Trattamento e Codifica di Dati Multimediali

Ilaria Battiston

Anno scolastico 2018-2019

Contents

1 Multimedia

Multimedia: utilizzo di diversi mezzi che concorrono insieme, solitamente in maniera interattiva, per trasferire informazione. I media hanno caratteristiche e standard differenti per essere catturati, immagazzinati, manipolati e trasmessi.

I segnali devono essere trattati tenendo conto del **sistema percettivo**. La tecnologia multimediale cerca di simulare il sistema percettivo umano, cioè minimizzando il quantitativo di dati da processare per ottenere un'informazione.

Il trattamento dei media si basa sull'analisi multimodale: essa permette di trarre conclusioni su contenuti di diversi tipi e di individuare stimoli legati a segnali. Per questo scopi devono essere introdotte misure oggettive e soggettive.

Un altro aspetto del multimodale è l'analisi da segnali che provengono da diverse fonti, come i dati fisiologici (ottenuti tramite sensori).

2 Segnali

I segnali possono essere classificati in base a **dominio** e **codominio**.

Dominio:

- $D = R$, segnale a tempo (spazio) continuo $x(t), t \in R$ (possono assumere tutti i possibili valori reali);
- $D = K$, segnale a tempo discreto, $x(t), t \in K$ con K numerabile, $K \in \{\dots, t_{-1}, t_0, t_1, \dots\}$. Il numero di valori t può essere comunque infinito.

Codominio:

- $C = R$, segnale continuo nelle ampiezze;
- $C = K$, segnale discreto nelle ampiezze con K numerabile e tipicamente finito $\{x_1, x_2, \dots, x_n\}$.

La variabile dipendente è definito sul dominio, quella indipendente sul codominio. I segnali possono essere reali o complessi (parte reale e immaginaria oppure modulo e fase, strettamente in relazione tra loro).

Le ampiezze tipicamente sono un insieme finito di valori, quindi al crescere di t (variabile indipendente, infiniti valori) il **range** di $s(t)$ (variabile dipendente) sarà limitato, anche se composto da infiniti valori. Il **passo** (distanza tra un campione e il successivo) dev'essere costante.

La quantizzazione è uno step intermedio della trasformazione tra segnale analogico e digitale, cioè la divisione del tempo (spazio) in passi e l'approssimazione nel dominio. La definizione del passo deve preservare la qualità del segnale, e va stabilita una finestra temporale da considerare.

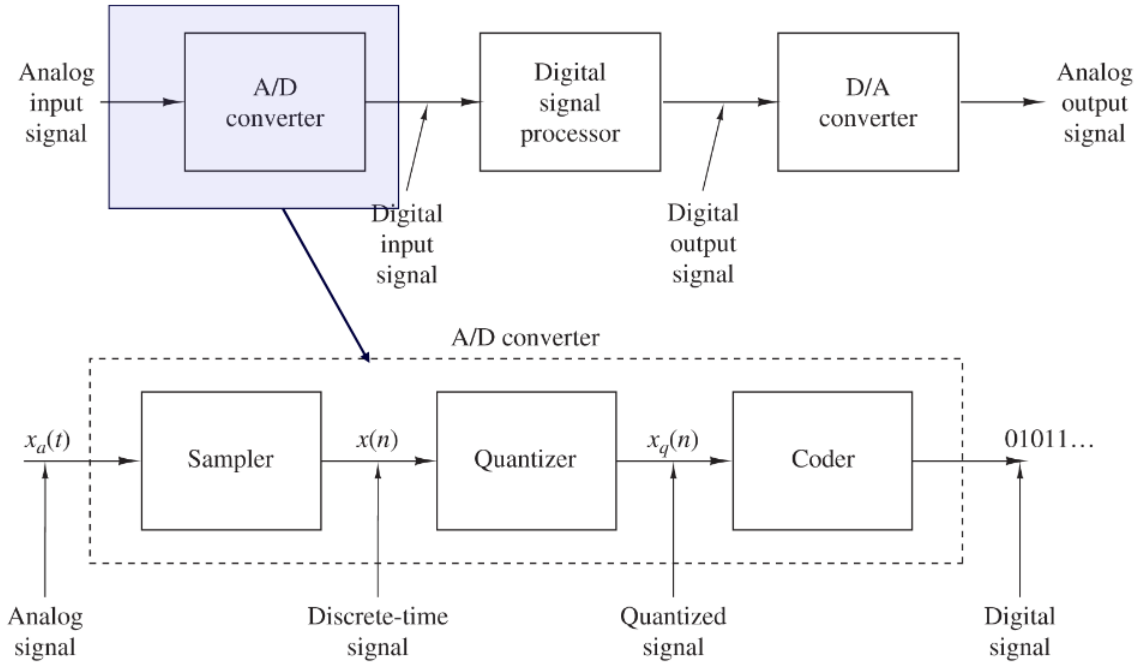
Le fasi della rappresentazione da analogica a digitale sono **campionamento**, **quantizzazione** e **codifica**.

Bisogna tenere conto di alcuni aspetti delicati, come i limiti di memoria, banda e tempi di processing: questi sono fattori importanti per determinare passo e finestra temporale. I valori cambiano anche a seconda del campo di analisi, per ottenere un'adeguata quantità di informazioni.

3 Multimedia processing

Per un efficace trattamento dei segnali multimediali è necessario minimizzare il quantitativo di dati processati, individuando solo quelli strettamente indispensabili. Oltre alla conversione e la compressione ci sono fasi come la memorizzazione e la trasmissione.

L'output è generalmente diverso dall'input: il segnale analogico x_a passa attraverso un campionatore, da cui esce come x_n (non più analogico). L'output $x_q(n)$ contiene un insieme discreto di valori (sottoinsieme del dominio) che poi verrà trasformato in segnale quantizzato $x_q(n)$ e poi in bit.



Il valore di $x(n)$ è il valore della funzione analogica preso nT volte, dove T è il passo di campionamento. La frequenza è impossibile da ottenere senza la dimensione del passo, ed è legata a quanto velocemente varia il segnale.

L'obiettivo è capire se esiste una soglia in grado di stabilire se le informazioni vengono perse in base alla dimensione del passo. Avere un'alta **frequenza di campionamento** significa avere una buona qualità ma un volume elevato dei dati, mentre una bassa frequenza produce fenomeni di aliasing (approssimazione a costante).

Un segnale qualsiasi è rappresentabile come integrali di infiniti termini (seni e coseni) con peso e ampiezza diversi. Tra essi bisogna preservare quello con frequenza massima, per evitare sovrapposizioni e cambiamenti. La scomposizione del segnale è effettuata tramite **analisi di Fourier**.

Il **campionamento** con frequenza massima consente inoltre di riconvertire il segnale digitale in analogico senza perdita di informazione, perché permette la conservazione delle frequenze.
La **quantizzazione**, invece, è un'operazione che comporta sempre perdite, quindi non è reversibile.

Se il numero di frequenze tende a infinito, com'è possibile individuare la massima? In questo caso si introduce il **filtering**: l'eliminazione dei valori esclusi da una certa soglia (superiore o inferiore). Un altro motivo del **filtro anti-aliasing** è l'eliminazione delle frequenze troppo basse.

Un esempio di elaborazione numerica è il filtraggio, l'eliminazione delle frequenze fuori dal range accettabile (filtro passa-basso o alto). I sistemi sono combinazioni di più operatori lineari, quindi nonostante complessi sono scomponibili a causa delle loro proprietà.

4 Segnali

Un segnale rappresenta il comportamento di grandezze fisiche in funzione di una o più variabili indipendenti. Sono monodimensionali se rappresentati da una sola variabile, per esempio il suono (continuo). I dati EEG sono multidimensionali in variazione al tempo, agli elettrodi e ai soggetti.

