

Un'**osservazione** interessante è la seguente. Un filtro gaussiano con una certa scala nel dominio delle frequenze, corrisponde ad un filtro gaussiano con scala inversa nel dominio dello spazio.

4.4.4 Sintesi

Qui di seguito si elencano i filtri passa basso più importanti:

- **Ideale:** è una brusca transizione in corrispondenza della frequenza di *cut-off*. Questo causa un fenomeno di Gibbs o di *ringing*;
- **Gaussiano:** transizione di *cut-off* dolce. Il parametro σ determina la frequenza di *cut-off*;
- **Butterworth:** ha una rapidità variabile e transizione *smooth*. La ripidità viene modellata dall'ordine del filtro. La frequenza di *cut-off* viene selezionata indipendentemente dall'ordine del filtro. Si può avere *ringing* per ordini elevati.

4.5 Filtri passa alto (*high-pass*)

Un **filtro passa alto** sopprime (blocca) le basse frequenze e lascia passare le alte frequenze. La costruzione di un filtro passa alto può essere eseguita come:

$$H_{PA} = 1 - H_{PB}$$

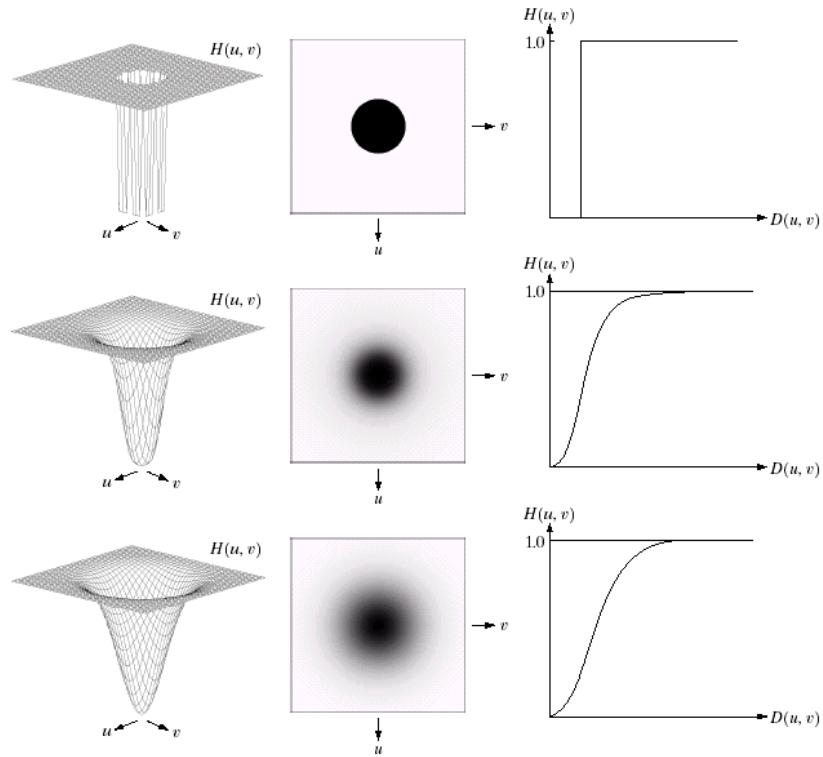
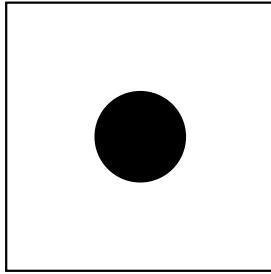


Figura 40: Filtri passa alto, dall'alto: ideale, di Butterworth, Gaussiano.

4.5.1 Filtri passa alto ideale

Dato il segnale H :



Le formule da applicare per il **filtro passa alto ideale** sono:

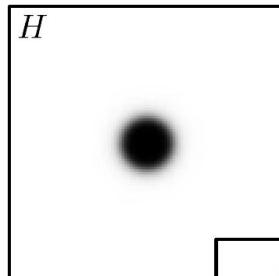
$$H(u, v) = \begin{cases} 0 & \text{se } D(u, v) \leq D_0 \\ 1 & \text{se } D(u, v) > D_0 \end{cases}$$

$$D(u, v) = \sqrt{\left(u - \frac{M}{2}\right)^2 + \left(v - \frac{N}{2}\right)^2}$$

Solamente le frequenze fuori dal cerchio di raggio D_0 vengono mantenute.

4.5.2 Filtri passa alto di Butterworth

Dato il segnale H :



Le formule da applicare per il **filtro passa alto di Butterworth** sono:

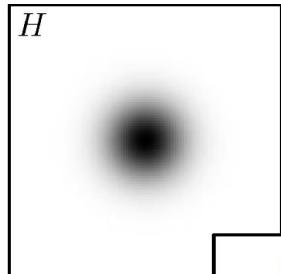
$$H(u, v) = \frac{1}{1 + \left[\frac{D(u, v)}{D_0}\right]^{-2n}}$$

$$D(u, v) = \sqrt{\left(u - \frac{M}{2}\right)^2 + \left(v - \frac{N}{2}\right)^2}$$

Per n tendente ad infinito si ha il filtro passa alto ideale.

4.5.3 Filtri passa alto Gaussiano

Dato il segnale H :



Le formule da applicare per il **filtro passa alto Gaussiano** sono:

$$H(u, v) = 1 - e^{-\frac{D^2(u, v)}{2D_0^2}}$$
$$D(u, v) = \sqrt{\left(u - \frac{M}{2}\right)^2 + \left(v - \frac{N}{2}\right)^2}$$

4.6 Filtri per enfatizzare le alte frequenze

Esiste anche il **filtro omeomorfo** che attenua le basse frequenze e aumenta quelle delle alte:

$$H(u, v) = (\gamma_H - \gamma_L) \left[1 - e^{-\frac{cD^2[u, v]}{D_0^2}} \right] + \gamma_L$$

Invece, usando il **filtro passa alto**, è possibile enfatizza aumentando la variabile k :

$$H(u, v) = (1 + k \cdot H_{PA}(u, v)) \cdot H$$

Dove k è il contributo delle alte frequenze, H lo spettro dell'immagine e H_{PA} il filtro passa alto

4.7 Filtri passa banda ideale 1D e ferma banda ideale

I seguenti filtri operano su una banda di frequenze, al contrario i precedenti filtri lavoravano su alte o basse frequenze.

Analogamente ai filtri passa basso e alto, anche qui è possibile derivare uno dei due filtri avendo solamente l'altro:

$$H_{PBn} = 1 - H_{FBn}$$

Dove PB indica passa banda e FB ferma banda.

Un **filtro passa banda ideale 1D** sopprime tutte le frequenze al di fuori di un intervallo di frequenze specificato. Più formalmente, esso è rappresentato con la seguente funzione:

$$G(\mu) = \begin{cases} 1 & \text{se } u_1 < |\mu| < u_2 \\ 0 & \text{altrimenti} \end{cases}$$

Le variabili μ rappresentano l'intervallo (come si vede in figura) e al di fuori di esso il valore è zero.

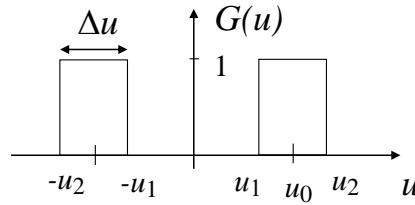


Figura 41: Rappresentazione grafica del filtro passa banda ideale 1D nel **dominio delle frequenze**.

