

Digital Signals and Images Management

Sommario

Il ricevimento avviene su appuntamento (per mail) in U14/1011. Il corso comprende due moduli: il primo di lezioni frontali da 4CFU e un laboratorio di 2CFU (organizzato da M. Buzzelli) con Python (Keras o PyThorch). Durante il corso sono presentati degli esercizi (da presentare entro la fine del corso) la cui valutazione peserà per il 40% del totale; per il restante 60% la valutazione è basata sulla discussione di un progetto (massimo 3 persone, 4 in via eccezionale). Il progetto è composto in tre parti: *processing* di segnali monodimensionali (audio), bidimensionali (video) e *retrieval*. Non ci sono vincoli dati dalla qualità dei dati (presenza di rumore di fondo). Oltre alle tracce proposte, è possibile proporre tracce alternative, a patto che contengano almeno due delle tre parti e che la difficoltà sia commisurata al numero di crediti. Il martedì sarà fatta lezione frontale, il giovedì laboratorio.

Indice

I	Introduzione	2
1	Digitalizzazione	2
2	Analisi di Fourier	3

Parte I

Introduzione

Esistono quattro macro-categorie di segnali, riconoscibili da dominio e codominio:

- segnale analogico: $\mathbb{R} \mapsto \mathbb{R}$;
- segnale campionato: $\mathbb{K} \mapsto \mathbb{R}$ (si ha informazione solo in determinati momenti);
- segnale quantizzato: $\mathbb{R} \mapsto \mathbb{K}$;
- segnale digitale: $\mathbb{K} \mapsto \mathbb{K}$.

Generalmente i segnali analizzati sono gli ultimi (file digitali), tipicamente generati tramite *digitalizzazione* di segnali analogici. Il processo di digitalizzazione avviene tramite una *codifica*, limitata dalla memoria del dispositivo digitale e dalla sua velocità (per effettuare il campionamento, che non può avere una latenza inferiore ad una certa soglia, perché altrimenti non potrebbe essere gestito). Il file così ottenuto può quindi essere compresso, processato o trasmesso.

Il segnale analogico quindi è dapprima convertito in formato digitale, poi processato e spesso codificato nuovamente in formato analogico (per poter essere fruito).

1 Digitalizzazione

Il campionamento del segnale nel processo di digitalizzazione è effettuato con cadenza regolare (*cadenza* o *frequenza di campionamento*). In base al numero di bit richiesti per archiviare l'informazione in presenza di ogni punto, il codominio del segnale analogico è discretizzato (in 2^{bit} valori) in modo tale da approssimarlo con la minor perdita di informazione possibile. La differenza tra il segnale analogico e la sua rappresentazione digitale è definita *rumore di quantizzazione*: una sorta di errore nella funzione a gradini di approssimazione. Il rumore diminuisce all'aumentare dei bit richiesti per la codifica del singolo campione.

Se la latenza del campionamento è breve (il passo è stretto) la qualità del segnale digitale è alta; tuttavia lo spazio di archiviazione richiesto aumenta; tuttavia con una frequenza di campionamento eccessivamente bassa, si ha una perdita di informazione (nel caso di immagini si generano degli artefatti). Si cerca quindi di effettuare il campionamento con una latenza tale da ottimizzare lo spazio richiesto e la qualità del prodotto finito. Con una frequenza del segnale massima nota, la frequenza minima di campionamento richiesta per non avere perdita di qualità ne è il doppio (*frequenza di Nyquist*).

Il campionamento avviene in base alla quantità di dimensioni richieste dal segnale: può avere una o più dimensioni. I segnali monodimensionali, generalmente audio, si presentano come funzioni del tempo; possono essere anche *multicanale* (come un elettrocardiogramma registrato su più parti diverse del corpo). I segnali bidimensionali (come le fotografie) invece sono campionate in base alla posizione del punto (pixel) rispetto agli altri; se a ogni punto è associato più di un valore (RGB) diventa un segnale *multicanale*. Spesso i segnali sono rappresentati però non in funzione del tempo o dello spazio ma in *frequenze*.

2 Riduzione delle dimensioni

L'analisi di Fourier può essere utile in vari modi: primo tra tutti dimostra che un segnale può essere considerato come la somma di armoniche elementari di *frequenza*, *ampiezza* e *fase* noti. Il numero delle dimensioni quindi scende notevolmente: anziché avere una dimensione per ogni osservazione, si inseriscono solamente le caratteristiche delle armoniche più importanti. Queste sono (per ogni armonica), l'*ampiezza* (l'altezza massima che il segnale assume, ovvero il *range* del codominio) e il *periodo* (la lunghezza in cui il segnale si ripete, o la distanza tra due picchi successivi).

Alternativamente si usano valori indice, come l'energia E di un segnale $f(t)$ che si estende all'infinito, definita come:

$$E_f = \int_{-\infty}^{+\infty} f^2(t) dt$$

(nel caso di segnali digitali l'integrale è sostituito dalla sommatoria). Se $E_f < \infty$ il segnale è un *segnale di energia*. La *potenza* P di un segnale $f(t)$ è definita invece come:

$$P_f = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} (|f(t)|)^2 dt$$

i cui primi due momenti sono:

$$\begin{aligned} \mu &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f(t) dt \\ \hat{\sigma}^2 &= \frac{1}{N-1} \sum_{i=0}^{N-1} (f(t) - \mu)^2 \end{aligned}$$

Questi valori possono essere usati al posto del segnale grezzo per poter allenare un classificatore.