Le immagini in bianco e nero sono segnali bidimensionali (coordinate spaziali) e monocanale (il grigio), mentre quelle a colori hanno 3 segnali dimensionali RGB. Il campionamento corrisponde al numero di pixel, e la quantizzazione è la profondità del colore (quanti bit per la codifica). Aumentando il numero di livelli, aumenta la capacità di rappresentare l'informazione.

Se la variabile indipendente continua viene discretizzata è stato effettuato un campionamento, in cui è necessario conoscere la distanza tra i campioni (digitali).

Il valore assunto dal segnale si definisce ampiezza (dipendente, codominio) mentre l'asse delle ascisse è il dominio (tempo o spazio). Si possono introdurre grandezze statistiche come media e varianza, indicate in modo diverso a seconda del tipo di segnale.

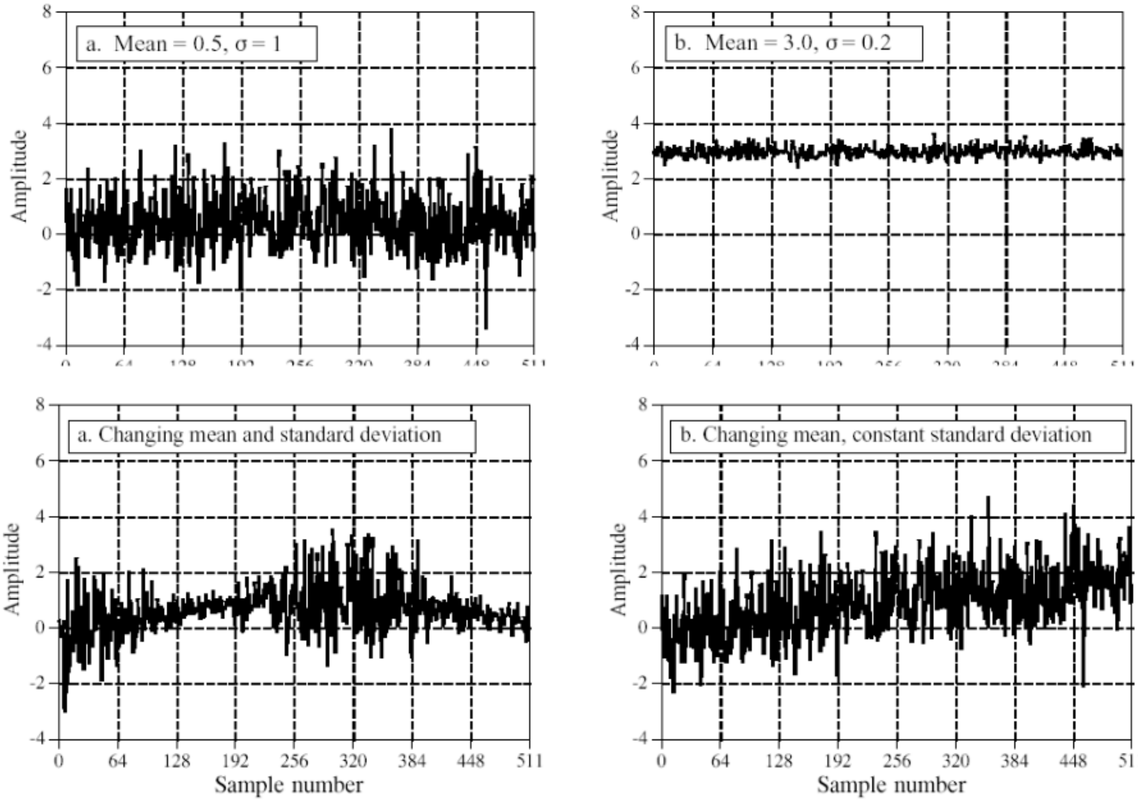
- Continuo:
  - $\mu = \lim\limits_{T \to \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} x(t)dt$;
  - $\mu = \frac{1}{T_1 - T_0} \int_{T_0}^{T_1} x(t)dt$;
- Discreto:
  - $\mu = \frac{1}{N} \sum_{i=0}^{N-1} x_i$.

Il segnale digitale ha solamente una casistica di μ perché non tende mai a infinito, e il livello dell'ampiezza è diverso. Se un segnale varia ha una componente continua DC (direct current), contributo a frequenza 0 e valore medio, e componenti AC a corrente alternata che variano in base a come il segnale fluttua intorno al valore medio.

Una forma d'onda ripetuta ha escursioni costanti, descritte da una grandezza chiamata ampiezza picco-picco A_{pp}. Il periodo è arbitrario.

Deviazione standard e varianza forniscono informazioni aggiuntive su quanto lontano (e con quale potenza) il segnale fluttua dal valore medio. Alla varianza è fortemente legata la potenza del rumore: un'alta varianza implica un forte rumore.
$\sigma^2 = \frac{1}{N-1} \sum_{i=0}^{N-1} (x_i - \mu)^2$

Se media e varianza di un segnale non cambiano nel tempo, esso è stazionario. Al contrario, se il segnale varia la media sarà diversa a seconda della finestra, e la media globale non dà informazioni.



La **periodicità** indica la ripetizione del segnale nel tempo, definito appunto in periodi. Non esistono segnali puramente periodici, ma si usano approssimazioni delle forme d'onda che assume il segnale. L'inversa del periodo è chiamata **frequenza fondamentale**: $f_0 = 1/T$.

4.1 Segnale sinusoidale

Un segnale sinusoidale è un seno o un coseno, monodimensionale in funzione del tempo o dello spazio.

$$A(T) = A_{med} + B \cdot \sin(2\pi f t + \varphi_0)$$

$$A_{med} = \frac{1}{T}\int_0^T A(t)dt \qquad A_{pp} = A_{max} - A_{min} \qquad B = A_{pp}/2$$

Parametri importanti in questi casi sono frequenza e fase.
La frequenza si misura in Hertz, e rappresenta la rapidità con cui varia l'ampiezza in un intervallo temporale T. La pulsazione (intera variazione di ampiezza) è proporzionale alla frequenza, si ha che $\omega = 2\pi f$.
La fase segna l'alternarsi di positività o negatività del segnale, in particolare è significativa la fase iniziale φ_0.

$$P(t) = |x(t)|^2 \qquad \text{potenza istantanea}$$

$$E_x = \int_{-\infty}^{+\infty} P(x)dt \qquad \text{energia: area di potenza istantanea}$$

Tanto più l'ampiezza si scosta dallo 0, la potenza aumenta. Se $E_x < \infty$ il segnale ha energia finita. Quando $E_x = \infty$ si definisce la potenza media: $P_x = \lim\limits_{T\to\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} |x(t)|^2 \, dt$

Potenza media di un segnale periodico T: $P_x = \frac{1}{T} \int_T |x(t)|^2 \, dt$

Entrambi i valori hanno una componente continua (valore medio) calcolato come il limite dell'integrale o l'integrale (segnale periodico) di $x(t)$.

4.2 Decibel

Il **decibel** dB è un'unità di misura logaritmica (quindi non lineare), di cui lo scopo del logaritmo è visualizzare meglio grandi scale di valori e avvicinarsi alla percezione umana. La misura è quindi relativa e adimensionale.

$$deciBel = dB \longrightarrow 10 \log_{10} \frac{P_1}{P_2} = 20 \log_{10} \frac{A_1}{A_2}$$

$$Bel \longrightarrow \log_{10} \frac{P_1}{P_2} \qquad P = \text{potenza} \qquad A = \text{ampiezza} \qquad P \propto A^2$$

Le componenti della formula sono due pressioni, di cui il numeratore è la potenza del suono e il denominatore è la soglia minima di udibilità. La potenza è proporzionale al quadrato dell'ampiezza, quindi trasformando in scala logaritmica essa diventa un fattore moltiplicativo. 6dB rappresentano un raddoppio dell'ampiezza.