Dalla rappresentazione grafica si ricava anche la relativa **funzione di trasferimento nel dominio delle frequenze**, ovvero la somma di due box traslate nel tempo ottenibili eseguendo la convoluzione di una box per due impulsi unitari (Dirac):

$$G(u) = \Pi\left(\frac{u}{\Delta u}\right) * [\delta(u - u_0) + \delta(u + u_0)]$$

Dove i termini u_0 e Δu indicano:

- u_0 indica dov'è centrata la box ed è ottenibile in questo modo:

$$u_0 = \frac{(u_1 + u_2)}{2}$$

- Δu indica la larghezza della box ed è ottenibile in questo modo:

$$\Delta u = u_2 - u_1$$

Invece, la **funzione di trasferimento nel dominio del tempo** è la seguente:

$$g(t) = 2\Delta u \cdot \frac{\sin(\pi t \Delta u)}{\pi t \Delta u} \cdot \cos(2\pi u_0 t)$$

Dato che $\Delta u < u_0$, allora la funzione $g(t)$ è un coseno modulato dal sinc. Pensandoci bene, la box viene rappresentata dalla funzione sinc⁶ e i due impulsi possono essere rappresentati, ricordando il corso di sistemi, dalla funzione coseno. Se u_0 è costante e Δu_0 tende a zero, allora si ha un coseno.

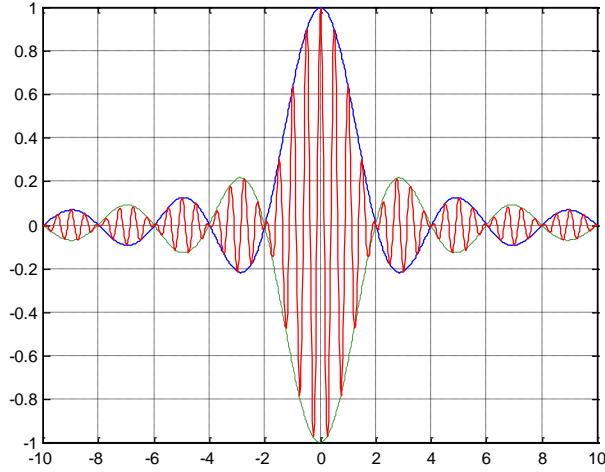


Figura 42: Rappresentazione grafica del filtro passa banda ideale 1D nel **dominio del tempo**.

⁶Si ricorda che la funzione sinc è così formata: $\text{sinc}(x) = \frac{\sin(\pi x)}{\pi x}$

Un **filtro ferma banda ideale** sopprime tutte le frequenze nell'intervallo specificato. Più formalmente, esso è rappresentato con la seguente funzione:

$$G(\mu) = \begin{cases} 0 & \text{se } u_1 < |u| < u_2 \\ 1 & \text{altrimenti} \end{cases}$$

Le variabili u rappresentano l'intervallo (come si vede in figura) e al di fuori di esso il valore è uno. Quindi, al contrario del filtro passa banda ideale, il ferma banda non consente al segnale di esistere nelle frequenze specificate.

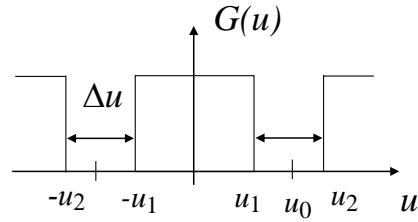


Figura 43: Rappresentazione grafica del filtro ferma banda ideale nel **dominio delle frequenze**.

Dalla rappresentazione grafica si ricava anche la relativa **funzione di trasferimento nel dominio delle frequenze**, ovvero la rimozione (segno meno) di due box traslate nel tempo ottenibili eseguendo la convoluzione di una box per due impulsi unitari (Dirac):

$$G(u) = 1 - \Pi\left(\frac{u}{\Delta u}\right) * [\delta(u - u_0) + \delta(u + u_0)]$$

I valori $u_0, \Delta u$ hanno lo stesso significato del filtro passa banda ideale. Invece, la **funzione di trasferimento nel dominio del tempo** è la seguente:

$$g(t) = \delta(t) - 2\Delta u \cdot \frac{\sin(\pi t \Delta u)}{\pi t \Delta u} \cdot \cos(2\pi u_0 t)$$

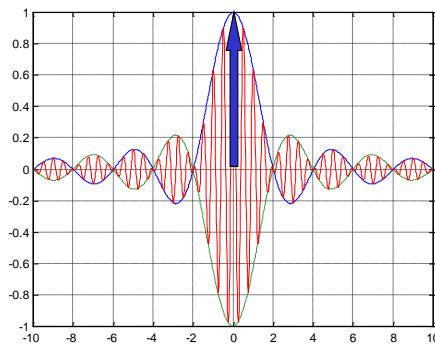


Figura 44: Rappresentazione grafica del filtro ferma banda ideale nel **dominio del tempo**.

5 Elaborazione di immagini - Estrazione di edge robusta

5.1 Estrazione di contorni (*edge*)

Si definisce come **feature** una proprietà dell'immagine. Quest'ultima può essere:

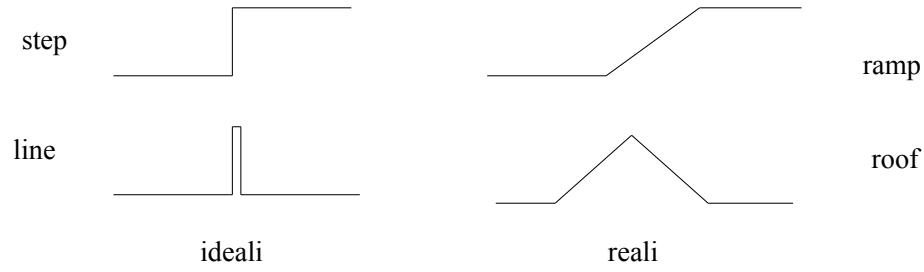
- **Globale**, ovvero dell'intera immagine;
- **Locale**, ovvero di una parte dell'immagine con speciali proprietà.

Con il termine **edge** si indica una forte variazione locale dei livelli di grigio, tipicamente associata ad un contorno o ad una separazione tra regioni.

Il fenomeno dell'*edge* provoca una serie di complicazioni. Si definisce il **problema dell'estrazione**: data un'immagine, la **rilevazione** (*detection*) e la **localizzazione** degli edge derivati dalla scena e non dal rumore.

La soluzione al problema non è affatto banale perché: il rumore può mascherare gli *edge* e possono esistere falsi *edge* positivi dovuti a ombre e variazioni di luminosità.

Esistono varie **tipologie di livello di grigio**:



Si definisce un **punto di edge**, un punto dell'immagine con coordinate (i, j) in cui è presente una variazione significativa locale dell'intensità.

Si definisce un **contorno**, una lista di punti di *edge* o un'altra struttura dati che modella i punti di *edge*, ordinandoli dal primo all'ultimo. In altre parole, è l'*edge* vero e proprio.

Ancora una volta, nasce una complicazione: il **problema della localizzazione**. Infatti, nonostante la posizione di un punto di *edge* si identifica tramite le coordinate (i, j) , le costanti i, j potrebbero non essere indici di pixel.

Inoltre, la localizzazione può richiedere anche l'**orientazione** di un punto d'*edge*. Questo implica l'introduzione del **gradiente**.

Per concludere, si **classificano gli edge**:

- **Corretti:** *edge* della scena;
- **Falsi:** non esistono nella scena (*false positives*);
- **Mancanti:** presenti nella scena, ma non rilevanti (*false negatives*).

