Introduzione
I segnali possono essere classificati in base alla dimensione del dominio:
- **Segnali 1D:** Variano rispetto a una sola dimensione (tipicamente il tempo). Esempi includono l'audio, la voce, la musica, segnali biomedici come l'ECG (elettrocardiogramma) e l'EEG (elettroencefalogramma), o i dati provenienti da sensori inerziali.
- **Segnali 2D:** Variano rispetto a due dimensioni spaziali. L'esempio principale è costituito dalle immagini digitali, che possono essere viste come matrici di numeri (pixel) dove la funzione $x = I(x, y)$ associa un'intensità o un colore a coordinate spaziali (righe e colonne).

Definizione operativa di Segnale
Un segnale descrive una grandezza fisica che varia rispetto a variabili indipendenti, quali possono essere il tempo o lo spazio. Lo strumento matematico utilizzato per modellare questi segnali sono le funzioni di una o più variabili, dove la variabile indipendente rappresenta le coordinate (es. tempo t) e la variabile dipendente rappresenta i valori della grandezza fisica $g = f(t)$.

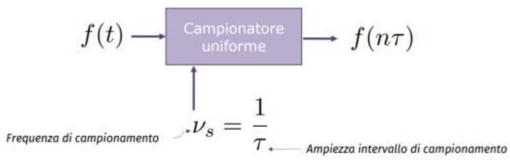
Continuo o Discreto
Una distinzione cruciale riguarda la natura delle variabili in gioco:
1. **Tempo (o Dominio):**
   - **Segnali a tempo continuo:** La variabile indipendente t, assume valori reali continui.
   - **Segnali a tempo discreto:** La variabile t, assume valori in un sottoinsieme discreto dei numeri reali, come risultato di un'operazione chiamata **campionamento**.
2. **Valori (o Ampiezza):**
   - **Segnali a valori continui:** La grandezza dipendente g, può assumere qualsiasi valore reale.
   - **Segnali a valori discreti:** La grandezza g, assume valori in un sottoinsieme discreto dei numeri reali, come risultato dell'operazione di **quantizzazione**.

Da queste combinazioni derivano le definizioni di **segnale analogico** (tempo continuo e valori continui) e **segnale digitale** (tempo discreto e valori discreti).

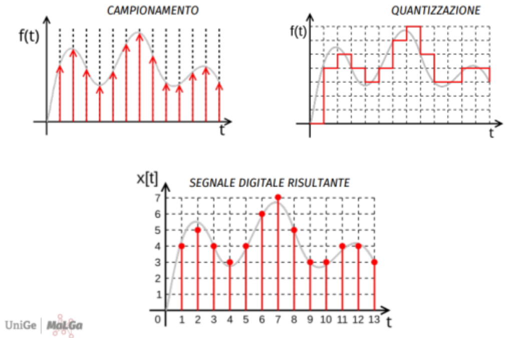
Segnale Continuo: un segnale continuo è un segnale in cui le variabili t (tempo) e g (grandezza) possono assumere qualsiasi valore reale. Il segnale è quindi definito in ogni istante del tempo e può essere rappresentato con una funzione: $g = f(t),\ t \in \mathbb{R}$. Un esempio è il segnale analogico.

Segnale Discreto: un segnale discreto è un segnale in cui la variabile indipendente t (tempo) assume valori in un sottoinsieme discreto dei numeri reali, come risultato di un'operazione chiamata campionamento; la grandezza dipendente g assume valori in un sottoinsieme discreto dei numeri reali come risultato dell'operazione di quantizzazione. Un esempio è il segnale digitale.

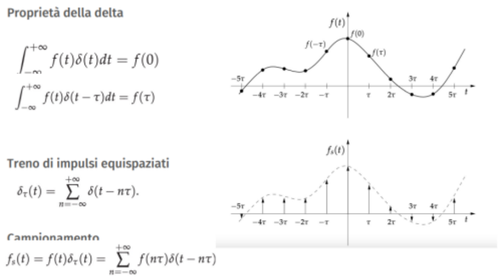
Il Processo di Digitalizzazione
Per trasformare un segnale analogico in digitale, si seguono tre fasi principali:
1. **Campionamento:** È il processo che discretizza il tempo. Un **campionatore uniforme** preleva valori dal segnale continuo a intervalli regolari τ . La frequenza con cui questo avviene è detta **frequenza di campionamento** $\upsilon_s = 1/\tau$, che trasforma il segnale continuo $f(t)$ in un segnale discreto $f(n\tau)$. Affinché il segnale sia fedele, l'intervallo di campionamento deve essere sufficientemente piccolo.



2. **Quantizzazione:** È il processo che discretizza i valori. Consiste nel prefissare un insieme finito di l valori numerici $\{x_1, \dots, x_l\}$ e associare a ogni numero x (valore reale del segnale campionato) il valore x_k che gli è più vicino.
3. **Codifica:** È il passo finale in cui i valori quantizzati vengono tradotti in parole binarie (bit). Ad esempio, con 8 bit si possono rappresentare 256 livelli di grigio in un'immagine.



Delta di Dirac (δ): è una distribuzione (strumento) matematica usata per rappresentare un impulso ideale concreto in un singolo istante del tempo. È fondamentale per descrivere matematicamente il campionamento, poiché permette di rappresentare un segnale continuo come una sequenza di impulsi discreti. Un impulso ideale che nel caso continuo vale infinito in 0 e 0 altrove, mentre nel caso discreto vale 1 in x=0 e 0 altrove. La delta è fondamentale per modellare matematicamente il campionamento attraverso un "treno di impulsi".

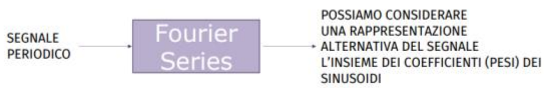


Rappresentazioni dei segnali nel dominio delle frequenze

Funzioni continue e periodiche possono essere rappresentate come somme pesate di seni e coseni. Questa rappresentazione è nota come Serie di Fourier, grazie alla quale possiamo ottenere una rappresentazione alternativa del segnale periodico e uno strumento utile per approssimarlo (con applicazione a compressione e riduzione del rumore).

Serie di Fourier

Una funzione continua e periodica può essere descritta attraverso una serie di sinusoidi (base di funzioni). Matematicamente, una funzione $f(t)$ di periodo T viene espressa sommando un termine costante a_0 a una serie infinita di termini coseno e seno, ciascuno moltiplicato per i propri coefficienti (pesi).



Consideriamo una funzione f(t) continua e periodica di periodo T

$$f(t) = a_0 + \sum_{k=1}^{\infty} \left(a_k \cos(\frac{2\pi}{T} tk) + b_k \sin(\frac{2\pi}{T} tk) \right)$$

Coefficienti

Applicando la formula di Eulero $e^{j\theta} = \cos\theta + j \sin\theta$ arriviamo a

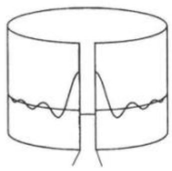
$$f(t) = \sum_{k=-\infty}^{\infty} c_k e^{j \frac{2\pi}{T} tk}$$

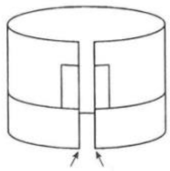
Coefficienti (complessi) di Fourier

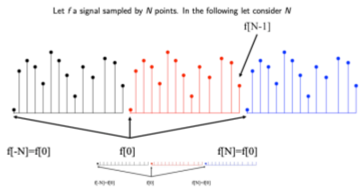
Trasformata di Fourier Discreta (DFT)

Quando si passa all'elaborazione digitale, si lavora con funzioni campionate. Consideriamo $f(t)$ campionata con N campioni equidistanti; chiamiamo $f[n]$ l'array che contiene i valori campionati. Possiamo considerare $f[n]$ una funzione discreta e finita con N campioni. Rendiamola periodica nell'intervallo $[0, T]$.

Cosa significa rendere periodica una funzione







La **DFT** trasforma i campioni originali $f[n]$ in coefficienti $F[k]$ nel dominio delle frequenze. Esiste anche l'operazione inversa, l'**Inversa della Trasformata di Fourier Discreta**, che permette di riottenere esattamente i valori originali del segnale partendo dai coefficienti. Un coefficiente particolare è $F[0]$, che rappresenta la componente continua: se la funzione f è reale, $F[0]$ è un numero reale e corrisponde al valore medio della funzione stessa.

$$f(t) = \sum_{k=-\infty}^{\infty} c_k e^{j \frac{2\pi}{T} tk}$$

Se la funzione è periodica e discreta:

$$f[n] = \sum_{k=0}^{N-1} F[k] e^{j \frac{2\pi}{N} kn}$$

La sommatoria è finita perche' nel caso di funzione discreta non mi occorrono infiniti sinusoidi per ricostruire tutti i dettagli

Data una funzione discreta e finita f[n] con N campioni, la sua DFT è

$$F[k] = \sum_{n=0}^{N-1} f[n] e^{-j \frac{2\pi}{N} kn}$$

coefficienti

Inversa della Trasformata di Fourier Discreta

$$f[n] = \sum_{k=0}^{N-1} F[k] e^{j \frac{2\pi}{N} kn}$$

Importante: restituisce esattamente i valori originali!