4.3 Trasformazioni di segnali

Operazioni molto comuni sono la traslazione, il cambio di scala e l'inversione temporale.

- Ritardo: fissato un tempo t_0, la traslazione trasforma il segnale $x(t)$ nel segnale $x(t - t_0)$;
- Anticipo: fissato un tempo t_0, la traslazione trasforma il segnale $x(t)$ nel segnale $x(t + t_0)$;
- Cambio di scala: fissato un numero reale $a > 0$, la scalatura trasforma il segnale $f(t)$nel segnale $f(at)$;
  - Se $a > 1$ si ottiene una compressione lineare;
  - Se $a < 1$ si ottiene un allungamento lineare;
- Inversione: trasforma il segnale $f(t)$ nel segnale $f(-t)$.

Se il segnale è reale, è possibile ritardarlo ma non anticiparlo. Un segnale si dice pari se $f(t) = f(-t)$, dispari se $f(t) = -f(-t)$: il coseno è una funzione pari, il seno è dispari.

4.4 Segnali continui

Gradino: usato per selezionare la parte positiva dei segnali che tendono a $\pm\infty$. Si definisce un gradino **unitario** $u(t)$:

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

Gradino **traslato** in t_0: definito quando la finestra di osservazione è finita, centrata rispetto a 0. Se il segnale è fuori dall'intervallo assume valore 0:

$$u(t - t_0) = \begin{cases} 1 & \text{se } t \geq t_0 \\ 0 & \text{se } t < t_0 \end{cases}$$

Impulso rettangolare unitario $rect(t)$:

$$rect(t) = \begin{cases} 1 & \text{se } |t| \leq 1/2 \\ 0 & \text{se } |t| > 1/2 \end{cases}$$

Quest'ultima è una funzione che forma un rettangolo di area unitaria. Generalizzando, si ha un rettangolo di altezza A, base T e traslato in t_0 sostituendo nella formula unitaria t con $\frac{t-t_0}{T}$ e confrontandolo con $1/2$.

$$|t - t_0| \leq T/2 \begin{cases} \text{per } (t - t_0) > 0 & (t - t_0) \leq T/2 & t \leq t_0 + T/2 \\ \text{per } (t - t_0) < 0 & (t - t_0) \geq -T/2 & t \geq t_0 - T/2 \end{cases}$$

La moltiplicazione di un segnale per un rettangolo lo approssima con segmenti verticali o orizzontali a seconda dell'asse considerato. Media e varianza non sono le stesse rispetto alla funzione originale.

Funzione **delta di Dirac** $\delta(t)$: distribuzione con rettangolo di base infinitesima e altezza infinita che abbia l'area unitaria, con la larghezza che tende a 0 e di conseguenza l'altezza che tende a infinito.

$$\int_{-\infty}^{\infty} \delta(t)dt = 1$$

Limite dell'impulso rettangolare di base Δ per $\Delta \to 0$, dove Δ è uno scalare sull'asse x:

$$\delta(t) = \lim_{\Delta \to 0} \frac{1}{\Delta} \text{ rect}\left(\frac{t}{\Delta}\right) \qquad \delta(t - x) = 0 \qquad \text{se } t \neq x$$

Viene introdotta per rappresentare fenomeni fisici di durata infinitesima (impulsi).

Il corrispondente di δ discreto nel dominio è la **delta di Kronecker** o impulso unitario, rappresentato con una freccia verticale di altezza (o peso) unitario. Questa funzione ha significative applicazioni nel trattamento dei segnali.

Si ha un impulso $A\delta(n - n_0)$ di ampiezza A e che occorre al tempo $n = n_0$. Con $n - 2$, δ è in ritardo di 2 (se fosse + sarebbe anticipo).

Esistono anche l'analogo discreto della delta di Dirac e del gradino unitario:

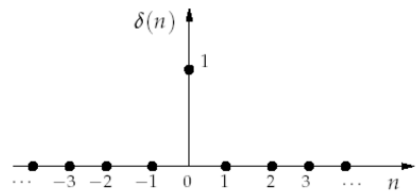
$$\delta(n) = \begin{cases} 1 & \text{se } n = 0 \\ 0 & \text{se } n \neq 0 \end{cases} \qquad u(n) = \begin{cases} 1 & \text{se } n \geq 0 \\ 0 & \text{se } n < 0 \end{cases}$$

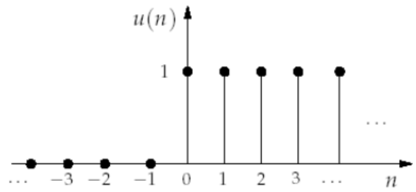
$x(n) \cdot \delta(n) = x(0)$

$f(n) \cdot \delta(n - n_0) = f(n_0)$

$\delta(n) = u(n) - u(n-1)$

Il gradino continuo, nel discreto diventa una successione di δ: $u(n) = \sum_{i=0}^{+\infty} \delta(n-i)$





5 Sequenze

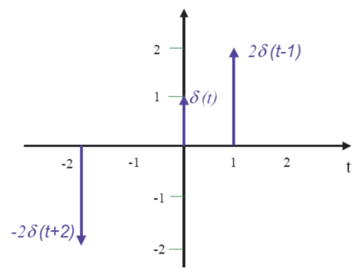
Le sequenze $x(n)$ sono formate da segnali a tempo discreto. Se essi sono quantizzati in ampiezza, si parla di segnale digitale.

Sequenza causale: $x(n) : n > 0$ (con numeri positivi).
Sequenza anticausale: $x(n) : n < 0$ (con numeri negativi).

Sequenza pari: $x(n) = x(-n)$ (coseno).
Sequenza dispari: $x(n) = -x(-n) \rightarrow x(0) = 0$ (seno).

Sequenza periodica: $x(n) = x(n+T)$
Sequenza limitata: $|x(n)| \leftarrow x_0 < \infty \quad \forall\, n$

Non è possibile tornare indietro nel tempo, quindi le sequenze considerate sono generalmente causali.



Nella figura, $\delta(n)$ rappresenta l'impulso unitario. Moltiplicando δ per un coefficiente k l'impulso avrà diversa lunghezza, o diversa direzione se k è negativo.
Un ritardo $t - t_0$ sposta l'impulso verso destra, un anticipo $t + t_0$ viceversa.

Se gli impulsi hanno un fattore moltiplicativo n, si definisce una funzione rampa: il segnale seguirà la diagonale.

6 Analisi di Fourier

L'analisi di Fourier decompone il segnale in costituenti sinusoidali di differenti frequenze. Il segnale non è più nel dominio tempo-spazio, ma delle frequenze: i dati sono gli stessi, cambia solo la rappresentazione.

Ogni funzione periodica e a quadrato sommabile può essere espressa come somma infinita e pesata di funzioni seno e coseno (combinazioni di funzioni armoniche).