5.2 Criteri per *edge detector* ottimi

Esistono 3 criteri per ottenere un *edge detector* ottimo:

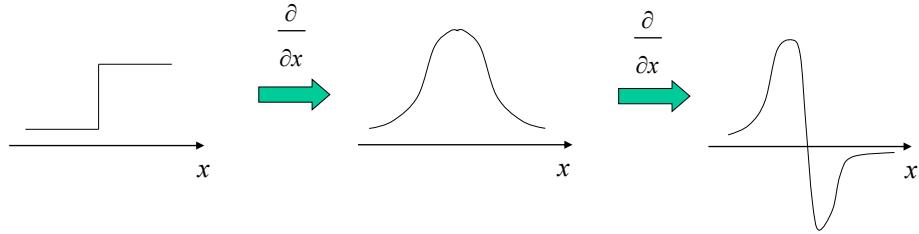
- **Buon rilevamento (*detection*)**. Deve minimizzare la probabilità di presenza di falsi positivi e la probabilità di presenza di falsi negativi. La massimizzazione del SNR (*Signal to Noise Ratio*, rapporto segnale-rumore; è una grandezza numerica che mette in relazione la potenza del segnale utile rispetto a quella del rumore⁷⁾) definito come il rapporto tra la risposta quadratica media (RMS) del filtro agli *step edge* ideali e il rumore;
- **Buona localizzazione**. Gli *edge* rilevati devono essere più vicino possibile agli *edge* reali;
- **Vincolo di singola risposta**. Per ogni punto di *edge* reale, il filtro deve rilevare un solo punto di *edge*. Ossia, minimizzare il numero di massimi locali creati dal rumore intorno all'*edge*.

I criteri di buon rilevamento e di buona localizzazione, identificano due modi di valutazione per l'estrazione di *edge* indipendentemente dalla lunghezza dell'*edge* e dalla potenza del rumore.

⁷Fonte: [WikipediaIT](#)

5.3 Il gradiente

Il **gradiente** è una misura della variazione in una funzione. Nel caso monodimensionale, corrisponde al picco della derivata prima:



Assumendo una funzione $f(x, y)$, allora:

$$\mathbf{G}[f(x, y)] = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

Il modulo è pari alla lunghezza del vettore G localizzato su x e y :

$$\|\mathbf{G}\| = G = \sqrt{G_x^2 + G_y^2}$$

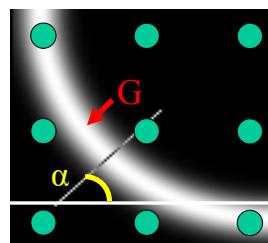
Approssimazioni del modulo:

$$\begin{aligned} G &\cong |G_x| + |G_y| \\ G &\cong \max(|G_x|, |G_y|) \end{aligned}$$

L'orientazione, misurata rispetto all'asse x , è data da:

$$\alpha(x, y) = \arctan \frac{G_y}{G_x}$$

In pratica, il vettore \mathbf{G} punta nella direzione di massimo scarto della funzione f :



5.4 Estrazione dell'edge

5.4.1 Il gradiente discreto

Per differenziare un'immagine digitale $F[x, y]$ è possibile seguire due strade:

1. Ricostruire un'immagine continua, quindi prendere il gradiente analitico;
2. Prendere la derivata discreta, tramite “differenze finite”, come visto nei capitoli precedenti.

Il **gradiente discreto** è approssimativamente la differenza tra pixel adiacenti:

$$\frac{\partial f}{\partial x}[x, y] \approx F[x+1, y] - F[x, y]$$

E le approssimazioni numeriche, invece, sono:

$$\begin{aligned} f_x &\approx f[i, j+1] - f[i, j] \longrightarrow [-1 \quad 1] \longrightarrow S_x = \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix} \\ f_y &\approx f[i, j] - f[i+1, j] \longrightarrow \begin{bmatrix} 1 \\ -1 \end{bmatrix} \longrightarrow S_y = \begin{bmatrix} 1 & 0 \\ -1 & 0 \end{bmatrix} \end{aligned}$$

5.4.2 Operatore di Roberts

L'**operatore di Roberts** è un'approssimazione con differenze incrociate:

$$|d[f(i, j)]| = |f(i, j) - f(i+1, j+1)| + |f(i, j+1) - f(i+1, j)|$$

La convoluzione dell'immagine (guardare le due matrici S_x e S_y nel paragrafo 5.4.1) con 2 maschere del tipo:

$$S_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad S_2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

5.4.3 Operatore di Prewitt

Data l'approssimazione con vicinato 3×3 :

$$\begin{bmatrix} a_0 & a_1 & a_2 \\ a_7 & (i, j) & a_3 \\ a_6 & a_5 & a_4 \end{bmatrix}$$

L'**operatore di Prewitt** è un'approssimazione del tipo:

$$|r[f(i, j)]| = \underbrace{|(a_2 + a_3 + a_4) - (a_0 + a_7 + a_6)|}_{S_x} + \underbrace{|(a_0 + a_1 + a_2) - (a_6 + a_5 + a_4)|}_{S_y}$$

Il risultato:

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

5.4.4 Operatore di Sobel

L'**operatore di Sobel** ha molte matrici predefinite, tra le più famose:

$$S_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad S_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

5.4.5 Operatore di Kirsch

L'**operatore di Roberts** ha 8 maschere direzionali, ognuna delle quali risponde massimamente per ciascun *edge* in una determinata direzione. Il **valore massimo** tra gli 8 valori stimati è il valore di output del gradiente in quel punto, mentre l'indice della maschera codifica le direzioni dell'*edge*.

Un esempio di maschere direzionali:

Horizontal line: $D_0 = \begin{bmatrix} -1 & -1 & -1 \\ 2 & 2 & 2 \\ -1 & -1 & -1 \end{bmatrix}$

45 line: $D_{45} = \begin{bmatrix} -1 & -1 & 2 \\ -1 & 2 & -1 \\ 2 & -1 & -1 \end{bmatrix}$

Vertical line: $D_{90} = \begin{bmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix}$

135 line: $D_{135} = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$

5.4.6 Operatore di Robinson

L'**operatore di Roberts** ha 8 maschere direzionali, simili a quelle di Kirsch.

5.4.7 Confronto tra operatori

Dati gli operatori per estrarre l'*edge*:

1. Gradiente
2. Roberts
3. Prewitt
4. Sobel

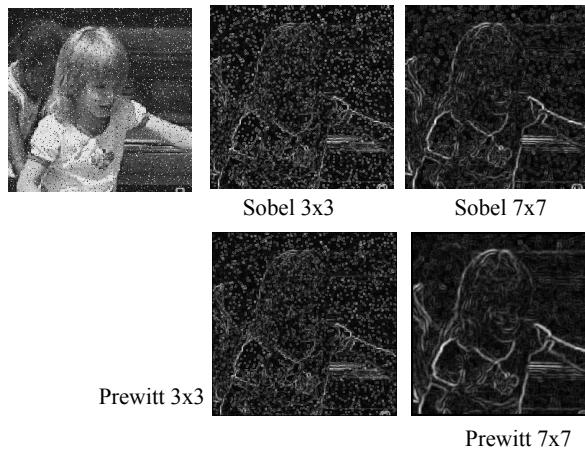
È possibile osservare come più si è **vicini al gradiente**, più:

- La localizzazione è buona;
- Aumenta la sensibilità al rumore;
- Diminuisce la rilevazione, diventa scarsa con il gradiente.

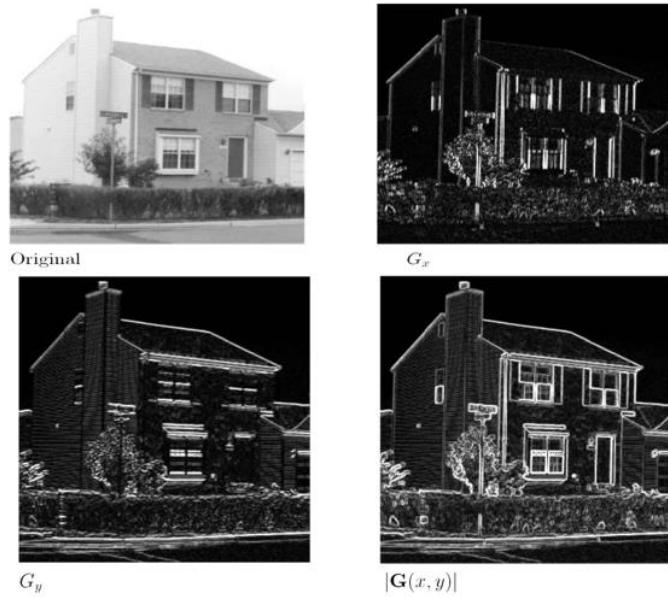
Al contrario, **vicino all'operazione di Sobel**:

- Diminuisce la localizzazione;
- Diminuisce la sensibilità al rumore;
- Aumenta la rilevazione.