La componente continua F[0]

$$F[k] = \sum_{n=0}^{N-1} f[n] e^{-i \frac{2\pi}{N} nk}$$

K=0

$$F[0] = \sum_{n=0}^{N-1} f[n]$$

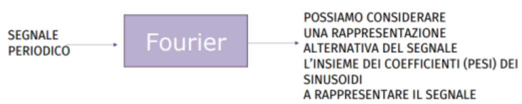
Se f è reale F[0] è reale e rappresenta il valore medio della funzione

Sebbene la DFT sia calcolabile direttamente (con complessità nell'ordine N^2), la FFT è un algoritmo molto più efficiente, utilizzato da tutte le moderne librerie di signal processing moderne.

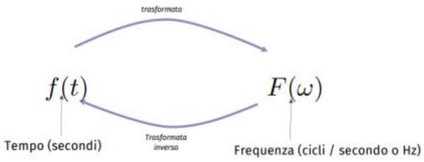
Quindi...

Cerchiamo rappresentazioni dei segnali che ci permettano di mettere in evidenza proprietà diverse dei segnali stessi. A questo scopo introduciamo la Trasformata di Fourier, grazie alla quale possiamo lavorare su due domini diversi ma collegati e passare da uno all'altro, tramite la trasformata e la sua inversa.

Una funzione continua e periodica può essere descritta attraverso una serie di sinusoidi (Serie di Fourier). Una funzione discreta e finita, dopo opportuna periodicizzazione, può essere rappresentata dalla Discrete Fourier Transform (DFT). Le funzioni non periodiche, sotto alcune ipotesi, possono essere rappresentate in modo analogo (**Trasformata di Fourier**).



I processi di rappresentazione nel dominio delle frequenze sono invertibili; questo ci permette di lavorare in un dominio alternativo (dominio delle frequenze o “dominio di Fourier”).



Consideriamo una funzione continua f(t)

In molti casi pratici f(t) è a valori reali, ma F è *solitamente* complessa

Trasformata di Fourier di f(t):

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

Inversa:

$$f(t) = \int_{-\infty}^{\infty} F(\omega)e^{j2\pi\omega t}d\omega$$

notate il cambio di variabile indipendente

Alcune proprietà importanti della FT (Fourier Transform)

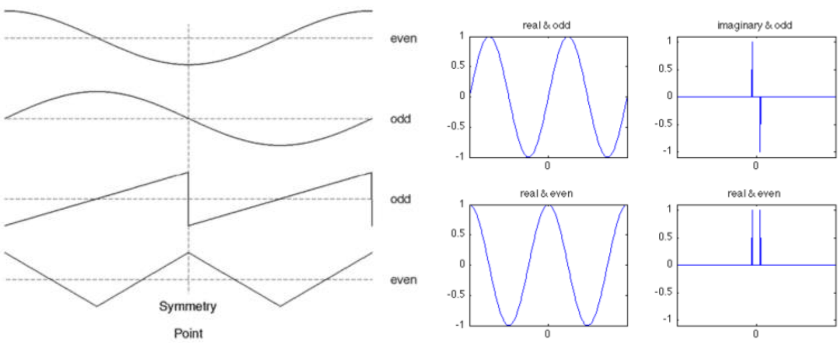
- **LINEARITÀ**: dati a e b numeri complessi, se $h(t)=a\,f(t)+b\,g(t)$, allora
$$H(\omega)=aF(\omega)+bG(\omega)$$
se $h[n]=af[n]+bg[n],\ 0\leq n\leq N-1$, allora $H[k]=aF[k]+bG[k],\ 0\leq k\leq N-1$
- **TRASLAZIONE** – time-shift: Per ogni numero reale t_0, se $h(t)=f(t-t_0)$ allora
$$H(\omega)=e^{-i2\pi\omega t_0}F(\omega)$$
Per ogni numero reale n_0, la DFT di $h(n)=(n-n_0)$ è
$$F_p[k]=F[k]e^{-i\frac{2\pi kp}{N}}$$
- **MODULAZIONE** – frequency-shift: Analogamente, se $h(t)=e^{-i2\pi\omega_0 t}f(t)$ allora
$$H(\omega)=F(\omega-\omega_0)$$
Analogamente, la DFT di $h[n]=f[n]e^{-i\frac{2\pi k_0 n}{N}}$ è
$$H[k]=F[k-k_0]$$

FT di segnali a valori reali

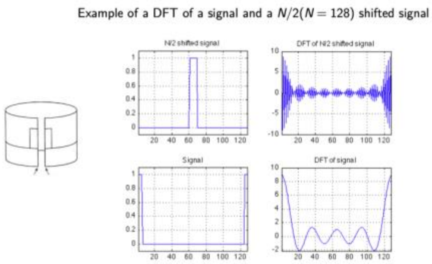
La FT di un segnale a valori reali ha una simmetria speciale:

- La parte reale è simmetrica pari
- La parte immaginaria è simmetrica dispari

La maggior parte dei segnali provenienti dal “mondo” sono reali.



DFT e shifting



Abbiamo visto che la DFT di un segnale traslato di una quantità p è

$$F_p[k] = F[k]e^{-j\frac{2\pi kp}{N}}$$

Nel caso particolare p=N/2 (N pari) abbiamo

$$F_{N/2}[k] = (-1)^k F[k]$$

Date due funzioni f(t) e g(t) o analoghe funzioni discrete, ci chiediamo quale sia $F\{f(t)g(t)\}$. Per poter rispondere ci serve un concetto nuovo: la Convoluzione.

Convoluzione

Considerando due funzioni continue f(t) e h(t), la convoluzione tra le due funzioni viene definita come:

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

Teorema di convoluzione

$$\mathcal{F}\{f(t)*h(t)\} = \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} f(\tau)h(t-\tau)d\tau\right]e^{-j2\pi\omega t}dt =$$
$$\int_{-\infty}^{\infty} f(\tau)\left[\int_{-\infty}^{\infty} h(t-\tau)e^{-j2\pi\omega t}dt\right]d\tau =$$
$$\int_{-\infty}^{\infty} f(\tau)H(\omega)e^{-j2\pi\omega\tau}d\tau =$$
$$H(\omega)\int_{-\infty}^{\infty} f(\tau)e^{-j2\pi\omega t}d\tau =$$
$$H(\omega)F(\omega)$$

Proprietà della FT

$$\mathcal{F}\{h(t-\tau)\} = H(\omega)e^{-j2\pi\omega t}$$

$g(t) = f(t) * h(t)$
$g(\omega) = F(\omega) \cdot G(\omega)$ IFT

$$f(t) * h(t) \iff H(\omega)F(\omega)$$
$$f(t)h(t) \iff H(\omega) * F(\omega)$$

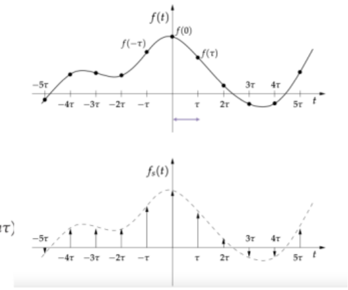
Campionamento

Treno di impulsi equispaziati

$$\delta_\tau(t) = \sum_{n=-\infty}^{+\infty} \delta(t - n\tau).$$

Campionamento

$$f_s(t) = f(t)\delta_\tau(t) = \sum_{n=-\infty}^{+\infty} f(n\tau)\delta(t - n\tau)$$



$$F_s(\omega) = F\big(f_s(t)\big) = \frac{1}{\tau}\sum_{n=-\infty}^{\infty} F\left(\omega - \frac{n}{\tau}\right)$$

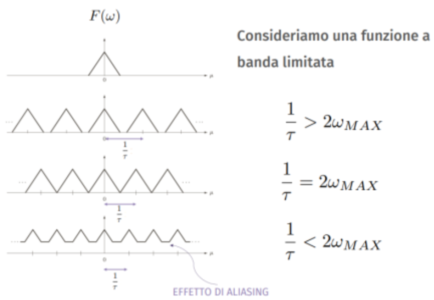
È una sequenza infinita e periodica di copie di $F(\omega)$ con ampiezza del periodo $\frac{1}{\tau}$. L'ampiezza del periodo nelle frequenze è duale all'ampiezza del periodo nel tempo.

Teorema del campionamento

Stabiliamo le condizioni sotto cui una funzione continua possa essere ricostruita a partire da un insieme di suoi campioni.