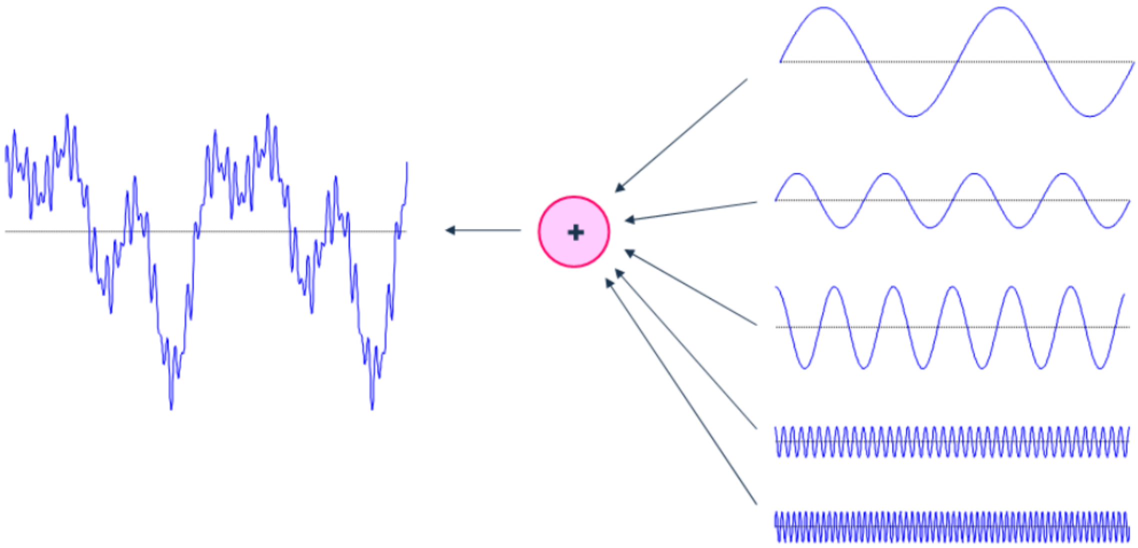
Si ricorda che una sequenza periodica è $x(n) = x(n+T)$. Una funzione armonica è una funzione periodica del tipo:

$$y = A\sin(\varpi x + \varphi) \qquad y = A\cos(\varpi x + \varphi)$$

Dove A è l'ampiezza, ϖ è la pulsazione, φ è la fase. Si ha che $\varpi = 2\pi/T$ dove $1/T$ è la frequenza, e π è 180°.

Sviluppando i seni e i coseni si ha, con $a = A\sin(\varphi)$ e $b = A\cos(\varphi)$:
$y = A\sin(\varpi x + \varphi) = a\cos(\varpi x) + b\sin(\varpi x)$
$y = A\cos(\varpi x + \varphi) = b\cos(\varpi x) + a\sin(\varpi x)$

Le armoniche vengono combinate una per volta, avvicinandosi man mano alla funzione originaria.



6.1 Serie di Fourier

La serie di Fourier è una formula utile per approssimare la scomposizione. Rappresenta una funzione periodica mediante combinazione lineare di funzioni sinusoidali.

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

$N \rightarrow$ periodo $\qquad (1/N) \rightarrow$ frequenza fondamentale f_0 $\qquad (1/N)k \rightarrow$ frequenze multiple kf_0

a_k e b_k sono numeri reali, k è un numero intero che funge da fattore moltiplicativo e N è l'ampiezza della parte di funzione che si ripete periodicamente (l'inverso è la frequenza fondamentale). x è la variabile indipendente.

Le frequenze (infinite), quindi, sono una parte fondamentale della formula. Con tempo $T = 1/f$ il dominio del tempo è t, e il dominio della frequenza è Hz. In altre parole, il segnale passa dal dominio spazio-tempo alle frequenze, con coppie frequenza-peso associate alla sinusoidale (ampia A).

Estendendo la formula con il concetto di pulsazione, si ha che una pulsazione è $\varpi_k = 2\pi f_k$, e di conseguenza:

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos(2\pi k f_0 x) + b_k \sin(2\pi k f_0 x)$$

$$a_k = \frac{2}{N} \int_{-\frac{N}{2}}^{\frac{N}{2}} f(x) \cos(2\pi k f_0 x) dx \qquad b_k = \frac{2}{N} \int_{-\frac{N}{2}}^{\frac{N}{2}} f(x) \sin(2\pi k f_0 x) dx$$

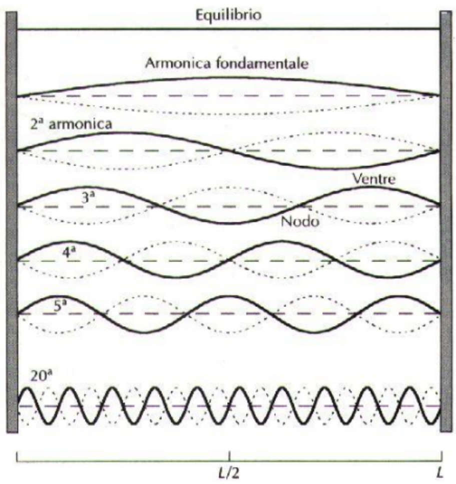
La variabile x e la funzione associata sono continue (periodiche), ma la variabile di frequenza è discreta (k intera), quindi dall'integrale si passa alla sommatoria.

Si ricorda che N è il periodo. Data la funzione $f(x)$ periodica, i coefficienti della serie sono **univocamente** determinati. I coefficienti sono i fattori moltiplicativi di seno e coseno, in relazione al tempo t.

Questo significa che esiste biunivocità, la trasformazione può essere effettuata da entrambi i versi senza perdere informazione.

6.2 Fourier e i suoni

I suoni elementari hanno andamento sinusoidale, periodico e con estensione indefinita; la maggior parte dei suoni in natura sono però caratterizzati da forme d'onda diverse.



Un segnale si definisce complesso quando è formato da più funzioni sinusoidali combinate, invece che una sola.

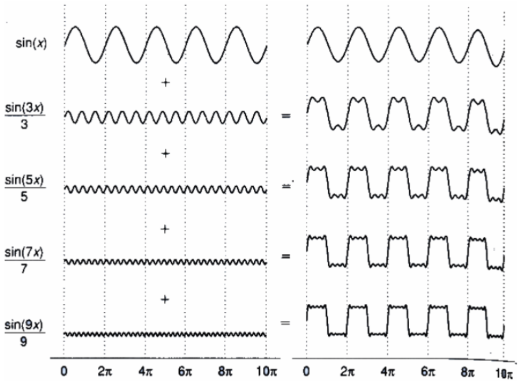
Si può dimostrare che, fatte alcune ipotesi di regolarità sull'andamento della forma d'onda, un generico suono complesso può essere descritto come una combinazione di suoni elementari (armoniche).

Le armoniche sono funzioni che, combinate fra loro, permettono di determinare il timbro degli strumenti musicali. Insieme costituiscono forme d'onda.

6.3 Onda quadra

L'onda quadra è un caso particolare in cui tutte le armoniche pari sono nulle, e l'onda quadra è data dalla forma delle componenti $F_0, 3F_0, 5F_0, \ldots$ cioè 200 Hz, 600 Hz, 1 Kz eccetera. Si ricorda che l'ampiezza di conseguenza è $1/3, 1/5, \ldots$ cioè la parte dispari della sommatoria.

Ognuna di queste componenti ha un'ampiezza diversa rispetto a quella dell'armonica fondamentale. La funzione è asimmetrica, e i termini pari sono appunto i coseni: ci sono solo seni non nulli.



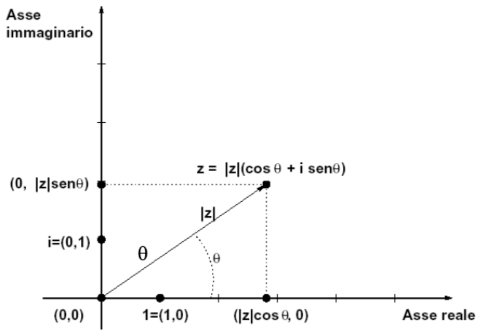
Esempio: un'onda quadra con periodo 5 ms può essere ottenuta sommando onde sinusoidali di opportuna frequenza, ampiezza e fase. Il contributo più rilevante è dato dalla prima sinusoide, si frequenza pari a quella dell'onda quadra (200 Hz, frequenza fondamentale F_0).