Un esempio di applicazione dei vari operatori:



Invece, l'applicazione del gradiente è visibile nella seguente figura. Interessante notare le singoli applicazioni, ovvero solo lungo x o y :



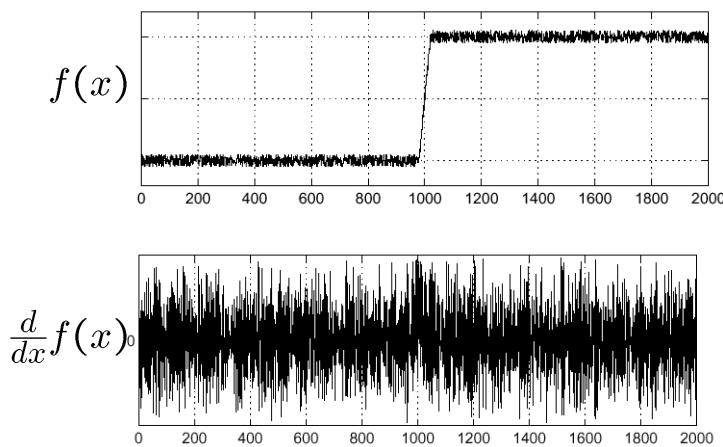
5.5 Derivata prima

5.5.1 Problemi

Gli operatori di **derivata prima** vanno applicati per ottenere il gradiente. Tuttavia, presentano dei problemi poiché **estraggono troppi punti di edge**.

5.5.2 Effetti del rumore

Si consideri una singola riga o colonna dell'immagine. Si traccia l'intensità in funzione della posizione del segnale. Si può vedere come la **derivata prima amplifichi il rumore**:



Per **risolvere** questo problema, si applica uno *smoothing* preliminare.

5.5.3 Rilevazione di *edge* o *detection* robusta

Per rilevare l'*edge* o eseguire una detection robusta, si eseguono tre fasi fondamentali:

- **Filtraggio** (*noise smoothing*): si vuole ridurre il rumore preservando gli *edge* reali;
- **Rinforzo** (*enhancement, sharpening*): facilita il rilevamento dell'*edge*, evidenziando le variazioni locali;
- **Rilevamento** (*detection*): decisione sulla presenza o meno dei punti di *edge*.

Tuttavia, si tenga in considerazione che:

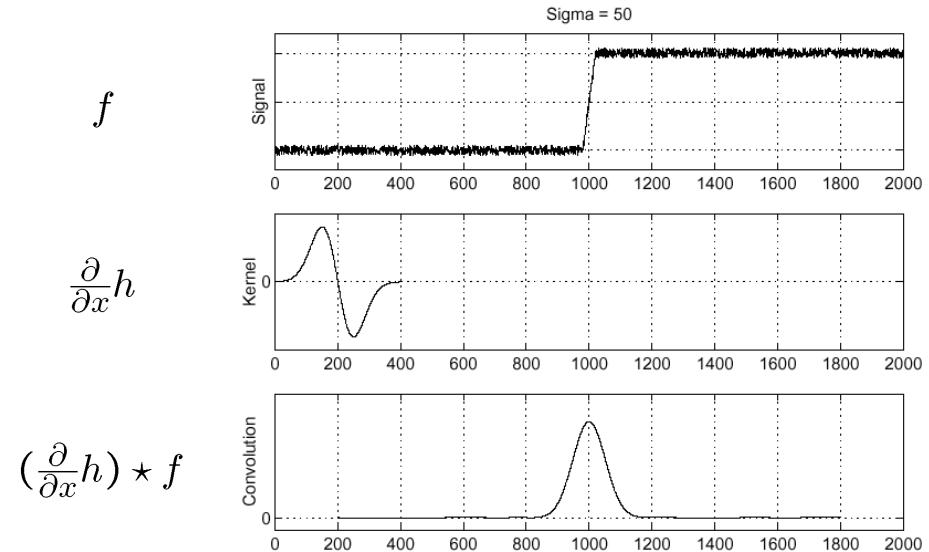
- ☛ Non tutti i punti aventi gradiente diverso da zero sono *edge*;
- ☛ È richiesto tipicamente una fase di sogliatura;
- ☛ È richiesto talvolta una fase di localizzazione a risoluzione *subpixel* e una fase di assottigliamento (*thinning*)

5.5.4 Teorema di convoluzione

Applicando il **teorema di convoluzione**:

$$\frac{\partial}{\partial x} (h * f) = \left(\frac{\partial}{\partial x} h \right) * f$$

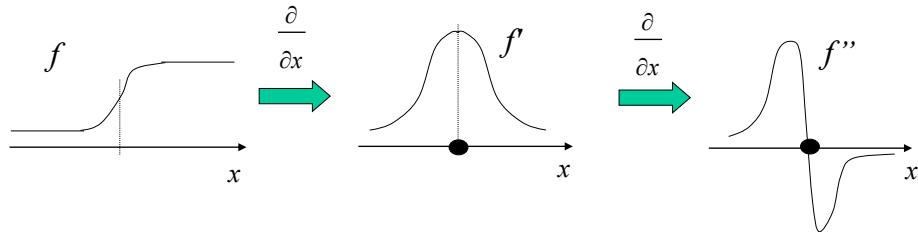
In questo modo, si risparmia un'operazione di filtraggio. Esempio:



5.6 Derivata seconda

5.6.1 Operatori di derivata seconda

Gli **operatori di derivata seconda** rilevano il punto di *edge* mediante il passaggio per lo zero (*zero crossing*) del valore calcolato, e per questo motivo sono **più precisi** (non necessitano di stabilire una soglia).



Nell'immagine è visibile l'applicazione delle varie derivate. Con la derivata prima si evidenzia la massima discontinuità tra i pixel antistanti, mentre con la derivata seconda c'è la massima variazione, ovvero il valore è prossimo allo zero.

5.6.2 Operatore Laplaciano

L'**operatore Laplaciano** equivale alla derivata seconda in 2 dimensioni:

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

Approssimando con l'equazione alle differenze e centrando la stima nel pixel (i, j) si ha:

$$\frac{\partial^2 f}{\partial x^2} = f(i, j+1) - 2f(i, j) + f(i, j-1)$$

$$\frac{\partial^2 f}{\partial y^2} = f(i+1, j) - 2f(i, j) + f(i-1, j)$$

Combinando in un singolo operatore, si nota che è isotropico:

$$\nabla^2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\nabla^2 = \begin{bmatrix} 1 & 4 & 1 \\ 4 & -20 & 4 \\ 1 & 4 & 1 \end{bmatrix}$$

$$\nabla^2 = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

$$\nabla^2 = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

5.7 Laplaciano di Gaussiana

L'**operatore Laplaciano di Gaussiana** consiste in un filtraggio Gaussiano e Laplaciano. È il metodo migliore per estrarre gli *edge*. Con questo operatore si può osservare che:

- Il filtro di *smoothing* è Gaussiano;
- La fase di rinforzo è la derivata seconda, cioè Laplaciano;
- Il rilevamento è la fase di attraversamento per lo zero;
- La localizzazione è fatta mediante interpolazione lineare.