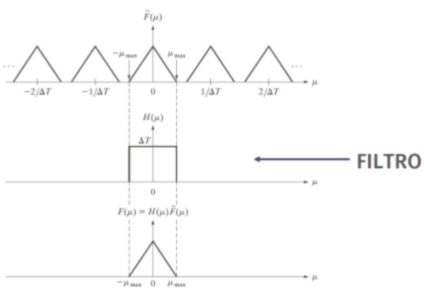
Partiamo da una <u>funzione continua $f(t)$ a banda limitata.</u> Una funzione $f(t)$ è a banda limitata se la sua trasformata di Fourier è nulla al di fuori dell'intervallo $[-\omega MAX,\ \omega MAX]$

$$F(\omega) = 0\ ,\ |\omega| > \ \omega MAX$$



Consideriamo una funzione continua $f(t)$ a banda limitata. Una ricostruzione perfetta del segnale è garantita se la frequenza di campionamento è almeno il doppio della frequenza massima del segnale. $\frac{1}{\tau} > 2\omega MAX$

Ricostruzione del segnale



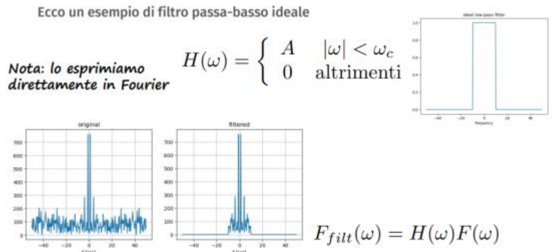
Filtraggio nelle Frequenze (segnali 1D)

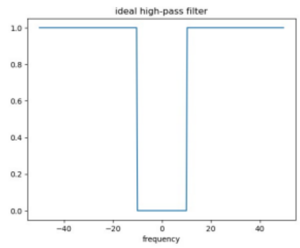
Un filtro è una funzione che lascia passare alcune componenti del segnale e ne elimina altre. Il concetto di filtro risulta molto chiaro nel dominio delle frequenze, e possiamo parlare di:

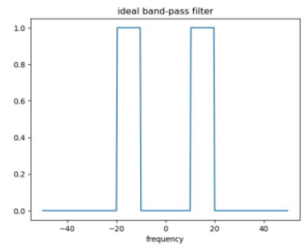
- Filtri passa basso: lasciano passare immodificate le basse frequenze, eliminano le alte frequenze
- Filtri passa alto: lasciano passare le alte frequenze
- Filtri passa banda: lasciano passare le frequenze comprese tra due valori

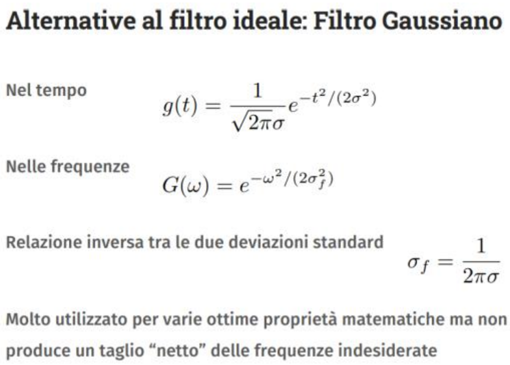
Filtrare un segnale corrisponde a moltiplicare un filtro H con la Trasformata di Fourier del segnale f : $F_{filt}(\omega) = H(\omega)F(\omega)$.

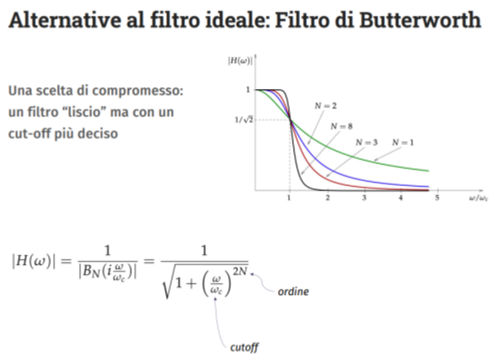
Un sistema che annulla "perfettamente" le armoniche in determinati intervalli di frequenza si chiama filtro ideale. Il modello del filtro ideale è molto adatto a descrivere l'azione di filtraggio. Nei casi reali un'implementazione del filtro ideale è impossibile.



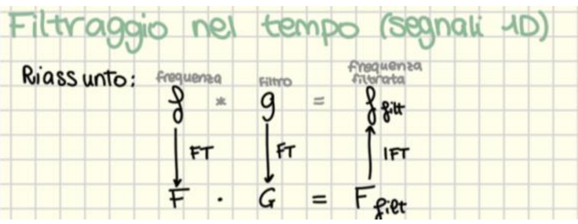








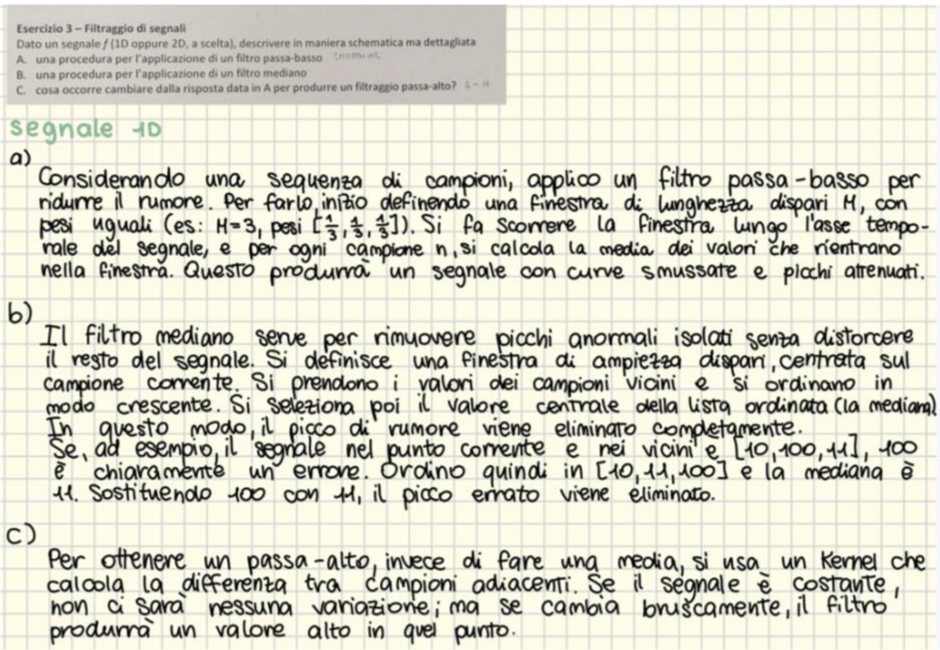
Il filtro ideale ha un taglio più netto ma è molto difficile da implementare senza effetti collaterali (disturbi, ecc). Si usa quindi il filtro gaussiano, che ha un taglio più tenue (smoothing), ma ha proprietà migliori (meno discrepanze). Si potrebbe usare anche il Butterworth, con un taglio più netto.



Rumore

Un termine incognito, imprevedibile e non desiderato che si aggiunge al segnale. Il rumore può essere "ambientale", come rumore di fondo nell'acquisizione di un suono o una luce al neon nello scattare una foto. Altrimenti può essere dovuto all'acquisizione, la digitalizzazione (campionamento e quantizzazione), la trasmissione, l'elaborazione (compresa la compressione o la conversione tra formati) del segnale stesso. La riduzione del rumore avviene tramite filtraggio del segnale.

Possiamo identificare molti modelli di rumore che possono essere adatti a descrivere diversi tipi di segnali rumorosi (rumori additivi, tra i quali il rumore bianco e il rumore gaussiano, modelli moltiplicativi, …).



Esercizio 3 – Filtraggio di segnali
Dato un segnale f (1D oppure 2D, a scelta), descrivere in maniera schematica ma dettagliata
A. una procedura per l'applicazione di un filtro passa-basso
B. una procedura per l'applicazione di un filtro mediano
C. cosa occorre cambiare dalla risposta data in A per produrre un filtraggio passa-alto?

segnale 1D

a)
Considerando una sequenza di campioni, applico un filtro passa-basso per ridurre il rumore. Per farlo, inizio definendo una finestra di lunghezza dispari M, con pesi uguali (es: $M=3$, pesi $[\frac{1}{3}, \frac{1}{3}, \frac{1}{3}]$). Si fa scorrere la finestra lungo l'asse temporale del segnale, e per ogni campione n, si calcola la media dei valori che rientrano nella finestra. Questo produrrà un segnale con curve smussate e picchi attenuati.

b)
Il filtro mediano serve per rimuovere picchi anormali isolati senza distorcere il resto del segnale. Si definisce una finestra di ampiezza dispari, centrata sul campione corrente. Si prendono i valori dei campioni vicini e si ordinano in modo crescente. Si seleziona poi il valore centrale della lista ordinata (la mediana). In questo modo, il picco di rumore viene eliminato completamente.
Se, ad esempio, il segnale nel punto corrente e nei vicini è [10, 100, 11], 100 è chiaramente un errore. Ordino quindi in [10, 11, 100] e la mediana è 11. Sostituendo 100 con 11, il picco errato viene eliminato.

c)
Per ottenere un passa-alto, invece di fare una media, si usa un kernel che calcola la differenza tra campioni adiacenti. Se il segnale è costante, non ci sarà nessuna variazione; ma se cambia bruscamente, il filtro produrrà un valore alto in quel punto.

Convoluzione

Consideriamo due funzioni continue $f(t)$ e $g(t)$. La convoluzione tra le due funzioni viene definita come:

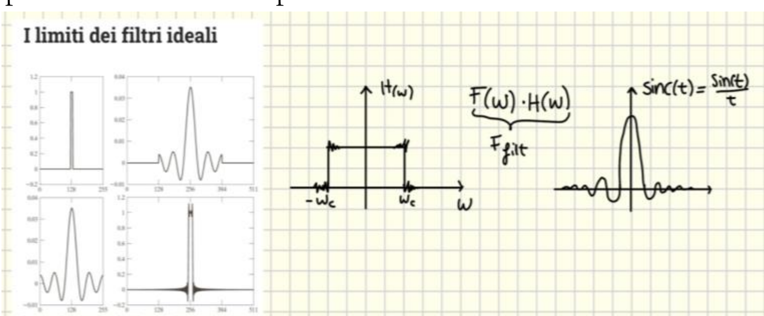
$$(f * g)(t) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} f(\tau) g(t - \tau)\, d\tau$$

To convolve a kernel with an input signal:
flip the signal, move to the desired time,
and accumulate every interaction with the kernel

$$f(t) * h(t) \iff H(\omega)F(\omega)$$
$$f(t)h(t) \iff H(\omega) * F(\omega)$$

La convoluzione è commutativa, quindi $f * g = g * f$

Lo stesso risultato può essere ottenuto filtrando nelle frequenze (moltiplicazione) oppure convolvendo un segnale con un opportuno filtro nel tempo.