Per costruire l'onda quadra sono necessarie anche altre componenti elementari di frequenza maggiore, le armoniche multiple di F_0.

6.4 Numeri complessi e serie di Fourier

I numeri complessi sono rappresentabili su un piano cartesiano come l'intersezione tra numeri reali sull'asse x e numeri complessi sull'asse y. Ogni numero complesso ha quindi una parte reale a e una immaginaria b.

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



Forma polare:
$a = \pi \cos(\theta)$
$b = \pi \sin(\theta)$
$\theta = \arctan\left(\frac{b}{a}\right)$

Le coordinate polari corrispondono a quelle cartesiane (proiezione).

Applicando i numeri complessi alla serie di Fourier e quindi cambiando dominio, si possono usare gli esponenziali complessi e la formula di Eulero.

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

Sia seno che coseno sono somme di due esponenziali complessi, con segno + e − rispettivamente. Il picco della frequenza è intero per il coseno, immaginario per il seno.

$$\cos\theta = \frac{1}{2}(e^{j\theta} + e^{-j\theta}) \qquad \sin\theta = -\frac{j}{2}(e^{j\theta} - e^{-j\theta})$$

La parte del seno è immaginaria, il coseno è reale. Le formule sono le stesse, ma applicate al sovrainsieme dei complessi $\mathbb{C}$ (j e i sono indicatori convenzionali della parte immaginaria).

Mettendo a sistema seno e coseno ottenuti in questo modo, si può estendere la serie di Fourier rappresentando le somme di contributi nello spazio degli esponenziali complessi.

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos(2\pi f_k x) + b_k \sin(2\pi f_k x)$$

$$f(x) = \sum_{-\infty}^{\infty} R_k e^{j(2\pi f_k x)}$$

Si ricorda che $\theta = 2\pi f_k$. Quando θ cambia segno, l'angolo si sposta. Essendo il coseno pari, questo non cambia; il seno è dispari quindi è l'opposto.

Per passare dalle variabili discrete alle continue, si usa il coefficiente R_k che raccoglie le due sommatorie (fino a infinito da entrambi i versi) e si applica l'integrale.

$$R_k = \int_{-\frac{N}{2}}^{\frac{N}{2}} f(x)e^{-j(2\pi f_k x)}dx$$

Il coefficiente $\frac{A_n}{2}$ (altezza del segnale sull'asse y) è dato appunto dalla divisione del coefficiente R in 2 parti, seno e coseno: l'ampiezza sarà la metà, e ogni ampiezza viene sommata insieme alle funzioni. C'è corrispondenza biunivoca fra dominio temporale e dominio delle frequenze. I seni hanno solo la parte complessa.

6.5 Trasformata di Fourier di funzioni continue

Per passare dalle funzioni periodiche a quelle reali, la sommatoria diventa un integrale. La trasformata di Fourier serve per rapppresentare le funzioni continue come sinusoidi.

Ogni funzione continua $f(x)$ anche se non periodica (purché abbia area finita) può essere espressa come integrale di sinusoidi complesse opportunamente pesate. L'antitrasformata ha dominio spaziale (temporale, diretto) mentre la trasformata ha dominio delle frequenze.

$$\text{Antitrasformata di Fourier: } f(x) = \int_{-\infty}^{\infty} F(u)e^{j2\pi ux}dx$$

$$\text{Trasformata di Fourier: } F(u) = \int_{-\infty}^{\infty} f(x)e^{-j2\pi ux}dx$$

$F(u)$ sostituisce il coefficiente discreto R_k, u è la frequenza. La trasformata di Fourier non si può applicare a tutte le funzioni, ma solo a quelle con quadrato sommabile (trovare la formula) e continue.

Dalla trasformata di Fourier, la funzione originiaria può essere ricostruita senza perdita di informazione attraverso il processo inverso: è possibile spostarsi dal dominio spaziale a quello delle frequenze e poi tornare indietro senza aver perso informazione.

Esiste un valore della trasformata (non più δ) per ogni funzione continua. Invertendo la funzione, quindi, se è a supporto finito non ci sarà 2δ.

Piuttosto che guardare parte reale e immaginaria, si considerano modulo e fase, ricavabili direttamente. Il modulo indica la potenza del segnale, e la fase assicura la biunivocità della corrispondenza.

$$F(u) = F[f(x)] = \Re(u) + j\Im(u) = |F(u)|\, e^{j\phi(u)}$$

$$\text{Spettro: } |F(u)| = [\Re(u)^2 + \Im(u)^2]^{1/2}$$

$$\text{Fase: } \phi(u) = \tan^{-1}\left[\frac{\Im(u)}{\Re(u)}\right]$$

$$\text{Potenza (densità) spettrale: } |F(u)|^2 = \Re(u)^2 + \Im(u)^2$$

La trasformata di Fourier è un operatore scomponibile, quindi per passare da una dimensione a n semplicemente si applica n volte la formula. In altre parole, si scompone rispetto a una direzione la funzione.

6.6 Trasformata di Fourier di funzioni discrete

Il primo step è il campionamento: in questo caso si utilizza la sommatoria, ma la funzione continua è rappresentata da un numero discreto di campioni. La funzione è definita per campioni i, e la variabile indipendente x assume valore complesso $i\Delta x$.

$$F(u) = \int_{-\infty}^{+\infty} f(x)e^{-j2\pi ux}dx \Longrightarrow F(u) = \sum_{i=-\infty}^{\infty} f(i)e^{-j2\pi ui\Delta x}$$

Limitando la funzione spazialmente, il passo Δx per convenzione diventa $1/N$, e la sommatoria arriva fino a $N-1$ invece che infinito. $f(i)$ è la funzione campionata (a partire da $f(x)$ continua): la trasformata è periodica. Si ha $f(i) = f(x_0 + i\Delta x)$.

$$F(u) = \frac{1}{N}\sum_{i=0}^{N-1} f(i)e^{-j2\pi u\frac{1}{N}i}$$

Per campionare le funzioni $f(x)$ e $F(u)$ si applicano le trasformate di Fourier discrete (DFT). Il segnale campionato è periodico, quindi la funzione inversa è comunque periodica con distanza $1/\Delta x$. Se il passo tende a 0, la funzione tende a infinito (senza repliche) e viceversa.

In altre parole, da una funzione campionata $\Delta x = 1/N$ si passa a una funzione continua e periodica di periodo N.

Il fattore $1/N$ davanti alla trasformata o all'antitrasformata è un fattore di normalizzazione per il dualismo fra spazio diretto e spazio trasformato. Di conseguenza, $\Delta u = \frac{1}{N\Delta x}$.

Un ragionamento analogo si applica per passare da una dimensione a due: la sommatoria è doppia e si divide per entrambi i fattori di normalizzazione. Entrambe le variabili rispettano la periodicità, cioè $F(u,v) = F(u+M, v+N)$ con $\Delta u = 1/M\Delta x$, $\Delta v = 1/N\Delta y$.

6.7 Elaborazione di immagini

Il segnale di un'immagine viene scomposto, come visto precedentemente, in modulo e fase. il modulo raccoglie le frequenza di un'immagine, mentre la fase contiene l'informazione posizionale di ogni in un'immagine (le frequenze sono le stesse a prescindere dall'organizzazione spaziale). L'ampiezza $\angle$ contiene informazioni relative alle posizione o meno delle strutture nell'immagine.

