

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	Riduzione delle dimensioni	3
3	Trasformata di Fourier	3
3.1	Trasformata discreta	4
3.2	Analisi di Fourier	5
3.3	Esempio: rimozione di frequenze	5
4	Operatori puntuali	5
5	Spazio dei colori	6

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:

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

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

3 Trasformata di Fourier

Data una funzione qualsiasi, questa può essere scomposta in armoniche fondamentali in modo tale da approssimarla con precisione arbitraria.

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

dove N è il periodo della funzione, $\frac{1}{N}$ è la *frequenza fondamentale* F_0 . Il calcolo dei coefficienti è effettuato come segue:

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

I coefficienti sono unicamente determinati: è impossibile scomporre la funzione in modo diverso (imponendo un limite a cui terminare la scomposizione).

La trasformata di Fourier permette di passare dal dominio temporale o spaziale al dominio delle frequenze, che facilita alcune operazioni necessarie all'analisi. Una sinusoide di ampiezza A e periodo T è rappresentata quindi come un picco di ampiezza A nel punto $F = \frac{1}{T}$.

$$F(u) = \int_{-\infty}^{+\infty} f(x) \cdot e^{-i2\pi ux} dx$$

La trasformata di Fourier è quindi una funzione con valori nel campo complesso \mathbb{C} . Con l'operazione di *antitrasformata* la funzione è riportata dal dominio delle frequenze al dominio originario senza perdita d'informazione (a livello teorico):

$$f(x) = \int_{-\infty}^{+\infty} F(u) \cdot e^{i2\pi ux} du$$

che restituisce il valore della funzione nel punto specifico.

Con la trasformazione di Fourier è quindi possibile portare i dati in un dominio differente che semplifica alcune operazioni e riportarlo indietro al dominio originario.

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

Il modulo della trasformata $|F(u)|$ è definito come:

$$|F(u)| = \sqrt{\Re(u)^2 + \Im(u)^2}$$

mentre la fase:

$$\phi(u) = \tan^{-1} \frac{\Im(u)}{\Re(u)}$$

Per il caso bidimensionale l'equazione della trasformata assume un'altra forma:

$$F(u, v) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) \cdot e^{-i2\pi(ux+vy)} dx dy$$

$$f(u, v) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} F(x, y) \cdot e^{i2\pi(ux+vy)} dx dy$$

che rimane molto simile a parte per il numero di assi.

3.1 Trasformata discreta

Operando con segnali digitali (che quindi assumono valori discretizzati), l'integrale è sostituito dalla sommatoria:

$$F(u) = \frac{1}{N} \sum_{j=0}^{N-1} f(j) \cdot e^{-i2\pi u \frac{1}{N} j}$$

3.2 Analisi di Fourier

La combinazione delle armoniche (tramite mera sommatoria) permette di stimare funzioni con precisione arbitraria. Ribaltando le armoniche nello spazio delle frequenze, si ottiene un fascio di linee parallele all'asse dell'ampiezza con altezza pari al contributo nel generare l'approssimazione; tuttavia un segnale reale non è così regolare da avere solamente dei picchi ma anche delle interferenze in prossimità. L'altezza delle rette è pari alla metà dell'ampiezza A dell'armonica corrispondente, e lo spazio è simmetrico rispetto all'asse delle ampiezze.

3.3 Esempio: rimozione di frequenze

Si intende filtrare il suono di un'onda eliminando le frequenze sopra ad una certa soglia, conservando solamente i bassi.

Effettuare l'operazione nel dominio del tempo è complicato: bisogna programmare un filtro di lunghezza arbitraria l che, partendo dal primo campione, moltiplichi gli l campioni successivi per dei pesi (la cui stima è oggetto di ricerca) per ottenere una proiezione del primo campione. Dunque si sposta il filtro alla posizione successiva e si ripete l'operazione fino a raggiungere la fine del file.

Effettuare l'operazione nello spazio delle frequenze è nettamente più semplice: è necessario solamente programmare un filtro *passa-banda* (ovvero una funzione che vale 0 sulle frequenze da eliminare e 1 su quelle da conservare) e moltiplicarlo per la trasformata $F(u)$, e dunque effettuare l'anti-trasformata per poter fruire nuovamente del file.

Discorso analogo può essere fatto selezionando arbitrariamente qualsiasi intervallo di frequenze. Spesso al posto di usare filtri ideali (ovvero che assumono solamente valore 0 o 1) si usano filtri più morbidi (come i filtri gaussiani) per evitare la generazione di artefatti: questi assumono valori compresi nell'intervallo e riducono tutte le frequenze in modo progressivo.