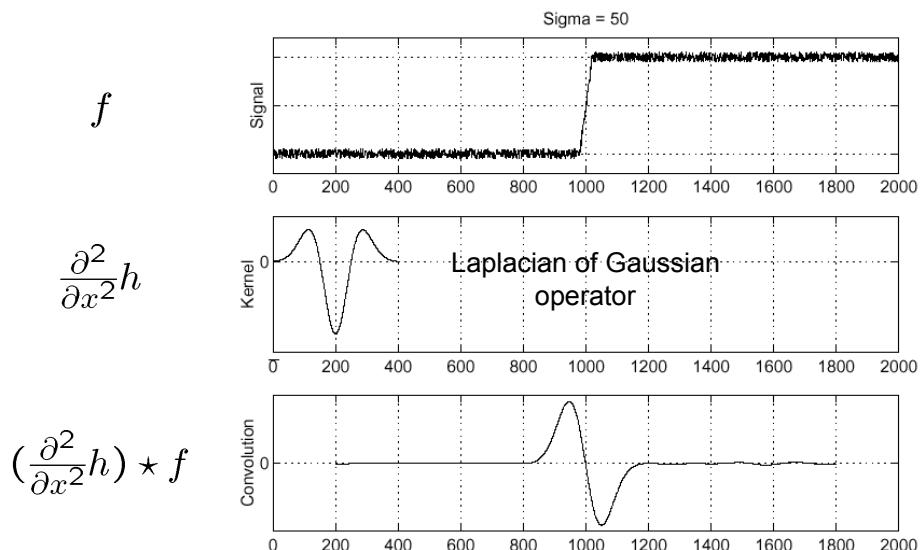
L'**applicazione** di questo filtro Gaussiano ha i seguenti fini:

- **Riduzione del rumore** per punti isolati e piccole configurazioni;
- “**Addolcimento** degli *step edge*”;
- “**Diffondere (blur)** gli *edge*”.

Considerando:

$$\frac{\partial^2}{\partial x^2} (h * f)$$

I grafici sono:



L'operatore laplaciano viene indicato con la lettera nabla al quadrato ∇^2 :

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

Considerando *edge* soltanto i punti in cui:

- Laplaciano passa per lo zero;
- Gradiente locale è massimo (maggiore di una soglia quindi).

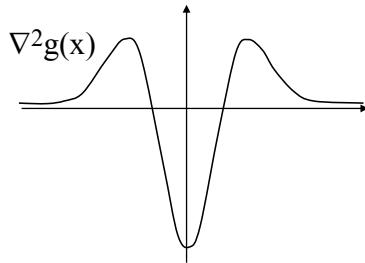
Allora l'uscita dell'operatore Laplaciano di Guassiana è $h(x, y)$:

$$h(x, y) = \nabla^2 [g(x, y) * f(x, y)] = [\nabla^2 g(x, y)] * f(x, y)$$

Dove:

$$\nabla^2 g(x, y) = \frac{x^2 + y^2 - 2\sigma^2}{2\sigma^4} \cdot \exp\left[-\frac{x^2 + y^2}{2\sigma^2}\right]$$

Quest'ultimo viene chiamato ***mexican hat*** poiché ha la forma di cappello messicano:



Adesso, grazie all'introduzione di questo operatore, le due seguenti operazioni sono **equivalenti**:

1. Eseguire la convoluzione dell'immagine con un filtro Gaussiano di *smoothing* e calcolare il Laplaciano del risultato;
2. Eseguire la convoluzione dell'immagine con un filtro lineare, ovvero il filtro Laplaciano di Guassiana.

Si introduce un **esempio** di maschera 5×5 :

$$\begin{bmatrix} 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & -2 & -1 & 0 \\ -1 & -2 & 16 & -2 & -1 \\ 0 & -1 & -2 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 \end{bmatrix}$$

Data la formula del filtro, con diversi valori della varianza σ , si ottengono diversi risultati:

- **σ grande:**

- ✓ Riduzione del rumore migliore;
- ✗ Perdita notevole di informazioni utili
- † Esempi: fusione di edge vicini, presenza di edge falsi, localizzazione imprecisa

- **σ piccolo:**

- ✓ Riduzione del rumore minore;
- ✗ Riduzione di perdita di informazioni utili
- † Esempio: molti edge spuri

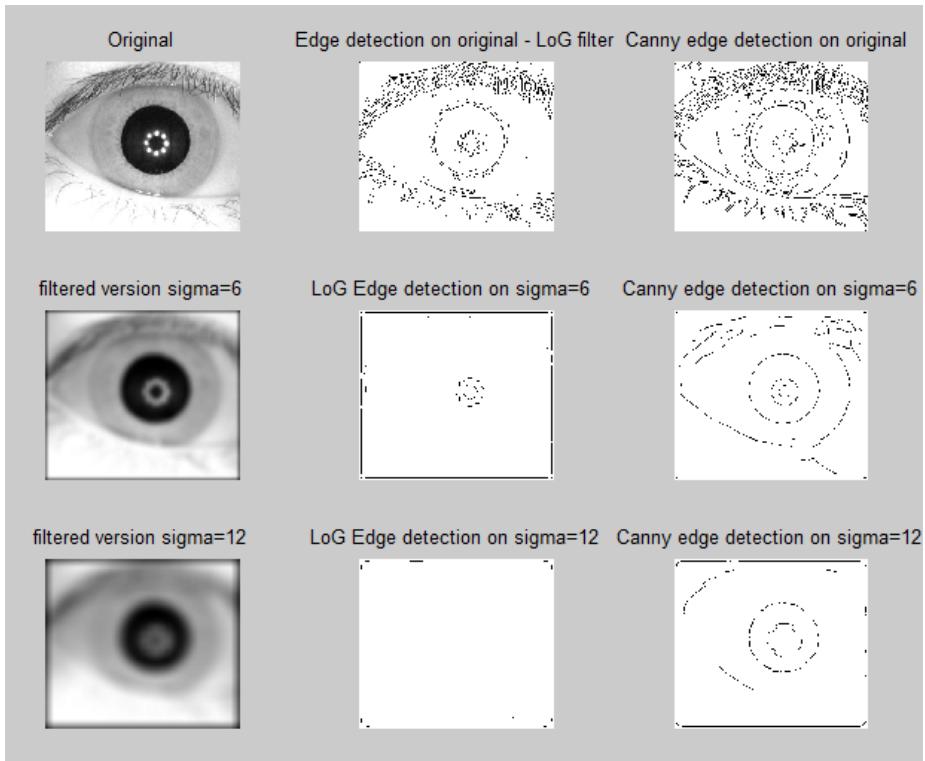
Per attenuare i difetti e cercare di ottenere un risultato ottimale, una **soluzione** comune è il **filtraggio a diversa scala** (diverse varianze):

- **σ grande:** pochi edge robusti, ma mal localizzati;
- **σ piccolo:** miglior localizzazione degli edge robusti.

Al **termine di un'applicazione di filtro Laplaciano di Gaussiana**, è necessario fondere i risultati delle varie scale ottenute, secondo alcuni criteri. Per esempio:

- Se ci sono *edge* in tutte le scale;
- Se ci sono *edge* almeno su K scale;
- E così via.

L'esempio nella prossima pagina, dovrebbe chiarire maggiormente la teoria di questo filtro:



5.8 Gradiente e Laplaciano di Gaussiana a confronto

L'operatore **gradiente** è efficace nel caso di immagini con poco rumore e transizioni nette.

Al contrario, l'operatore **Laplaciano di Gaussiana**, insieme allo *zero crossing*, offre un'alternativa quando il rumore è elevato e le transizioni non sono nette. Infatti, lo *zero crossing* viene aggiunto poiché offre un modo affidabile per localizzare l'edge. Mentre le proprietà di *smoothing* del Laplaciano di Gaussiana riducono gli effetti del rumore.