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

I pixel possono assumere valore da 0 a 255, mappati in una trasformata di Fourieri discreta. Data la periodicità e la simmetria dello spettro rispetto all'origine, è possibile scegliere il periodo dello spettro centrato sull'origine $(0,0)$.

In questo modo, il picco rappresenza la componente continua (valore medio della funzione), e le frequenze crescono allontanandosi dal picco. Per convenzione lo spettro (in alto a sinistra) si trasla in modo da essere al centro, e le alte frequenze sono verso gli angoli esterni.

Il processo è reversibile: l'immagine, di solito sotto forma di segnale, se viene ricostruita dalla trasformata è identica all'originale.

Un operatore puntuale ha una corrispondenza biunivoca non indipendente dall'ambiente, cioè ogni pixel dell'immagine originale viene mappato nel modulo nello stesso modo a prescindere dai pixel intorno.

Al contrario, la trasformata di Fourier funziona applicando sommatorie (per il segnale), quindi ogni valore dipende da tutti i precedenti. Ciò si definisce operatore locale: ogni frequenza ha una sua influenza, e sebbene ci possano essere corrispondenza tra linee e forme, ogni pixel nel dominio è ricavato in base a tutti i pixel nel codominio (e viceversa).

L'istogramma di un'immagine a livelli di grigio è una funzione discreta $h(r_k) = n_k$ dove r_k è il livello di grigio k-simo, e n_k è il numero di pixel nell'immagine con intensità r_k (per ogni livello).

Il numero di campioni stabilisce l'informazione che si può rappresentare: se è sufficientemente alto, la distribuzione dei toni di grigio sarà regolare. Questo non permette di trarre conclusioni sullo spazio, ma semplicemente su quanti campioni hanno un certo valore (e di conseguenza la bontà del quantizzatore).

Se ci sono picchi nella distribuzione, esso è dovuto alla ripetizioni di pattern nell'immagine (come lo sfondo). Il numero di campioni nella trasformata è uguale a quello nell'immagine, e con frequenza di campionamento bassa la Δf sarà piccolo e si verificherà l'aliasing.

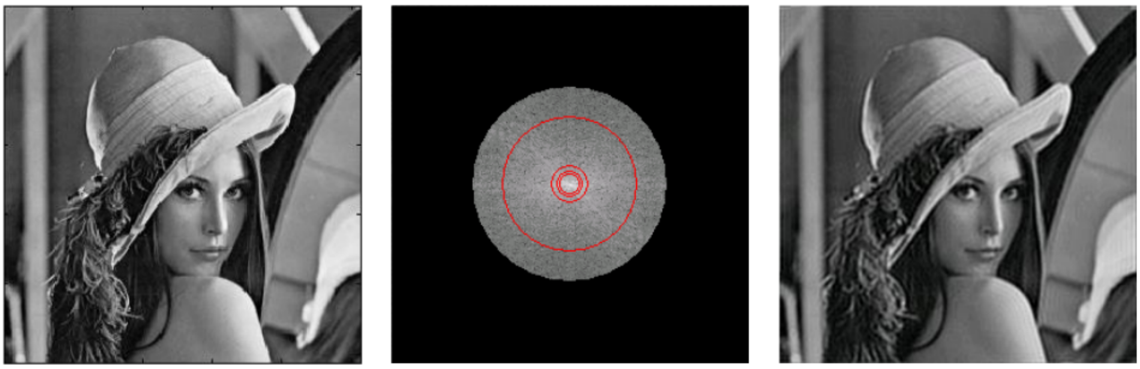
Per rimuovere i rumori si acquisiscono porzioni di segnali più lunghe, eventualmente analizzando le loro medie in modo da preservare le componenti più importanti.

Lo spettro è il risultato della trasformata, di cui una parte è reale e una immaginaria. Lavorando sul modulo, è possibile individuare circonferenze che indicano dove si concentra la potenza del segnale: manipolando lo spettro il risultato dell'inversione sarà differente.

Alcune metodologie di filtraggio comportano l'applicazione di maschere, finestre (monodimensionali, rappresentate come un rettangolo nel piano) a sezione circolare. Modificare lo spettro, quindi, implica ricostruire una diversa immagine che sarà più o meno accurata rispetto all'originale.

Spesso si possono eliminare le ridondanze senza perdere informazioni. Su questi principi si basano gli algoritmi di compressione (es. JPEG), riducendo la complessità a discapito della qualità.

Si ha che il 99.5% dello spettro è contenuto in una circonferenza di dimensione limitata.



Originale Spettro (99.5%) Ricostruita

Eliminando parte dello spettro e delle frequenze, il filtro diventa parte integrante del segnale che viene antitrasformato: ciò introduce fenomeni di ringing, quindi è necessario osservare anche il comportamento dell'antitrasformata del filtro.

Per ottenere diversi effetti si selezionano diverse parti dello spettro da tenere, per esempio solo l'interno o l'esterno, o una banda di un certo intervallo (mosaico, smoothing).

6.8 Proprietà della trasformata

Un'importante proprietà è la linearità, che vale per ogni dominio: la trasformata di Fourier applicata a una funzione che è combinazione lineare (somma pesata) di funzioni, il risultato è uguale alla trasformata di Fourier di ogni funzione presa singolarmente.

$$F(f)x) + g(x) = F(f(x)) + F(g(x)) = F(u) + G(u)$$

La linearità permette di ridurre (scomporre) la complessità della trasformata come sovrapposizione di sistemi più semplici e monodimensionali, senza introdurre dispersione del segnale.

Vale anche la traslazione nello spazio (tempo) e nelle frequenze, ma solo per la trasformata discreta: una traslazione equivale a una modulazione nelle frequenze con un esponente complesso. In altre parole, la trasformata del modulo è identica, ma la fase cambia di x_0 (sfasamento).

$$F(f(x + x_0)) = e^{j2\pi ux_0} \qquad F^{-1}(F(u \pm \varpi)) = f(x)e^{\pm j2\pi\varpi x}$$

A oggetti più grandi nel dominio diretto, corrisponderanno oggetti simili ma compressi nel dominio trasformato e viceversa, a causa dell'aumento delle frequenze.

$$\text{Scala: } F(f(ax)) = \frac{1}{|a|}F\left(\frac{u}{a}\right) \implies f(\alpha x, \beta y) = \frac{1}{|\alpha\beta|}F\left(\frac{u}{\alpha}, \frac{v}{\beta}\right)$$

$$\text{Inversione: } F(f(-x)) = F(-u)$$

La trasformata di una funzione reale gode di simmetria Hermitiana: la parte reale e il modulo sono simmetrici rispetto all'origine (lo 0), mentre la parte immaginaria e la fase sono antisimmetriche rispetto all'origine.

Campionando il segnale di partenza, la trasformata diventa periodica e il periodo permette di ricostruire il segnale, a sua volta periodica.

$$\text{Periodicità (discreta): } F(u) = F(u + M) \implies F(u) = \frac{1}{M}\sum_{x=0}^{M-1} f(x)e^{-j\frac{2\pi}{M}ux} \quad u = 0, \ldots, M-1$$

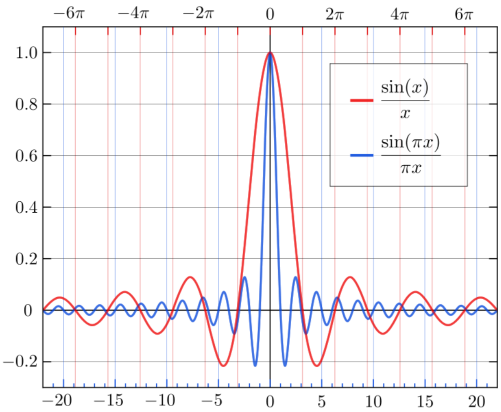
Altre proprietà sono già state enunciate, come l'esistenza dell'inversa della trasformata discreta e la reversibilità.