Applicazioni tipiche:

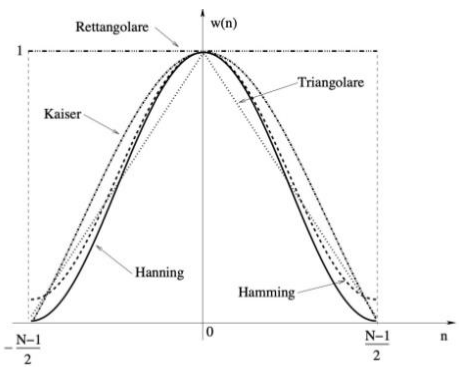
- Ridurre il rumore (filtri passa basso, nel tempo li chiamiamo filtri di smoothing)
- Mettere in evidenza punti di cambiamento "rapido" del segnale (filtri passa alto, nel tempo li chiamiamo filtri di enhancement)

Convoluzione Discreta:

Consideriamo $f(t)$ funzione di periodo T discreta (o discretizzata) con N punti nell'intervallo $[0,T] \rightarrow f[n]$. Consideriamo un filtro $h[n]$. La convoluzione discreta viene definita come:

$$g[n] = \sum_{k=0}^{N-1} h[k]f[n-k] = \sum_{k=0}^{N-1} h[n-k]f[k]$$

Filtri rappresentati mediante finestre:



Filtri di enhancement e differenze finite

In matematica discreta le differenze finite sono definite come $f'(x) = \frac{f(x+h)-f(x)}{h}$.

Nell'analisi dei segnali discreti diventano uno strumento indispensabile per rilevare punti di cambiamento del segnale.

- La derivata prima:
  - 0 in zone costanti del segnale
  - Diversa da zero in corrispondenza di variazioni
- La derivata seconda:
  - 0 in zone costanti del segnale
  - 0 in zone a crescita o decrescita costante
  - Diverse da 0 in punti di variazione della pendenza

Differenze finite e rumore

Schema generale:

- Filtrare il segnale con un filtro passa basso per attenuare il rumore
- Successivamente filtrare il segnale con un filtro passa alto per mettere in evidenza i punti di cambiamento

Esistono filtri che sono in grado di svolgere entrambi i compiti, per esempio la derivata della Gaussiana (che è una convoluzione tra Gaussiana e filtro passa alto).

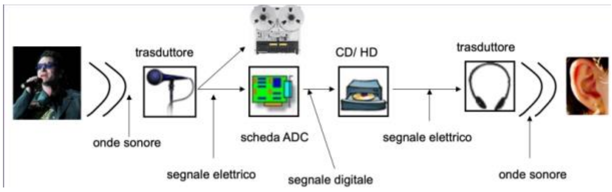
Operatori di rango (non lineari)

Altri tipi di filtro utili che non vengono realizzati tramite convoluzione ma sono basati su un passo intermedio di ordinamento del segnale o di una porzione di esso. Esempi di filtri di questo tipo sono il filtro minimo, il filtro massimo o il filtro mediano. Il filtro mediano, in particolare, è utile a curare rumore impulsivo che si riscontra quando il segnale presenta valori errati e scorrelati dagli elementi vicini.

Sound processing

Il suono che percepiamo corrisponde ad un complesso fenomeno fisico di variazione della pressione dell'aria vicino all'apparato uditivo. Se le variazioni sono ampie, sentiamo suoni forti; se le variazioni sono rapide sentiremo suoni con toni alti. La sorgente emette un suono, che si propaga nel mezzo di trasmissione (di solito, l'aria).

Come nasce un segnale audio



Il trasduttore elettroacustico (es. microfono) è in grado di tradurre le vibrazioni del suono in un segnale elettrico. Il convertitore Analogico-Digitale (ADC) lo trasforma in un segnale digitale applicando campionamento e quantizzazione. Questo segnale digitale viene archiviato fisicamente (es. CD) o digitalmente (es. Cloud).

Per l'ascolto è necessario seguire il procedimento opposto: il segnale digitale diventa elettrico e il trasduttore (es. Cuffie) emette onde sonore.

Alcuni segnali nascono direttamente da un dispositivo elettronico (es. Sintetizzatore) e quindi non necessitano il trasduttore iniziale.

Rappresentazione del suono

Un primo modo per rappresentare il suono è la forma d'onda, che rappresenta come la pressione sonora (o la corrispondente tensione elettrica) varia nel tempo in un punto fissato dello spazio. È il modo più diretto per descrivere un segnale sonoro nel dominio del tempo.

- Frequenza: numero di vibrazioni al secondo (raddoppiando il numero di vibrazioni salgo di un'ottava)
- Unità di misura: Hertz, numero di vibrazioni al secondo
- Determina l'altezza (pitch) del suono, che è una caratteristica percettiva del suono:
    - frequenze alte, molte oscillazioni al secondo: suono acuto
    - frequenze basse, poche oscillazioni al secondo: suono grave

Per essere percepite come suono, le vibrazioni devono cadere all'interno di un intervallo predefinito:

- per un essere umano con un buon udito, il numero di variazioni al secondo deve essere nel range [20 Hz – 20 KHz]
- Sopra ai 20 KHz sono considerati ultrasuoni (udibili da delfini e pipistrelli)
- Sotto i 20 Hz sono considerati infrasuoni (udibili da elefanti e pesci)

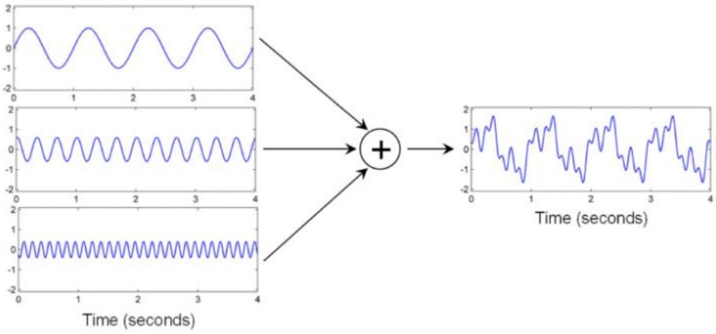
La gamma di suoni coperta dalla musica è molto più limitata (do grave 65 Hz, do acuto ~8 KHz). L'orecchio umano è particolarmente sensibile nell'intervallo [2 KHz – 5 KHz].

Intensità del suono

L'intensità del suono è legata all'ampiezza della vibrazione, Più è ampia la vibrazione più il suono è forte. L'intensità si misura in decibel e ha un range di "accettabilità" compreso tra Soglia di udibilità (0 dB) e Soglia del dolore (130 dB).

Source	Intensity	Intensity level	× TOH
Threshold of hearing (TOH)	10^{-12}	0 dB	1
Whisper	10^{-10}	20 dB	10^{2}
Pianissimo	10^{-8}	40 dB	10^{4}
Normal conversation	10^{-6}	60 dB	10^{6}
Fortissimo	10^{-2}	100 dB	10^{10}
Threshold of pain	10	130 dB	10^{13}
Jet take-off	10^{2}	140 dB	10^{14}
Instant perforation of eardrum	10^{4}	160 dB	10^{16}

Suoni semplici e suoni complessi



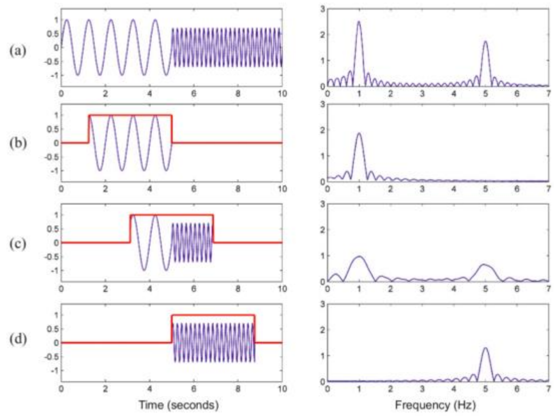
Un suono complesso (uno strumento musicale o la voce) non è una singola sinusoide ma è formato dalla somma di più componenti sinusoidali di frequenze multiple della frequenza fondamentale (armoniche). La frequenza fondamentale (che determina l'altezza/pitch) è la frequenza più bassa f_0 presente nel segnale. La nota è l'effetto percepito della frequenza fondamentale. Il segnale può essere scomposto nelle sue componenti in frequenza tramite la trasformata di Fourier. Le armoniche sono multipli interi della fondamentale e caratterizzano il suono di strumenti armonici.

Timbro

È una caratteristica che consente di distinguere due suoni della stessa frequenza fondamentale e ampiezza, ma generati da sorgenti diverse, per esempio strumenti diversi. È legato alla presenza di diverse armoniche con ampiezze diverse.