5.9 Filtro di Canny

5.9.1 Definizione

Il **filtro di Canny** è uno degli strumenti più efficienti. È il contrapposto all'operazione di Laplaciano di Gaussiana e viene considerato il “più esoso computazionalmente”. L'**algoritmo** consiste in 3 passi:

1. **Blur**, ovvero il filtro di rinforzo;
 2. **Edge thinning**, ovvero *non maxima suppression*;
 3. **Sogliatura ad isteresi + edge linking**.
-

5.9.2 Modello di *edge* per Canny

Il **modello di *edge* per Canny** è lo *step edge*, i cui descrittori sono:

- **Posizione o centro**, la posizione in cui l'*edge* è rilevato lungo la perpendicolare all'*edge* (solitamente memorizzata come un'immagine binaria);
- **Normale o orientazione o direzione**, versore della massima variazione di intensità nel punto di *edge* (perpendicolare al contorno di cui l'*edge* fa parte);
- **Forza**, modulo gradiente.

Il **modello di *step edge*** è il seguente:

$$G(x) = \begin{cases} 0 & x < 0 \\ A & x \geq 0 \end{cases}$$

Le ipotesi di base sono tre:

- Il filtro di rinforzo è lineare;
 - Il filtro deve essere ottimo per gli *edge* con rumore;
 - Il rumore è additivo, bianco e Gaussiano.
-

5.9.3 Algoritmo di Canny

L'**algoritmo di Canny** si divide in tre passaggi. Dato in ingresso l'immagine I :

1. Si applica il **filtro di rinforzo** ad I ;
2. Si applica l'algoritmo **Non-Maxima Suppression** all'output del passaggio precedente;
3. Si applica la **Sogliatura ad Isteresi** all'output del passaggio precedente.

Il filtro di rinforzo, la non-maxima suppression e la sogliatura ad isteresi, saranno argomenti trattati nei prossimi paragrafi.

5.9.4 Filtro di Rinforzo

Il **filtro di rinforzo** per essere applicato necessita di una serie di passaggi:

1. Si applica all'immagine I il filtro di *smoothing* Gaussiano:

$$J = G * I$$

Ovvero la convoluzione, con G che è la Gaussiana con media nulla e varianza σ .

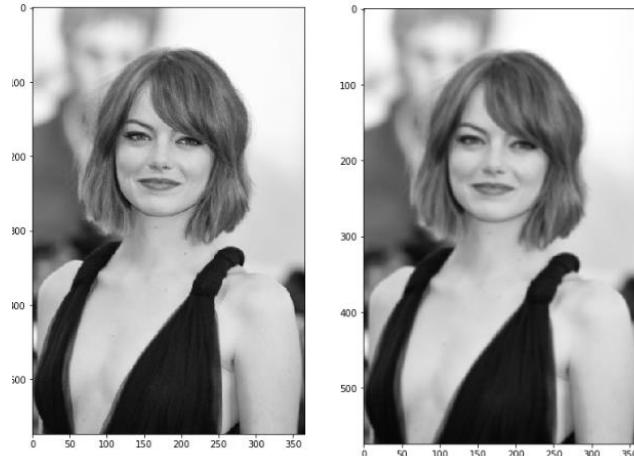
2. Si calcola il gradiente $dJ = [J_x, J_y]$ e se ne stima il modulo e direzione:

$$e_s(i, j) = \sqrt{J_x^2(i, j) + J_y^2(i, j)} \implies \text{img } e_s$$

$$e_o(i, j) = \arctan \frac{J_y(i, j)}{J_x(i, j)} \implies \text{img } e_o$$

Il valore della varianza σ da utilizzare, dipende dalla lunghezza dei contorni, dal livello di rumore e dal compromesso tra localizzazione e rilevamento.

Esempio di immagine con $\sigma = 1$:



Esempio di immagine con $\sigma = 1.3$:



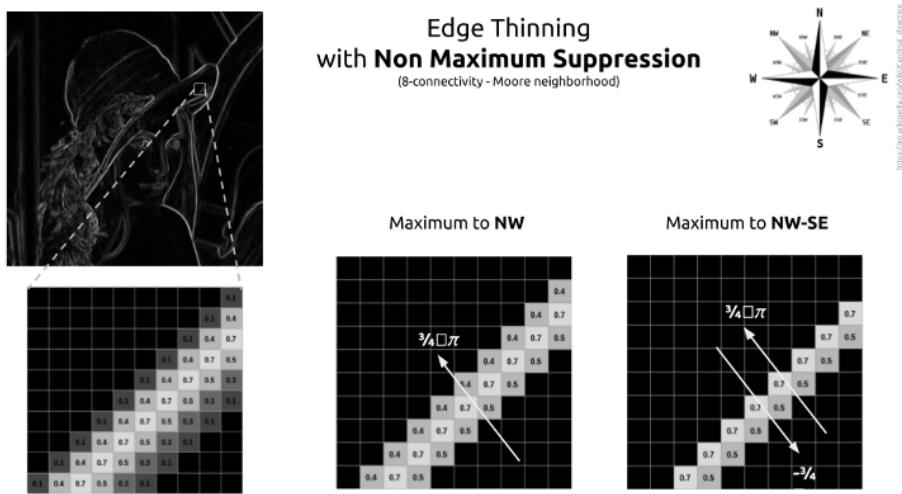
5.9.5 Non-Maxima Suppression

Il **filtro Non-Maxima Suppression** produce *edge* di un solo pixel scartando i massimi locali.

Il suo algoritmo è così strutturato:

- Riceve in input:
 - e_s , ovvero l'immagine della forza degli *edge*;
 - e_o , ovvero l'immagine con il gradiente orientato;
 - d_1, d_2, d_3, d_4 , che rappresentano le direzioni degli *edge* (per esempio, 0, 45, 90, 135).
- Per tutti i pixel con coordinate (i, j) :
 - Trovare la direzione d_k che approssima meglio la direzione del gradiente in quella posizione $e_o(i, j)$;
 - Se $e_{i,j}$ è minore di almeno uno dei suoi due vicini lungo d_k , allora assegna al risultato $I_N(i, j) = 0$, altrimenti $I_N(i, j) = e_s(i, j)$.
- Il risultato (output) è $I_N(i, j)$, ovvero l'immagine con *edge* più sottili (così da evitare *edge* troppo vicini o antistanti).

Un esempio di applicazione:



Partendo dall'immagine di sinistra e andando verso l'immagine di destra, viene eseguita un'eliminazione dei valori più piccoli (spiegato in modo approssimativo, si guardi l'algoritmo per capire meglio).

5.9.6 Sogliatura ad isteresi

La **sogliatura ad isteresi** non è altro che un confronto tra diverse soglie. Tuttavia, questa operazione nasce da un problema.

Il problema della sogliatura: l'immagine creata dalla soppressione dei non massimi (Non-Maxima Suppression, paragrafo precedente), contiene ancora punti spuri creati dal rumore.

Solitamente, viene definita una soglia TH (*threshold high*) e si scartano i pixel la cui forza è minore di TH , ma questo comporta degli svantaggi a seconda della soglia:

- Con **soglia piccola**, vengono rilevati anche gli *edge* deboli e di conseguenza si ottengono anche massimi spuri (falsi contorni);
- Con **soglia grande**, i valori dei veri massimi lungo i contorni possono fluttuare intorno alla soglia, frammentando così il contorno.

Ecco che la **soluzione** che si applica è la **soglia ad isteresi**.