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

$$\text{Valore medio di una trasformata discreta: } F(0) = \frac{1}{M}\sum x = 0^{M-1}f(x) = \overline{f}(x)$$

$F(u_0)$ è il peso con cui l'onda complessa $e^{j2\pi u_0 x}$ di frequenza u_0 concorre a formare il segnale $f(x)$. Se $f(x)$ è un'onda complessa di frequenza u_k, $F(u)$ sarà 0 per ogni frequenza diversa da u_k.

Nel dominio del coseno trasformato non ci saranno solo le δ se il segnale non è periodico e infinito, ma ci sarà anche l'effetto del *sinc* (finestra). $sinc(x) = \frac{\sin(x)}{x}$, con un fattore π a per normalizzarla [immagine: Georg-Johann].

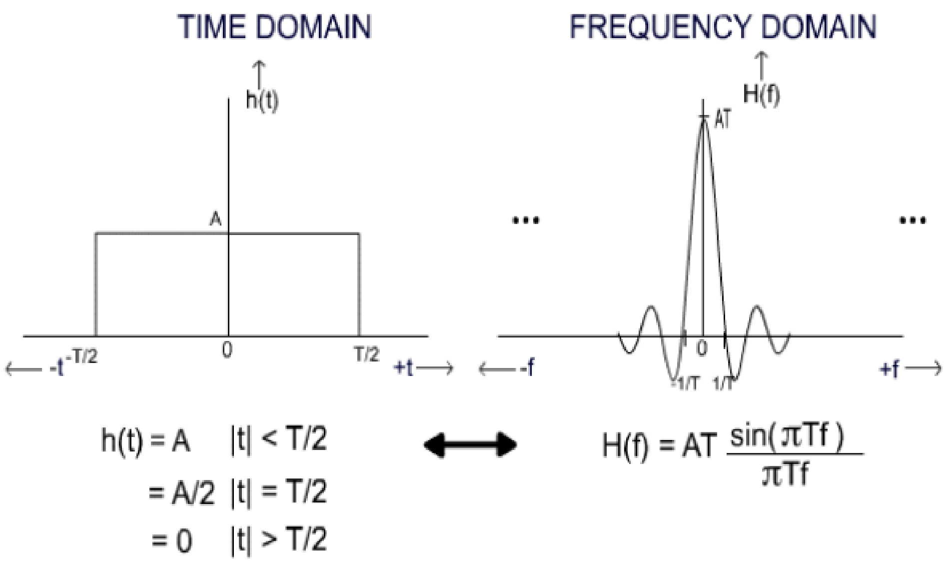


Questa funzione rappresenta il limite con x che tende a 0 della funzione seno, cioè 1. Nel processamento dei segnali è la trasformata di Fourier della funzione rettangolare (finestra) non scalata, ricostruendo il segnale continuo a partire dai campioni.

Per il seno, c'è una δ positiva e una negativa, e se trasformando entrambe diventano positive non c'è più la possibilità di distinguerlo dal coseno. In realtà le due funzioni sono la stessa traslata (proprietà di traslazione). Bisogna però ricordare che la δ positiva della parte immaginaria in realtà corrisponde alla parte negativa.

L'altezza di δ nel punto 0 è uguale all'altezza del segnale, che varia sempre meno fino a diventare costante. La trasformata di una costante, quindi, è rappresentabile come un coseno le cui δ si sovrappongono nel centro (limite della funzione, con frequenza che tende a 0).

La finestra di osservazione nel tempo ha frequenza δ, con due picchi di seno e coseno, e il contributo di *sinc* cresce con l'aumento dell'ampiezza della finestra (ringing). Al contrario, *sinc* si avvicina sempre di più a δ con il restringimento.



7 Campionamento

Il campionamento è un processo che legge valori distanziati l'uno dall'altro con un passo Δt approssimato al tempo infinitesimo, ma in realtà discreto. La scelta dell'ampiezza del passo deve evitare sia lo spreco di risorse che la perdita di informazioni.

Il segnale, in pratica, passa dall'analogico al digitale e dal continuo al discreto, e viene rappresentato come una serie di impulsi.

A una funzione campionata (con passo Δx) corrisponde nel dominio trasformato uno spettro periodico. Il periodo è inversamente proporzionale al passo di campionamento (ricostruzione della funzione continua dalla trasformata).

Così come a una funzione continua nel dominio dello spazio ne corrisponde una nel dominio delle frequenze espressa come somma pesata di armoniche (analisi e sintesi), a una funzione campionata con passo Δx corrisponde nel dominio trasformato uno spettro periodico, con periodo in funzione del passo di campionamento.

$$F(u) = \sum_{k=-\infty}^{\infty} f(k\Delta x)e^{-j2\pi uk\Delta x}$$

La frequenza di campionamento minima, quindi, garantisce un periodo sufficientemente grande per contenere l'intero spettro. Esso è simmetrico, quindi $-f_{max} = f_{max}$ come visto nel dominio della trasformata di Fourier.

Se il passo di campionamento è troppo alto ($\Delta x' = 1/M > \Delta x = 1/N$), le repliche dello spettro si sovrappongono (periodo M troppo breve) e non è possibile risalire alla funzione originale.
Questo succede perché le alte frequenze che non vengono considerate (a causa dell'alta variabilità del segnale, che non può essere interpretata in un periodo) ma si sovrappongono alle più basse, causando una errata rappresentazione.

Al contrario, una bassa frequenza di campionamento produce fenomeni di aliasing: la funzione viene approssimata male, come costante o una differente sinusoide.

Il teorema di Shannon afferma che, data la frequenza massima del segnale f_{MAX}, la frequenza massima di campionamento deve essere:

$$F_s = {}^1\!/_{\Delta} > 2f_{MAX}$$

Il periodo minimo di ripetizione dello spettro perché non vi sia sovrapposizione $N = 2f_{MAX}$. La f_{MAX} viene definita frequenza di Nyquist, e rappresenta la frequenza massima del segnale.

Lo spazio pertanto sarà $2f_{MAX}$, il che è il minimo valore per evitare le sovrapposizioni delle repliche. Nella formula viene utilizzata la disuguaglianza stretta per assicurare che non ci siano due frequenze nello stesso punto.

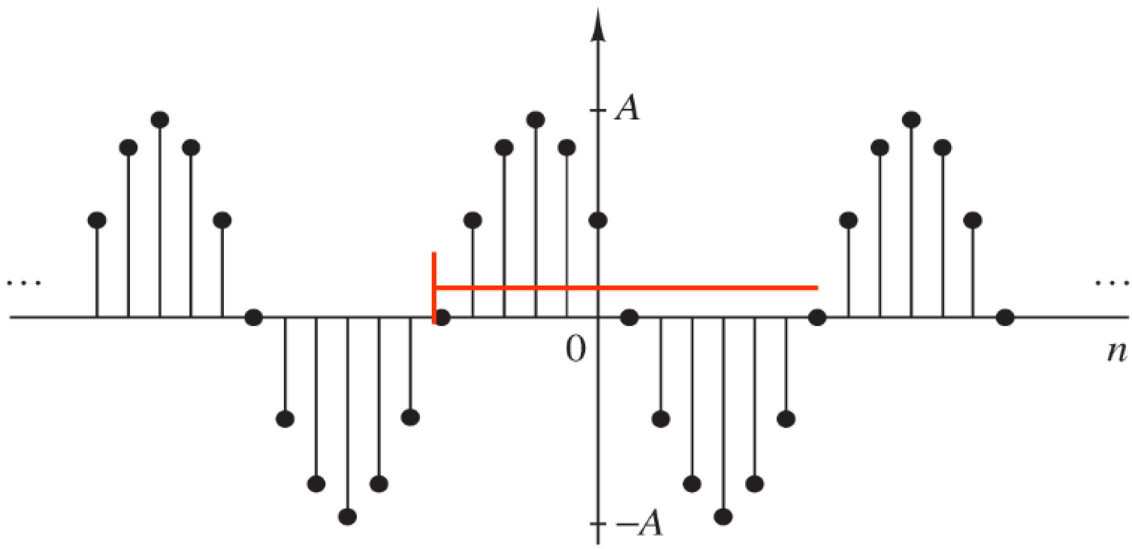
Per eliminare le frequenze alte e di conseguenza ridurre il passo di campionamento si ricorre al filtraggio (filtro anti-aliasing), che comunque non assicura il recupero della funzione originale ma evita le sovrapposizioni.