Short Time Fourier Transform (STFT)

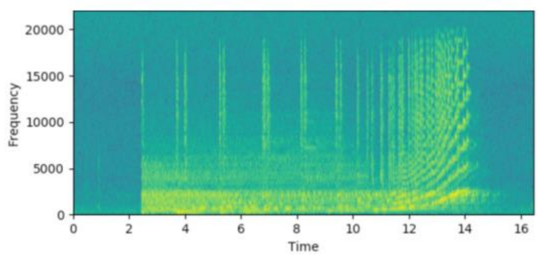
La FT descrive il contenuto "globale" del segnale in termini di frequenza, ma non ci permette di localizzare un fenomeno all'interno del segnale. Si usa quindi la STFT, uno strumento matematico che analizza i segnali non stazionari, ossia il cui spettro varia nel tempo. La STFT usa una porzione del segnale, detta finestra, di durata finita, che funziona come una successione di FT calcolate su segmenti del segnale. I risultati ottenuti dipendono dalla dimensione della finestra e dalla forma (tagli bruschi introducono artefatti).



Lo Spettrogramma

È una rappresentazione bidimensionale del modulo della STFT, visualizzato come un'immagine:

- L'asse orizzontale è il tempo
- L'asse verticale le frequenze
- Il colore l'intensità / ampiezza (modulo)



Immagini digitali

Un'immagine Digitale è definita come una matrice di numeri. La dimensione dell'immagine corrisponde al numero di pixel, seguendo la convenzione righe per colonne. I valori nella matrice rappresentano i livelli di intensità luminosa in uno spazio (x, y).

Tipologie di Immagini:

- **Pittoriche (foto):** Rappresentano l'intensità (bianco e nero) o il colore.
- **Range images:** Contengono informazioni sulla profondità.
- **Immagini mediche:** Rappresentano il livello di assorbenza delle radiazioni.
- **Immagini termiche:** Rappresentano il calore.

Trasformata di Fourier in 2D

È una somma di proiezioni di onde piane a diverse angolazioni e frequenze. Matematicamente è espressa come:

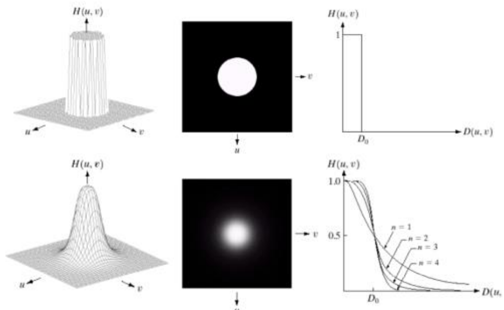
$$F(u, v) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} f[m, n] e^{-j2\pi(um+vn)}.$$

Lo spettro di un oggetto traslato è identico allo spettro dell'immagine originale (cambia solo la fase, non il modulo). Se un oggetto nell'immagine viene ruotato, anche il suo spettro ruota corrispondentemente.

Filtraggio nel Dominio delle Frequenze (Passa Basso)

I Filtri Passa Basso (Low Pass) sono filtri che attenuano le alte frequenze mantenendo quelle basse, usati per sfocare o ridurre il rumore. Il Filtro Ideale taglia nettamente le frequenze oltre una certa soglia D_0.

- **Effetto "Ringing":** È un artefatto visivo (oscillazioni o cerchi concentrici attorno ai bordi) causato dall'uso di filtri ideali nel dominio di Fourier.
- **Filtro Gaussiano:** È un filtro "smooth" (morbido) che evita l'effetto ringing, producendo una sfocatura più naturale.



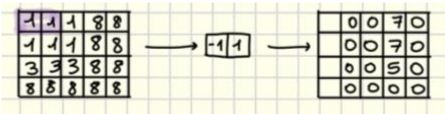
Filtri passa alto e derivate

I Filtri Passa Alto vengono utilizzati per stimare le differenze finite, analizzando le variazioni locali dell'immagine (analogia con la derivata).

In analogia a quanto visto nel caso 1D possiamo stimare le differenze finite tramite convoluzione con un filtro S.

S

-1	1

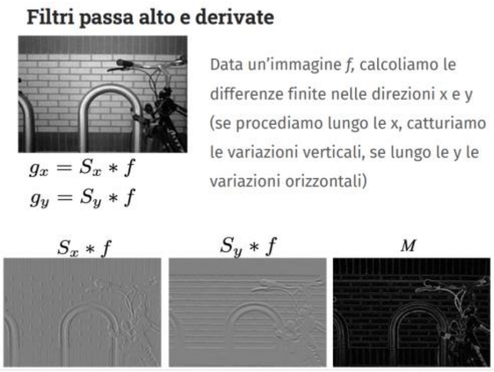


Poiché l'immagine è 2D, l'analisi viene applicata lungo le due direzioni principali:

- **Direzione x ($S_x = S$):** Cattura le variazioni verticali.
- **Direzione y ($S_y = S^T$):** Cattura le variazioni orizzontali.

Il Gradiente dell'Immagine

Il gradiente $\nabla f = \left(g_x, g_y\right)$ è un campo vettoriale che descrive la variazione nelle direzioni x e y; le sue componenti sono chiamate derivate parziali. Ogni elemento è un vettore bidimensionale che possiamo usare per calcolare il modulo $M = \sqrt{g_x^2 + g_y^2}$. Il modulo descrive l'intensità delle variazioni del segnale senza considerare il segno (direzione).



Filtro di Sobel

È un filtro specifico utilizzato per approssimare la stima del gradiente. Oltre a calcolare le differenze finite, applica un piccolo livello di smoothing per ridurre il rumore durante il calcolo dei contorni.

S_y

-1	-2	-1
0	0	0
1	2	1

$$g_x = S_x * f$$
$$g_y = S_y * f$$
$$S_y = S_x^\top$$

Edge (Contorni)

Sono definiti come cambiamenti significativi e locali all'interno di un'immagine. Si verificano tipicamente ai bordi di diverse regioni, ma non corrispondono necessariamente ai contorni degli oggetti fisici reali (possono essere causati da ombre o texture). Gli edge sono punti dell'immagine dove si è verificata una alta variazione dei valori di intensità luminosa: possono essere rilevati usando filtri passa alto e, in modo semplice, applicando una sogliatura (thresholding) sul modulo del gradiente.

Colore e immagini

I colori non sono proprietà intrinseche dei corpi ma sensazioni attivate nel sistema nervoso dell'osservatore L'esperienza del colore è causata dal fatto che il sistema visivo umano risponde in modo diverso ad una varietà di lunghezze d'onda. Inoltre, la percezione del colore da parte dell'uomo è una funzione complessa del contesto: illuminazione, proprietà dell'oggetto osservato, memoria, identità dell'oggetto, emozione.

I nostri occhi percepiscono una parte molto limitata delle radiazioni elettromagnetiche: quelle con lunghezze d'onda (lambda) compresa tra i 400 nanometri (ultravioletti) e i 700 nm (infrarossi).

Il colore percepito a livello della retina è condizionato da due fattori fondamentali: la superficie su cui la luce è riflessa, il tipo di luce e il colore della luce.

Il colore delle superfici colorate è il risultato di una grande varietà di meccanismi. Questi dipendono dal materiale che costituisce la superficie e come interagisce con la luce che lo colpisce, quindi quanto assorbe le lunghezze d'onda, rifrazione, diffrazione...

La luce può venire da sorgenti naturali, come il sole o il cielo nuvoloso, oppure sorgenti artificiali: luci incandescenti, luci fluorescenti, luci ai vapori di mercurio, luci alogene, ...

Creazione del colore

La misura del colore è oggetto di studio della colorimetria. Ci sono due metodi usati per formare il colore:

- **Sintesi additiva:** si parte dall'assenza di luce (nero) e si aggiungono colori. Un nuovo colore viene ottenuto attraverso la miscelazione (somma) di emissioni di luce relative ad altri colori
- **Sintesi sottrattiva**: si parte dal bianco e si "eliminano" dei colori per ottenere altri colori. Si fa passare luce bianca attraverso dei filtri che lasciano passare solo delle radiazioni di una determinata lunghezza d'onda (cioè un dato colore):
  - La vernice blu riflette quasi tutta la luce a lunghezze d'onda corte e assorbe la luce a lunghezze lunghe
  - La vernice gialla riflette quasi tutta la luce a lunghezze lunghe e assorbe quella a lunghezze corte.
  - Entrambe riflettono le lunghezze d'onda medie, che appaiono circa verdi

Gli spazi di colore

Gli spazi di colore sono normalmente inventati per motivi pratici, quindi ne esistono molti. Possono essere derivati:

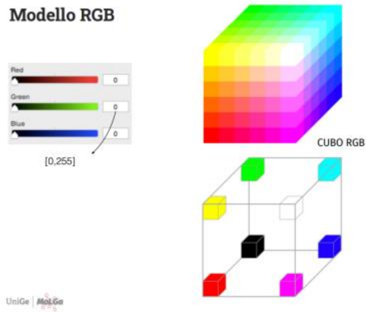
- da dispositivi: RGB, CMYK
- dallo studio della percezione: HSV

Modello RGB

È basato su 3 colori primari monocromatici:

- R red (lunghezza d'onda l=700nm)
- G green (l=546.1nm)
- B blue (l=435.8nm)

Tipico dei dispositivi di acquisizione e visualizzazione. Viene solitamente usato per lo storage ed è quindi il modello di riferimento delle immagini digitali. Una vasta percentuale dello spettro visibile può essere rappresentata miscelando luce rossa, verde e blu in diverse proporzioni e intensità. Se i colori primari sono sommati producono il bianco: per questo motivo il modello RGB è di tipo additivo. Geometricamente, è rappresentato da un cubo. Per convertire un'immagine a colori in scala di grigi, dovremmo fare $\omega_R R + \omega_G G + \omega_B B$, dove i coefficienti rappresentano la tricromatica umana.



