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.

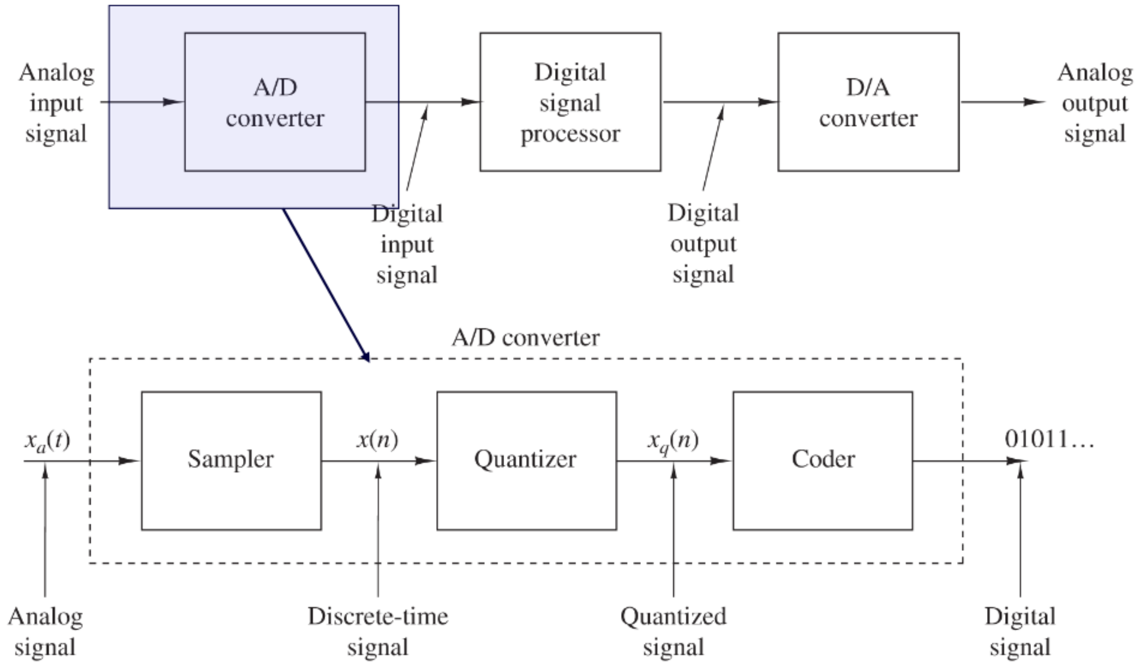
Le fasi della rappresentazione da analogica a digitale sono **campionamento**, **quantizzazione** e **codifica**.

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.

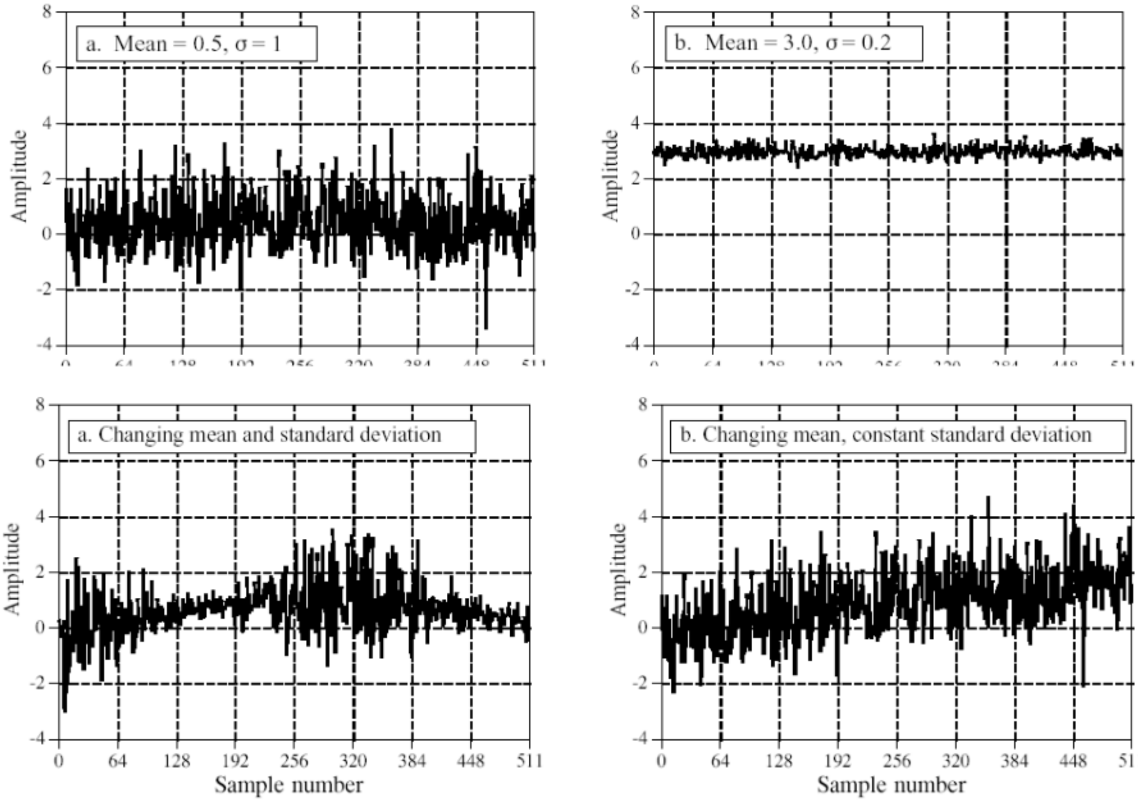
- Continuo:
  - $\mu = \lim_{T \to \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} x(t)dt$;
  - $\mu = \frac{1}{T_1 - T_0} \int_{T_0}^{T_1} x(t)dt$;
- Discreto:
  - $\mu = \frac{1}{N} \sum_{i=0}^{N-1} x_i$.

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.

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

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

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: $P_x = \lim_{T\to\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} |x(t)|^2 \, dt$

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:

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

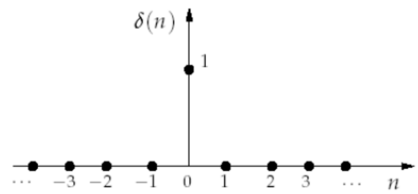
Viene introdotta per rappresentare fenomeni fisici di durata infinitesima (impulsi).

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:

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

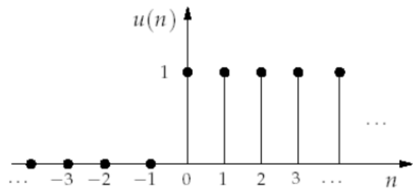


$x(n) \cdot \delta(n) = x(0)$

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

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

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:: $x(n) \ : \ n > 0$ (numeri positivi).
Sequenza anticausale: $x(n) \ : \ n < 0$ (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.

diagonale = funzione rampa, fattore moltiplicativo n

6 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.