L'**algoritmo** è il seguente:

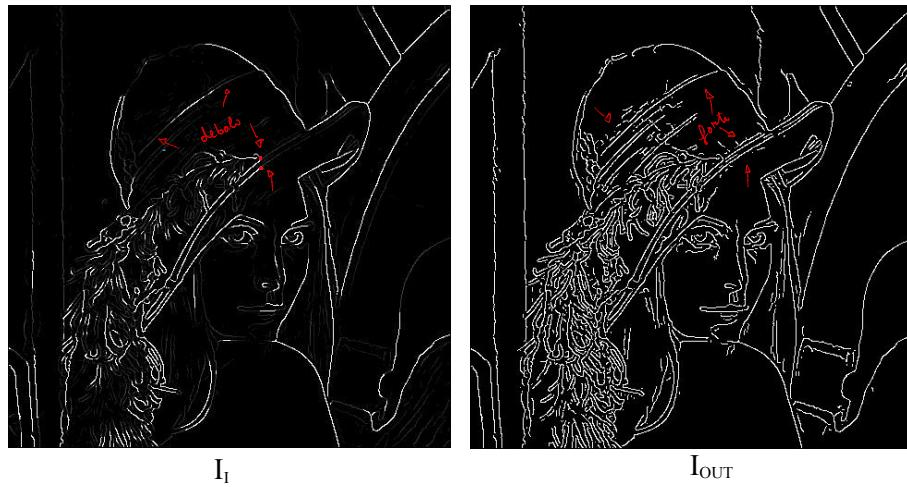
- Riceve in input:
 - I_N , ovvero l'immagine di uscita dell'algoritmo di Non-Maxima Suppression (paragrafo precedente).
- Vengono scelti due valori: soglia forte (TH , *threshold high*) e soglia debole (TL , *threshold low*)
- Viene analizzato ogni punto di *edge* a seconda che esso sia:
 - $> TH$, che corrisponde ad un punto di *edge* forte;
 - $\geq TL$ and $\leq TH$, che corrisponde ad un punto di *edge* debole;
 - $< TL$, che corrisponde ad un punto di *edge* ininfluente e da non considerare.
- Il risultato (output) è l'immagine I_I .

5.9.7 Edge linking

L'operatore **edge linking** viene utilizzato nell'algoritmo di Canny e ha il seguente algoritmo:

- Per tutti i punti di *edge* ottenuti dalla sogliatura ad isteresi, quindi dall'immagine I_I :
 - Si localizza il prossimo punto di *edge* $I_I(i, j)$ tale che $I_I(i, j) > TH$;
 - Partendo da $I_I(i, j)$ si segue la catena di massimi locali connessi nelle due direzioni perpendicolari alla normale dell'*edge* finché $I_I(i, j) > TL$ (memorizzando i risultati).
- Il risultato finale è l'immagine I_{OUT} e un insieme di liste, ognuna descrivente la posizione degli *edge*, la forza e l'orientazione dell'*edge* linkato a cui si fa riferimento.

Un **esempio** di *edge linking* con delle piccole frecce rosse per indicare gli effetti. A sinistra, l'immagine dopo la sogliatura ad isteresi e a destra l'immagine dopo l'*edge linking*:



5.10 Prestazioni di un *edge detector*

Per misurare le **prestazioni di un edge detector**, si seguono alcuni **criteri**:

- Probabilità di falsi *edge* (*false positives*);
- Probabilità di *edge* mancanti (*false negatives*);
- Errore nella stima dell'orientazione;
- Distanza media al quadrato della stima dal vero *edge*.

I **metodi** di valutazione sono principalmente due: (1) si eseguono delle prove sulle immagini sintetiche e sul rumore noto, (2) molti parametri da considerare.

5.11 Figura di merito per gli *edge*

Inoltre, si esegue una **misura quantitativa e oggettiva**, valida per ogni *edge detector*, così da ottenere una **figura di merito per gli *edge***.

Infine, esistono tre **tipi di errore**:

- *Edge* validi mancanti;
- Errore nella localizzazione;
- Classificazione di rumore come *edge*.

La **formula** della figura di merito è:

$$F_M = \frac{1}{\max(I_A, I_I)} \sum_{i=1}^{I_A} \frac{1}{1 + d_i \alpha^2}$$

Con:

- I_A , *edge* rilevati
- I_I , *edge* ideali
- d , distanza minima tra rilevati e ideali
- α , costante di penalizzazione

5.12 Trasformata di Hough

5.12.1 Definizione

La **trasformata di Hough** viene utilizzata dopo aver estratto gli *edge* da un'immagine ed è conosciuta come una tecnica di *object detection*⁸, basata su criteri geometrici.

Dato che l'individuazione di un oggetto è difficile per vari motivi, questa trasformata propone una soluzione. Infatti, trasforma il problema dalla ricerca di una curva, ad una banale **ricerca di massimi**.

In generale, una curva piana è definita in forma analitica tramite un insieme di parametri. Un'equazione lega tale parametri alle coordinate cartesiane con la seguente relazione:

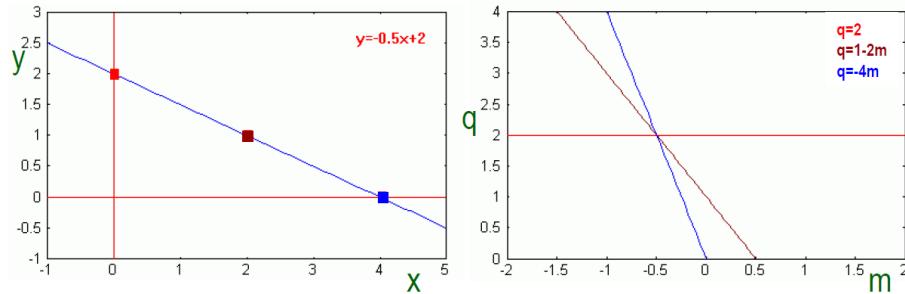
$$f((x, y), (a_1, a_2, \dots, a_n)) = 0$$

In cui (x, y) è un punto della curva nello **spazio immagine** (SI), e (a_1, a_2, \dots, a_n) è una n -upla di valori che individuano un punto nello **spazio dei parametri** (SP). Un punto nello spazio dei parametri individua univocamente una curva analitica.

Per **esempio**, la **retta** $y = mx + q$ può essere riscritta come:

$$f((x, y), (m, q)) = y - mx - q = 0$$

Quindi, fissato un punto (x_i, y_i) nello spazio immagine SI, l'equazione $q = y_i - mx_i$ descrive un insieme di curve (che rimangono ancora una retta) nello spazio dei parametri SP. Graficamente:



⁸Per *object detection* si intende l'individuazione precisa della classe di un oggetto in un'immagine

I **cerchi di raggio noto**, generano uno spazio dei parametri SP bidimensionale e la curva generata da ogni punto nello spazio immagine è essa stessa un cerchio:

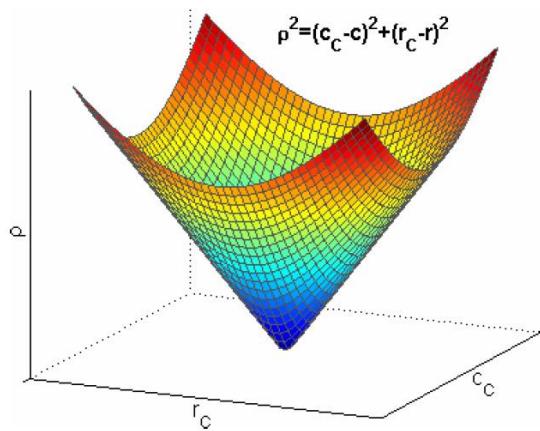
$$(y - y_c)^2 + (x - x_c)^2 = r^2$$

Lo spazio dei parametri SP è generato dalle coordinate del centro (x_c, y_c) .