Modello CMY

A partire dal cubo RGB ritroviamo immediatamente i colori ciano (Cyan), Magenta, giallo (Yellow) (posizionati sui 3 vertici rimanenti). Il modello CMY rappresenta lo stesso spazio di colore del RGB ma utilizza una base alternativa. Il modello CMY si basa sul modello sottrattivo. È tipico dei dispositivi di stampa ed è basato sulla capacità propria dell'inchiostro su carta di assorbire luce. Quando la luce bianca colpisce gli inchiostri, alcune lunghezze d'onda visibili vengono assorbite, mentre altre vengono riflesse e quindi viste.

Modello CMYK

K rappresenta il nero (blacK): il nero può essere derivato direttamente dalla combinazione di C M e Y (ossia assorbendo tutti e tre i colori base). Purtroppo, però, gli inchiostri di stampa contengono molte impurità; quindi, questo modello di combinazione del colore invece di produrre il nero produce un marrone scuro. CMYK è lo standard delle stampanti.

Trasformazione da RGB a CMY

Assumendo che R, G, B appartengano all'intervallo [0,1]

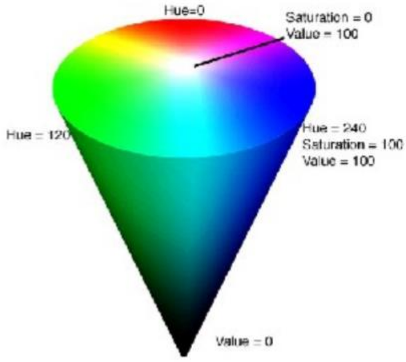
$$\begin{pmatrix} C \\ M \\ Y \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

Gli spazi RGB e CMYK sono stati concepiti in funzione dei dispositivi di visualizzazione (monitor e stampanti), ma non rispecchiano affatto la nostra percezione. Un modello alternativo molto più vicino al nostro modo di "vedere" i colori e valutare la similarità tra colori è il modello HSV. Il modello HSV viene usato molto spesso nell'ambito dell'analisi di immagini.

Modello HSV

Lo spazio di colore HSV descrive i colori in modo più vicino alla percezione umana. È rappresentato geometricamente da un cono. Le sue 3 componenti sono Hue, Saturation e Value.

- Hue (tinta) rappresenta il tipo di colore puro; è misurato come un angolo tra 0 e 360 gradi lungo la circonferenza del cono.
- Saturation rappresenta la 'purezza' del colore, ovvero quanto è diluito con il bianco. Un valore basso indica un colore sbiadito, un valore alto un colore vivido.
- Value rappresenta la luminosità, brillantezza o intensità del colore. È l'altezza del cono, con 0 nero assoluto e 1 massima luminosità del colore.



Trasformazione da RGB a HSV

V = (R+G+B)/3

S = 1-min(R,G,B)

H=... $H = \begin{cases} \theta & \text{se } g \geq b \\ 2\pi - \theta & \text{se } g < b \end{cases} \quad \text{con } \theta = \cos^{-1}\left(\frac{\frac{1}{2}(r-g)+(r-b)}{\sqrt{(r-g)^2+(r-b)(g-b)}}\right)$

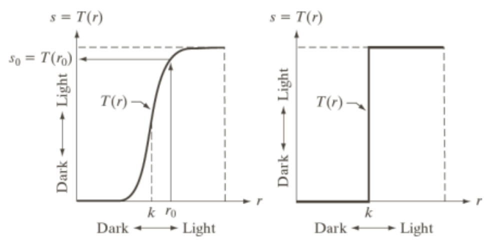
Elaborazione di immagini a colori

Proprio perché è il risultato di un fenomeno di percezione, il colore è spesso ignorato nelle applicazioni di elaborazione di immagini. La maggior parte degli algoritmi utilizzano solo informazione di brightness, codificabile attraverso livelli di grigio. Il colore diventa più importante in applicazioni di medio/alto livello (identificazione/classificazione di oggetti). Una tipica operazione di elaborazione di immagini a colori è la segmentazione, ossia l'identificazione di aree uniformi all'interno di un'immagine.

Istogrammi e Operazioni sui pixel

Trasformazioni di intensità luminosa

Data un'immagine I, consideriamo un operatore T che produce un'immagine di output J dove gli elementi hanno subito variazioni nell'intensità luminosa $J(p) = T[I(p)]$. Un esempio può essere il negativo di un'immagine. È una semplice trasformazione lineare. Se i livelli di grigio assumono valori nel range [0, L-1] (con L=256) allora $s = L - 1 - r$.



Operazioni tra immagini: esempi di applicazione

Date due immagini I1 e I2 (della stessa dimensione), ha senso applicare ad esse operatori aritmetici (pixel-wise).

- Addizione: può essere utile per fondere immagini diverse
- Sottrazione: per mettere in evidenza differenze/cambiamenti
- Moltiplicazione: per pesare in modo diverso elementi di un'immagine o per applicare maschere

a b c

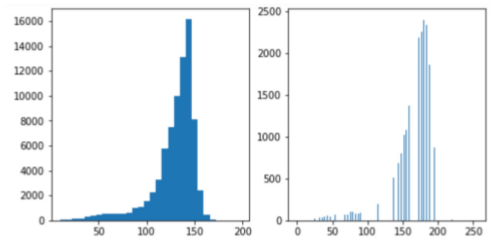
FIGURE 2.30 (a) Digital dental X-ray image. (b) ROI mask for isolating teeth with fillings (white corresponds to 1 and black corresponds to 0). (c) Product of (a) and (b).

Istogrammi di immagini di intensità

L'istogramma di un'immagine I con valori di intensità nell'intervallo [0, L-1] è una funzione discreta $h(r_k) = n_k$, dove:

- r_k è il k-esimo valore di intensità del range
- n_kè il numero di pixel dell'immagine I con intensità r_k

Può essere molto utile raggruppare elementi con valori simile. Questo corrisponde ad un'operazione di quantizzazione, dove i valori [0, L-1] vengono raggruppati in bin.



Calcolo dell'istogramma

Data l'immagine I per calcolare l'istogramma H con M bin, H={H1, ..., HM}:

```
for each p in I
   g = I(p);
   bin_g = g/bin_size;
   H(bin_g)= H(bin_g)+1;
end
```

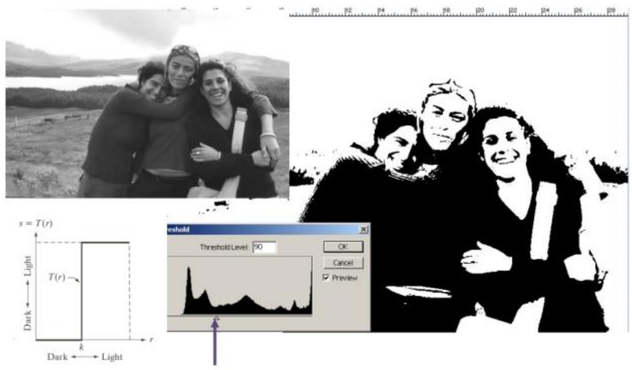
Normalizzazione dell'istogramma

```
for each b in H
   H(b)= H(b)/N;
end
```

N: numero dei pixel dell'immagine I. L'istogramma normalizzato può essere pensato come una stima della probabilità dei valori di intensità in un'immagine $\sum_r H(r) = 1$

Istogrammi e sogliatura

Un istogramma è una rappresentazione alternativa dell'immagine che mette in evidenza la distribuzione delle intensità perdendo informazione spaziale. A partire dalla stessa immagine, si possono costruire istogrammi diversi, cambiando il numero di bin (elementi dell'istogramma) o scegliendo di normalizzare l'istogramma (in questo modo lo possiamo trattare come una stima di probabilità). Possiamo analizzare l'istogramma ottenuto per scegliere un'opportuna soglia da applicare all'immagine ad intensità luminosa per binarizzarla.



operazioni H che modificano il contenuto del pixel $H(I(\vec{p})) = I_{out}(\vec{p})$

→ modifico l'istogramma

operazioni geometriche che "spostano" i pixel $T(\vec{p}) = \vec{q}$ $I(\vec{p}) \rightarrow I_{out}(\vec{q})$

→ non modifico l'istogramma

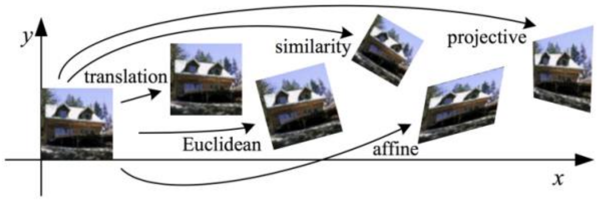
Contrast stretch

È un'operazione che ci permette di espandere il range di valori di intensità luminosa utilizzati in un'immagine.

```
def my_contrast_stretch(img):
    M=np.max(img)
    m=np.min(img)
    return np.multiply(np.divide((img-m),(M-m)),255).astype(np.uint8)
```

Trasformazioni geometriche

Le trasformazioni geometriche sono operazioni che modificano le coordinate spaziali di un'immagine. A differenza di altre operazioni di elaborazione, esse agiscono sulla posizione dei pixel (x, y) e non direttamente sul loro valore di intensità (colore).