L'aliasing è appunto il fenomeno per cui il segnale originale non è ricostruibile dato che il campionamento è avvenuto con frequenza inferiore a quella di Nyquist, e il segnale risultante ha frequenza inferiore all'originale. Questo può gravemente compromettere la qualità di immagini e video: nella pratica i segnali non sono limitati in frequenze, ma nel tempo, quindi le repliche di uno spettro vanno separate attentamente tramite filtri passa-basso.

Quando il segnale è campionato, la variabile tempo/spazio non è più esplicita e la frequenza f_N rappresenta il numero di cicli per campione.

7.1 Esempio

$$x(n) = A\cos(\omega n + \theta)$$



$\omega = 2\pi f_N$, dove ω è la pulsazione e f_N la frequenza normalizzata, cioè il numero di cicli al secondo (continuo), con campioni al posto dei secondi nel mondo campionato.

Per capire la frequenza f_N del segnale rappresentato, bisogna guardare quanti cicli ci sono per campioni (o quanti campioni ci sono in un ciclo, cioè prima che il pattern si ripeta) e calcolare l'inverso. In questo caso un periodo è 12 campioni, quindi $f_N = 1/12$.

La frequenza al secondo (Hz) è ottenibile sapendo quanta distanza c'è tra un campione e l'altro, altrimenti per ogni rappresentazione ci sarebbero infinite sinusoidi.

La pulsazione ω è uguale a $2\pi f_N$, quindi $\frac{2\pi}{12} = \frac{\pi}{6}$.
Lo sfasamento con anticipo di 2 campioni si trova effettuando la proporzione: un periodo è 2π, se 12 campioni sono in2π allora $2\pi : 12 = \theta : 2 \rightarrow \theta = \frac{\pi}{3}$.

La frequenza normalizzata indica quanto bene si vedono i campioni e la loro variazione in un periodo, per poi decidere la finestra di osservazione. La frequenza dipende dal numero di campioni al secondo (f_c): se la frequenza di campionamento è $f_c = 1$ campione/secondo, si ha che $f = 1/12$ ciclo/secondo, cioè $f = f_c \cdot f_N$.

Per ricavare la formula più facilmente si possono utilizzare le unità di misura. Un altro modo per arrivare allo stesso risultato è sostituire a $n\Delta t$ la t.

Con 2 campioni al secondo, cioè $f_c = 2$ e $\Delta t = 0.5 = {}^1\!/_2$, la frequenza f in cicli al secondo (Hz) è ${}^1\!/_2$:

$$f_N \cdot f_c = \frac{1}{12}\frac{\text{cicli}}{\text{campione}} \cdot \frac{2\ \text{campioni}}{\text{sec}} = \frac{1}{6}\frac{\text{cicli}}{\text{sec}}$$

$$x(n) = A\cos\left(2\pi\frac{1}{12}n + \frac{\pi}{3}\right) \implies x(t) = A\cos\left(2\pi\frac{1}{6}t + \frac{\pi}{3}\right) \qquad \Delta tn \implies t$$

7.2 Frequenze

Il range tra minimo e massimo di una frequenza di oscillazione è 1, tutto ciò che è al di fuori dell'intervallo in realtà si trova comunque normalizzato all'interno. Aumentando ω, infatti, dopo un giro (periodo 2π) si torna ad avere la stessa funzione.

Frequenza minima di oscillazione di una sinusoide a tempo discreto: $f_N = 0 \implies \omega = 0$
Frequenza massima di oscillazione di una sinusoide a tempo discreto: $f_N = \pm{}^1\!/_2 \implies \omega = \pm\pi$
Frequenza di oscillazione di una sinusoide generica: $f = \frac{1}{\text{n. campioni}} \implies \omega = 2\pi \cdot f_N$

I segnali sinusoidali a tempo discreto con pulsazioni separate da multipli di 2π sono identici:
$\omega + k2\pi = \omega$ con $k \in \mathbb{Z}$
$x(n) = \cos(\omega n + \theta)$
$\cos((\omega + 2k\pi)n + \theta) = \cos(\omega n + \theta)$
$\cos(\omega n + kn2\pi + \theta) = \cos(\omega n + \theta)$

Per convenzione si utilizza la frequenza di Nyquist, $f_n = [-{}^1\!/_2, {}^1\!/_2]$, oppure intervalli di pulsazione $[-\pi, \pi]$ o $[0, 1] \to [0, 2\pi]$.

Il segnale sinusoidale è periodico solo se la sua frequenza normalizzata f è un numero razionale (rapporto tra due interi): $f = {}^k\!/_N$. I numeri razionali quindi impongono la periodicità.

8 Convoluzione

La convoluzione è un operatore matematico che descrive l'interazione tra il segnale e il filtro, e ne osserva il comportamento durante il movimento nello spazio. Si definisce nel dominio continuo fra due funzioni $g(x)$ e $f(x)$ tale che:

1. L'asse di rappresentazione di uno dei due assi è invertita, $g(t) \to g(-t)$;
2. Il segnale invertito viene fatto traslare tra ∞ e $-\infty$;
3. Per ogni traslazione si calcola il prodotto tra il segnale traslato e l'altro non traslato;
4. Si calcola l'area del prodotto, cioè la somma degli infiniti prodotti composti da una funzione che scorre e una statica.

La convoluzione è l'operatore con cui sono descritti i filtraggi lineari nel dominio spaziale, cioè l'applicazione di una funzione f a una funzione h chiamata filtro (filter kernel, descrive il sistema) per ogni valore di x. Gode della proprietà di linearità e commutativa.

$$g * f = \int_{x=-\infty}^{\infty} g(x-s)f(s)ds$$

Nel caso di funzioni discrete, essa è definita come una somma di prodotti tra gli elementi di f e i coefficienti di h, con funzioni che sono in realtà sequenze:

$$g(x) = \sum_m f(m)h(x-m)$$

Se i segnali non si sovrappongono, la somma dei loro prodotti è 0. Quando iniziano a sovrapporsi progressivamente, se sono positivi, i valori crescono fino a raggiungere un massimo, per poi diminuire.

La lunghezza finale della funzione ottenuta dipende dai due segnali che la compongono: assumendo che non ci siano valori nulli, con A punti di $f(m)$ e B di $h(-m)$ si ha $g(x)$ con $A+B-1$ punti.

Applicando la definizione di trasformata di Fourier, si può dimostrare il fondamentale teorema della convoluzione. La trasformata della convoluzione di due funzioni è il prodotto delle trasformate delle due funzioni:

$$G(u) = \mathcal{F}[g(x)] = \mathcal{F}[f(x) * h(x)] = F(u)H(u)$$