Per effettuare la stessa cosa con un'immagine, si considera la trasformata bidimensionale del segnale: di questo si considera solamente il modulo (si conserva però la fase per poter tornare all'immagine originale). Più la rappresentazione del modulo è regolare, più l'immagine sarà ordinata. Per costruire un filtro basta considerare una circonferenza di raggio arbitrario per considerare solamente determinate frequenze. Considerando solamente le frequenze alte sono conservati i bordi dell'immagine, con frequenze basse le informazioni sul contenuto (si ottiene una sfocatura dell'immagine).

4 Operatori puntuali

Contrast Stretching Serve per aumentare il contrasto compreso tra due valori di grigio in un'immagine. Con una funzione particolarmente angolata (caso estremo) tutti i valori sopra a una certa soglia sono mappati col nero ottico e sotto col bianco ottico: si ha quindi una binarizzazione.

$$s = T(r) = \frac{1}{1 + \left(\frac{m}{r}\right)^E}$$

Tipico caso è l'aumento nella ristrutturazione di negativi fotografici e pellicole.

Smoothing Sono filtri che fanno perdere nitidezza all'immagine rendendo i contorni meno decisi. Sono usate per appiattare transazioni rapide che possono esserci tra un pixel e il successivo, rimuovere dettagli (*blurring*) o collegare linee vicine che non sono collegate o ancora ridurre il rumore. I filtri possono essere costruiti o riassegnando il valore del pixel in base a una maschera di pesi o sfruttando momenti del campione estratto dalla maschera. I filtri con pesi sono usati con rumore gaussiano, mentre i filtri mediani sono molto efficaci nei casi di *salt and pepper* (picchi molto scuri o chiari).

Sharpening Sono filtri usati per migliorare la nitidezza dell'immagine, tramite un processo di derivazione (che può anche essere calcolata come la differenza tra un pixel e il successivo). Avendo a che fare con strutture bi-dimensionali, si sfruttano maschere bi-dimensionali, i cui pesi hanno somma nulla; in base a come sono distribuiti i pesi, sono evidenziati i dettagli orizzontali o verticali (ma non entrambi). Un filtro trasposto passa da evidenziare i dettagli da una direzione all'altra. Per evidenziare tutti i dettagli, si combinano entrambe le immagini combinando il gradiente bidimensionale.

$$\begin{aligned}\partial^2 f &= \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \\ \frac{\partial^2 f}{\partial x^2} &= f(x+1, y) + f(x-1, y) - 2f(x, y) \\ \frac{\partial^2 f}{\partial y^2} &= f(x, y+1) + f(x, y-1) - 2f(x, y)\end{aligned}$$

Andando a sommare la derivata seconda all'immagine originale, si ottiene un'immagine i cui dettagli sono molto più marcati rispetto all'originale.

Aliasing Riducendo la risoluzione di un'immagine, si eliminano frequenze che prima erano visibili: prima di effettuare questa operazione è buona norma effettuare uno *smoothing* dell'immagine di partenza per eliminare le frequenze che comunque verrebbero tagliate.

5 Spazio dei colori

Il modello classico di codifica dei colori prevede l'uso di tuple (di 3 o 4 valori). Lo *spazio dei colori* presuppone inoltre una definizione su come interpretare le informazioni. Esistono più spazi colore diversi, scelti a seconda dell'uso che se ne deve fare.

RGB È uno spazio colore dipendente dal dispositivo: una volta rappresentati i colori “puri”, tutti i colori ottenibili abitano nel cubo così ottenuto. Il successo è dato dal fatto che sia uno standard, e quindi può essere usato su ogni dispositivo. L'informazione è codificata tramite tre numeri che definiscono le coordinate del colore nello spazio.

Spazi colore intuitivi Lo spazio è rappresentato come la distanza del colore rispetto ad un centro di simmetria radiale. La saturazione aumenta all'aumentare della distanza dal centro radiale; mentre l'intensità è data da quanto è scuro lo spazio. Esistono più tipologie

di spazi colore di questo tipo in base alla forma che può ottenere (cilindrica, cono o doppio cono), ma gli assi rappresentano sempre le medesime informazioni.

Molti dispositivi sono in grado di codificare i colori in modo sia lineare (in ingresso) che non lineare (in uscita): per poter ottenere risultati migliori si preferisce progettare dispositivi che codificano i colori in entrata in modo non lineare.

CYCbCr Questa codifica separa la luminosità Y dall'opponenza blu-giallo Cb e rosso-verde Cr .