$$I_{in}(x,y) \rightarrow I_{out}(u,v) \quad \text{con} \quad (u,v) = \mathcal{T}(x,y)$$

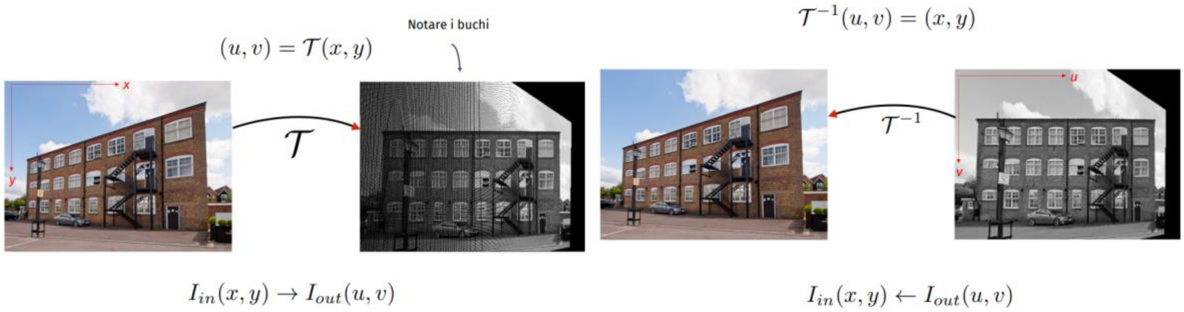
Formalmente, una trasformazione mappa un'immagine di input I_{in} in un'immagine di output I_{out} secondo la relazione: $I_{in}(x, y) \rightarrow I_{out}(u, v)$, dove le nuove coordinate sono ottenute tramite una funzione di trasformazione T : $(u, v) = \mathrm{T}(x, y)$.

Definizioni matematiche per tre trasformazioni fondamentali:

- Traslazione: Sposta ogni punto di un vettore costante $t = (t_1, t_2)$.
$$q_1 = p_1 + t_1$$
$$q_2 = p_2 + t_2$$
- Rotazione: Ruota i punti di un angolo θ attorno all'origine degli assi.
$$q_1 = p_1 \cos\theta + p_2 \sin\theta$$
$$q_2 = -p_1 \sin\theta + p_2 \cos\theta$$
- Scalatura: Ridimensiona l'immagine moltiplicando le coordinate per fattori di scala c e d.
$$q_1 = c \cdot p_1$$
$$q_2 = d \cdot p_2$$

Trasformazioni geometriche su immagini discrete

- La trasformazione descrive il punto d'arrivo di un pixel (non necessariamente sulla griglia discreta)
- Occorre ricampionare gli elementi sulla griglia discreta (mapping diretto o inverso)



Applicare trasformazioni geometriche a immagini digitali (discrete) presenta sfide specifiche:

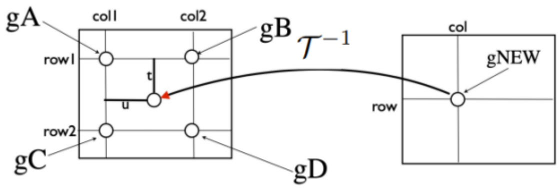
1. La trasformazione calcolata può restituire coordinate che non cadono esattamente sulla griglia intera dei pixel.
2. È necessario **ricampionare** gli elementi per adattarli alla griglia discreta.

Per gestire questo problema esistono due approcci di mapping:

- **Mapping Diretto:** Si prendono i pixel dell'immagine di input (x,y) e si calcola la loro nuova posizione (u,v) tramite T .
  $$(u, v) = \mathrm{T}(x, y)$$
  Problema: Questo metodo può creare "buchi" (pixel vuoti) nell'immagine di destinazione se la trasformazione espande l'immagine o se le coordinate mappate non sono contigue.
- **Mapping inverso:** Si parte dai pixel della griglia dell'immagine di output (destinazione) e si calcola da quale posizione dell'input provengono, usando la trasformazione inversa T^{-1}.
  $$\mathrm{T}^{-1}(u, v) = (x, y)$$
  $$I_{in}(x, y) \leftarrow I_{out}(u, v)$$
  Questo metodo garantisce che ogni pixel dell'immagine finale abbia un valore, evitando i buchi.

Interpolazione bilineare su mapping inverso

Per ogni pixel p dell'immagine in costruzione, calcoliamo il livello di grigio come combinazione bilineare degli elementi dell'immagine di input più vicini al mapping inverso di p.



gNEW = (1-t)(1-u)gA + u(1-t)gB + t(1-u)gC + utgD

Quando si utilizza il mapping inverso, la coordinata calcolata (x,y) nell'immagine di input cade spesso tra quattro pixel esistenti (interi). Per assegnare un valore a quel punto, si usa l'interpolazione bilineare. Il livello di grigio del nuovo pixel viene calcolato come una combinazione lineare pesata dei 4 pixel vicini nell'immagine originale.

Dati i quattro vicini $gA,\ gB,\ gC,\ gD$ e le distanze frazionarie t e u :

$$gNEW = (1 - t)(1 - u)gA + u(1 - t)gB + t(1 - u)gC + utgD$$

Esempi di risposte a domande d'esame

d) nel passaggio tra il B e C abbiamo eseguito il campionamento del segnale continuo, prelevando il suo valore a intervalli di periodo $V_S = 1/T$, ottenendo così una sequenza discreta di campioni.

Esercizio 2 – Convoluzione

A. Definisci l'operazione di convoluzione tra due segnali f(t) e g(t)

Date le due funzioni f e g

Consideriamo h(t)=f(t)*g(t).

B. a quale valore di t corrisponde il valore massimo di h?

C. qual è l'intervallo di t dove h ha valore diverso da 0?

$t_{max} = t_{centro-f} + t_{centro-g}$

$f(t) = [A, B] \quad g(t) = [C, D]$

$h(t) = [A+C; B+D]$

a) La convoluzione tra le due funzioni viene definita come:

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

b) Il valore massimo di h(t) si ha quando i due segnali sono perfettamente sovrapposti, cioè quando l'area di intersezione tra f e g è massima. Osservando i grafici questo avviene per $t = 0$.

c) h(t) ha valore diverso da 0 nell'intervallo tra [-1.5; 1.5], perchè f = [-1; 1] e g = [-0.5; 0.5].

b) Guardando l'istogramma, i valori di intensità sono concentrati tra circa 0 e 200, mentre l'intervallo possibile va da 0 a 255, non sfruttando tutto il range dinamico. Con una procedura di context stretch possiamo espandere il range di intensità da [0,200] a [0,255], aumentando il contrasto e rendendo più visibili le differenza di luminosità.

Contrast stretch: operazione che ci permette di espandere il range di valori di intensità luminosa utilizzati in un'immagine

c)

```
def context_stretch(image):
    M = max (image)        # esempio: 200
    m= min (image)         # esempio: 0
    val = ((image-M)/(M-m))*255
    stretch = val.astype(np.uint8) #conversione valori in interi a 8 bit
    return stretch
```

Esercizio 4 – Ridurre il rumore da un'immagine digitale

A. Descrivere i passi principali di un algoritmo per ridurre il rumore di un'immagine nel dominio delle frequenze

B. Fornire l'esempio di un filtro (con tutti i dettagli necessari) che possa operare in questo senso

C. Qual è un effetto indesiderato del filtraggio di riduzione del rumore sulle immagini?

a)

1) conversione nel dominio delle frequenze: applico la trasformata di Fourier 2D (FFT) all'immagine per ottenere lo spettro delle frequenze.

2) analisi dello spettro: individuo le componenti ad alta frequenza, che tipicamente rappresentano il rumore.

3) applicazione del filtro: modifico lo spettro moltiplicandolo per una maschera filtrante (ad esempio passa-basso) che attenua o elimina le alte frequenze indesiderate.

4) ricostruzione dell'immagine: applico la trasformata inversa (IFFT) per tornare nel dominio spaziale e ottenere l'immagine filtrata.

b)

Un esempio di filtro per ridurre il rumore di un'immagine è il filtro passa-basso, che attenua le alte frequenze dove risiede il rumore. Si distingue tra:

- Filtro ideale: pone a 0 i valori di tutte le frequenze oltre una certa distanza D_0 dal centro. Genera un effetto indesiderato detto 'ringing'.
- Filtro Gaussiano (smooth): A differenza di quello ideale, attenua le frequenze in modo graduale. Riduce il rumore sfocando l'immagine.

c)

Gli effetti indesiderati sono:

- l'effetto 'ringing' nel filtro ideale, delle oscillazioni, dei bordi 'fantasma', come dei bordi doppi attorno ai contorni degli oggetti
- perdita di dettagli con il filtro Gaussiano, che riduce il rumore attenuando le alte frequenze, rendendo i bordi e i dettagli dell'immagine meno nitidi.

segnale 2D

A. Procedura per filtro passa-basso (Smoothing)

Il filtro passa-basso serve a ridurre il rumore e i dettagli fini, producendo un effetto di sfocatura (smoothing). È un'operazione **lineare**.