Nel caso in cui il **raggio è incognito**, lo spazio dei parametri SP è tridimensionale:

$$f((c, r), (c_c, r_c), \rho) = (r - r_c)^2 + (c - c_c)^2 - \rho^2 = 0$$

In questo caso, la figura generata è un cono:



Ricapitolando:

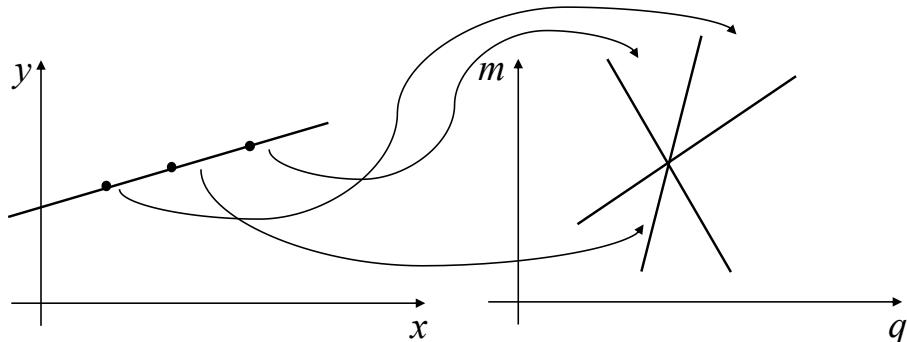
- ✓ Ogni punto nello spazio immagine SP genera una curva o superficie nello spazio dei parametri SP;
- ✓ Nello spazio immagine SP, un'intersezione di molte curve o superfici è l'indizio della presenza di una particolare istanza della curva analitica cercata;
- ✓ In generale, occorre un numero di punti almeno pari al numero dei parametri per individuare una curva, cioè le sue caratteristiche;
- ✓ La soglia alta *HT* permette quindi di convertire un problema di ricerca di curve in quello più semplice di ricerca di intersezioni.

5.12.2 Trasformata per le rette

Un punto P vota per il fascio di rette che passano per P . Due punti P e Q votano ciascuno per un fascio diverso, e vi è una sola retta in comune ai due fasci (che prende due voti): quella passante per P e Q .

I parametri che identificano una retta sono m e q nell'equazione $y = mx + q$ e un fascio è individuato da un insieme di valori m e q . Dunque, un punto $P = (x, y)$ vota per un insieme di valori m, q che rappresentano il fascio di rette passanti per P .

Il punto P vota per tutti i valori di p e q che soddisfano la solita equazione della retta $y = mx + q$ con x, y fissati. Il punto P vota per la retta nello spazio m, q dove il coefficiente angolare e il termine noto sono dati da x e y . Quindi, si passa da una rappresentazione nello spazio cartesiano x, y ad un'altra rappresentazione nello spazio dei parametri m, q :

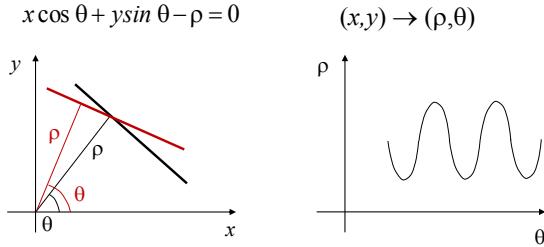


Nel caso n punti giacenti su una retta, tutti vengono mappati in un insieme di rette nello spazio dei parametri, tutte passanti per il punto (m, c) che quindi sarà maggiormente votato. A questo punto, il **problema si trasforma da line detection a peak detection** nello spazio dei parametri.

Tuttavia, questo implica:

- Delimitazione e discretizzazione dello spazio dei parametri;
- La ricerca di massimi assoluti per trovare le rette;
- Soppressione dei massimi locali per far fronte a picchi spuri nell'intorno di quelli veri.

In parole poche, viene **utilizzata la forma polare** perché più comoda, dato che: θ è delimitato, ρ si può delimitare specificando una regione circolare attorno all'origine.



Quindi, ad un fascio di rette passante per x, y , corrisponde una curva sinusoidale nello spazio dei parametri SP.

Per concludere, l'**algoritmo** corrisponde a:

- In input riceve l'immagine binaria $M \times N$;
- Viene definito lo spazio dei parametri SP (1 *edge* e 0 *no edge*):

$$\rho \in [0, \sqrt{M^2 + N^2}], \quad \theta \in [0, \pi] \rightarrow 0, 180$$

- Viene eseguita la discretizzazione di (ρ, θ) in (ρ_d, θ_d) usando come “passo” un valore accettabile a seconda del problema da risolvere (più precisione o più computazione);
- Sia A la matrice $R \times T$ risultante, inizializzata a zero;
- Allora, per ogni pixel $I(x, y) = 1$ e per $h = 1, \dots, T$:
 - Sia $\rho = x \cos(\theta_d(h)) + y \sin(\theta_d(h))$
 - Si trova l'indice k tale che $\rho_d(k)$ è l'elemento più vicino a ρ ;
 - Viene incrementato $A(k, h)$ di un'unità.
- Infine, vengono trovati tutti i massimi locali (k_ρ, h_ρ) tale che $A(k_\rho, h_\rho) > \tau$ soglia.

L'output finale è un insieme di coppie $[\rho_d(k_\rho), \theta_d(h_\rho)]$ che identifica le rette.

Alcune **considerazioni**:

- ✓ La HT (*high threshold*) non richiede una fase di *linking* o raggruppamento preliminare;
- ✓ È robusta al rumore;
- ✗ Data la presenza del rumore e dell'inesistenza della discretizzazione, non esiste un picco ben marcato. Di conseguenza, è necessario un piccolo insieme di accumulatori con valore elevato e viene usato il calcolo del centroide;
- ✗ Lo spazio dei parametri SP cresce col numero dei parametri, quindi computazionalmente oneroso per modelli complessi (risolvibile parzialmente con l'uso del gradiente).

5.12.3 Trasformata per i cerchi

L'equazione del cerchio in forma implicita è:

$$(x - a)^2 + (y - b)^2 = r^2$$

Con a, b coordinate del centro e r il raggio.

Lo spazio dei parametri SP è multidimensionale.

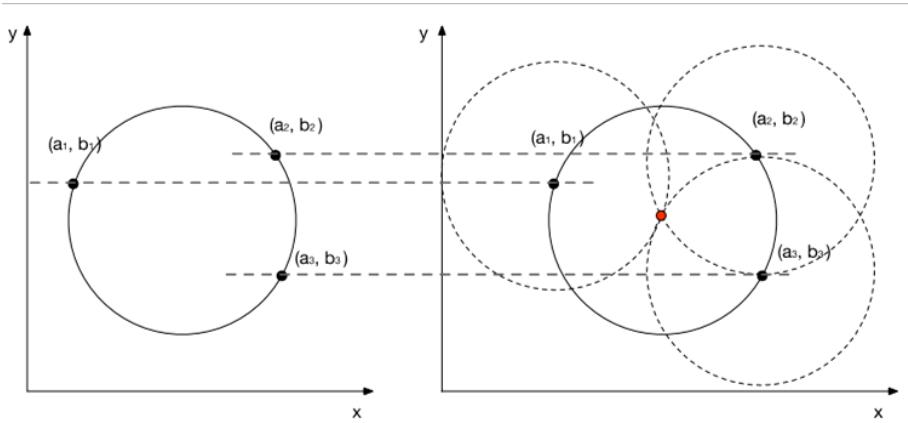
Talvolta, l'**equazione parametrica** del cerchio attraverso seni e coseni permette di rendere la trasformata **più semplice** da trattare:

$$\begin{cases} r = \frac{x - a}{\cos \theta} \\ b = y - (x - a) \tan \theta \end{cases}$$

Per **esempio**, se si è a conoscenza del raggio del cerchio r con a, b coordinate del centro:

$$(x - a)^2 + (y - b)^2 = r^2$$

Ogni punto genera un cerchio e intersezioni tra cerchi nello spazio dei parametri SP indicando il probabile centro.



Attenzione! Se non si è a conoscenza del raggio del cerchio r , ma si è in grado di **calcolare il gradiente del cerchio** nel punto:

$$b = a \tan \theta - x \tan \theta + y$$

Che costituisce una linea nello spazio dei parametri. Inoltre, le intersezioni tra rette negli spazi dei parametri, indicano il probabile centro.

Infine, se non si è a conoscenza del cerchio e non è possibile calcolare gli angoli rispetto alla circonferenza, si utilizza l'intersezione di coni.