Procedura:

1. **Definizione del Kernel:** Si definisce una maschera (o kernel) di dimensione dispari (es. 3×3 o 5×5). I coefficienti della maschera devono essere tutti positivi e la loro **somma deve essere uguale a 1** (per mantenere la luminosità media dell'immagine).
   - Esempio (filtro media): Una matrice 3×3 dove ogni elemento vale $1/9$.
2. **Scansione:** Si fa scorrere la finestra del kernel su ogni pixel dell'immagine originale $f(x, y)$.
3. **Convoluzione:** Per ogni posizione, si calcola la somma pesata dei pixel vicini coperti dalla maschera. Matematicamente, si moltiplica il valore di ogni pixel sottostante per il corrispondente coefficiente del kernel e si sommano i risultati.
4. **Assegnazione:** Il valore risultante sostituisce il valore del pixel centrale nella nuova immagine filtrata $g(x, y)$.

Nota: Questo processo attenua le variazioni brusche di intensità tra pixel adiacenti.

B. Procedura per filtro mediano

Il filtro mediano è un filtro **non lineare**, molto efficace per rimuovere il rumore "sale e pepe" (valori estremi isolati) preservando però i bordi (edges) dell'immagine.

Procedura:

1. **Definizione della Finestra:** Si definisce una finestra di vicinato (es. 3×3) centrata sul pixel da elaborare.
2. **Raccolta dei valori:** Si sovrappone la finestra al pixel corrente $f(x, y)$ e si raccolgono in una lista i valori di intensità di tutti i pixel che cadono all'interno della finestra (per un 3×3, avremo 9 valori).
3. **Ordinamento (Sorting):** Si ordinano i valori raccolti in ordine crescente (o decrescente).
4. **Selezione della Mediana:** Si estrae il valore che si trova esattamente al centro della lista ordinata (il 5° elemento su 9).
5. **Assegnazione:** Si sostituisce il valore del pixel originale con il valore mediano appena trovato.

C. Modifiche per il filtro passa-alto

Il filtro passa-alto serve a evidenziare i bordi e i dettagli rapidi, attenuando le zone uniformi. Rispetto alla procedura descritta in **A** (filtro lineare), occorre cambiare la natura del **Kernel**:

1. **Somma dei coefficienti:** A differenza del passa-basso (dove la somma è 1), in un filtro passa-alto la **somma dei coefficienti del kernel deve essere 0**. Questo fa sì che nelle aree di colore uniforme (frequenza zero o DC component) l'uscita sia nera (zero).
2. **Distribuzione dei valori:** Il kernel ha tipicamente un valore positivo elevato al centro circondato da valori negativi (es. Laplaciano).

Alternativa concettuale (riferimento alla nota "1 - H"):
Un modo alternativo per ottenere un passa-alto è sottrarre il segnale "passa-basso" (smoothato) dal segnale originale:

$$HighPass = Originale - LowPass$$

Questo concetto rimuove le basse frequenze dall'immagine originale, lasciando solo i dettagli (alte frequenze).

[transformations]
4a) Cosa sono le trasformazioni geometriche? fornire esempio.

4b) Qual è la differenza tra trasformazione geometrica e non geometrica? fornire esempi.

4c) Data una trasformazione a scelta, fornire algoritmo che la implementi elencando i passi.

a) Le trasformazioni geometriche sono definite come trasformazioni di coordinate. Modificano la posizione dei pixel all'interno dell'immagine, ma non il loro valore (intensità o colore). Matematicamente, si esprimono come una funzione che mappa le coordinate di input (x,y) in coordinate di output (u,v): $I_{in}(x,y) \rightarrow I_{out}(u,v)$ con $(u,v) = T(x,y)$. Un esempio è la traslazione, dove le coordinate vengono spostate di un vettore $t = (t_1, t_2)$, rendendo $q_1 = p_1 + t_1$, $q_2 = p_2 + t_2$

b) Una trasformazione geometrica agisce sulla posizione spaziale dei pixel. I pixel vengono spostati, ma il valore numerico che trasportano rimane idealmente invariato rispetto all'originale. Esempi sono traslazione, rotazione, scalatura, affine, euclidea.

Una trasformazione non geometrica modifica il valore del pixel e non la sua posizione. Esempi sono operazioni sull'istogramma (equalizzazione), modifica della luminosità o contrasto, ...

C)

1. **Iterazione sull'output:** Si considerano le coordinate (u, v) di ogni pixel dell'immagine di destinazione (output) che si vuole costruire.
2. **Calcolo coordinate inverse:** Per ogni pixel (u, v), si applica la trasformazione inversa $\mathcal{T}^{-1}$ per trovare le coordinate corrispondenti (x, y) nell'immagine originale (input):

   $$\mathcal{T}^{-1}(u, v) = (x, y)$$

   Nota: Il punto (x, y) risultante spesso non cade esattamente su coordinate intere della griglia discreta.
3. **Identificazione dei vicini:** Si individuano i pixel interi vicini al punto calcolato (x, y) nella griglia originale. Per l'interpolazione bilineare, si considerano i 4 vicini: gA, gB, gC, gD.
4. **Calcolo dei pesi:** Si calcolano le distanze frazionarie t e u del punto (x, y) rispetto ai vicini interi.
5. **Interpolazione (Assegnazione valore):** Si calcola il valore del nuovo pixel ($gNEW$) combinando i valori dei vicini pesati dalle distanze calcolate. La formula dell'interpolazione bilineare è:

   $$gNEW = (1-t)(1-u)gA + u(1-t)gB + t(1-u)gC + utgD$$

6. **Assegnazione:** Il valore $gNEW$ viene assegnato al pixel (u, v) dell'immagine di output.

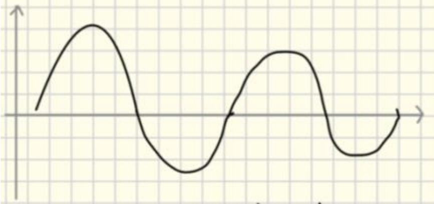
1a) Disegnare il grafico di un segnale continuo a tempo continuo, un segnale continuo a tempo discreto, un segnale discretizzato a tempo discreto.

1b) Definire l'operazione di quantizzazione.

1c) Definire la delta di dirac e spiegare come può essere usata nell'operazione di campionamento.

a)

segnale continuo a tempo continuo



segnale continuo a tempo discreto

segnale discretizzato a tempo discreto

b)

c)

La Delta di Dirac è uno (distribuzione) strumento matematico usato per rappresentare un impulso ideale concentrato in un singolo istante del tempo. È fondamentale per descrivere matematicamente il campionamento, poichè permette di rappresentare un segnale continuo come una sequenza di impulsi discreti.

continuo

$$\delta(t) = \begin{cases} \infty & \text{se } t=0 \\ 0 & \text{se } t \neq 0 \end{cases}$$

discreto

$$\delta(x) = \begin{cases} 1 & \text{se } x=0 \\ 0 & \text{se } x \neq 0 \end{cases}$$

[sound-processing]
2a) Spiegare cos'è la STFT (Short-Time Fourier Transform).

a)

La STFT è uno strumento matematico che analizza segnali non stazionari, ossia il cui spettro varia nel tempo. La STFT usa una finestra di durata finita, e funziona come una successione di FT calcolate su segmenti del segnale

[images]
3a) ~~Data la seguente immagine filtrata nei 3 colori RGB~~, definire cos'è lo spazio RGB e ~~stabilire di che colore è l'immagine originale.~~

3b) Se si volesse separare l'immagine dell'oggetto dallo sfondo, che algoritmo si potrebbe usare? fornirlo elencando i passi.

3c) Definire la tecnica addittiva e sottrattiva, fornire esempi di entrambe.

a)

Lo spazio RGB è un modello di colore basato su tre colori: rosso, verde e blu. Viene di solito rappresentato tramite un cubo dove i valori di intensità variano tipicamente nell'intervallo [0, 255]. È un modello di tipo additivo: i colori primari vengono sommati per produrre nuove tinte e, se sommati tutti insieme, producono il bianco. È tipico dei dispositivi di acquisizione e visualizzazione, come i monitor.

b)

Per separare un oggetto dal contorno, ovvero identificare i contorni che delimitano regioni dell'immagine, si possono usare filtri passa-alto e il calcolo del gradiente. Un algoritmo semplice è la sogliatura sul modulo del gradiente.

1) Convoluzione: data un'immagine f, si calcolano le differenze finite nelle direzioni x e y tramite convoluzione.
2) Calcolo delle componenti: si ottengono le componenti del gradiente:
   - $g_x = S_x * f$ (variazioni verticali procedendo lungo le x)
   - $g_y = S_y * f$ (variazioni orizzontali procedendo lungo le y)
3) Calcolo del modulo: si calcola il modulo M del gradiente, che descrive l'intensità delle variazioni del segnale, usando la formula: $M = \sqrt{g_x^2 + g_y^2}$.
4) Sogliatura: si applica una soglia al modulo M per mettere in evidenza i punti dove si è verificata un'alta variazione di intensità luminosa, ossia i bordi.

c)

La tecnica additiva parte dall'assenza di luce (nero) e si aggiungono i colori. Un nuovo colore viene ottenuto attraverso la miscelazione (somma) di emissioni di luce relative ad altri colori. Un esempio è il modello RGB.

La sintesi sottrattiva parte dal bianco, ossia la presenza di tutte le lunghezze d'onda, e si "eliminano" dei colori per ottenerne altri. Questo avviene facendo passare la luce attraverso filtri o pigmenti che assorbono determinate lunghezze d'onda e ne riflettono altre. Il modello CMY e CMYK (delle stampanti) sono esempi di sintesi sottrattiva.