

ALMA MATER STUDIORUM · UNIVERSITÀ DI  
BOLOGNA

---

SCUOLA DI SCIENZE  
Corso di Laurea in Matematica

**METODI MATEMATICI  
PER L'ANALISI  
DELL'ANDATURA**

**Relatore:**  
Chiar.mo Prof.  
Massimo Ferri

**Presentata da:**  
Giuseppe Murolo

Sessione unica  
a. a. 2017-2018



*“È difficile restare arrabbiati  
quando c'è tanta bellezza nel mondo.  
A volte è come se la vedessi tutta insieme, ed è troppa.  
Il cuore mi si riempie come un palloncino che sta per scoppiare.  
E poi mi ricordo ...di rilassarmi,  
e smetto di cercare di tenermela stretta.  
E dopo scorre attraverso me come pioggia,  
e io non posso provare altro che gratitudine.”  
(American Beauty, Sam Mendes, 1999)*

*Ai miei genitori,  
che sempre hanno  
creduto in me*



# Indice

<b>Introduzione</b>	<b>5</b>
<b>1 Prerequisiti matematici</b>	<b>9</b>
1.1 Prerequisiti per il capitolo 2 . . . . .	9
1.2 Prerequisiti per il capitolo 3 . . . . .	13
<b>2 Riconoscimento tramite la GEI</b>	<b>19</b>
2.1 Rappresentazione del cammino con la Gait Energy Image . . . . .	19
2.2 Uso della Gait Entropy Image . . . . .	21
2.3 Selezione delle caratteristiche a coppie . . . . .	22
2.4 Cenni su ulteriori metodi di riduzione della dimensione del problema . . . . .	23
2.5 Esperimenti . . . . .	25
<b>3 Riconoscimento tramite la persistenza</b>	<b>29</b>
3.1 Il complesso simpliciale $\partial K(I)$ . . . . .	29
3.2 Le filtrazioni sul complesso simpliciale $\partial K(I)$ . . . . .	31
3.3 Omologia persistente per la firma topologica . . . . .	32
3.4 Stabilità della firma topologica . . . . .	34
3.5 Esperimenti . . . . .	35
<b>Conclusioni</b>	<b>37</b>
<b>Bibliografia</b>	<b>39</b>

Lista delle figure

43

# Introduzione

Il riconoscimento di una persona è una componente essenziale nella sicurezza o, più nello specifico, nella videosorveglianza. In particolare sono pochissime le tecniche biometriche utilizzabili in una situazione in cui c'è bisogno di un riconoscimento a distanza, o in cui serve poter non interferire con l'attività del soggetto da identificare.

L'analisi dell'andatura, della deambulazione, è tra queste ed ha un importante potenziale: le informazioni utili possono essere estratte da molti punti di vista diversi e, soprattutto, a distanze maggiori rispetto ad altri approcci biometrici, come l'analisi dell'iride o il riconoscimento facciale.

In più, a differenza del riconoscimento tramite impronta digitale, non c'è bisogno di un'interazione dell'individuo con l'ambiente.

Purtroppo però le caratteristiche che si estraggono dalla sequenza video di un'andatura, o l'andatura stessa, sono sensibili a molte condizioni, che includono ad esempio l'entità della superficie sulla quale si cammina, l'indossare un particolare abbigliamento (un cappotto lungo, un ampio vestito da donna, una certa calzatura...) o il trasportare oggetti (un pacco, uno zaino, una valigia, una borsa...). In generale quindi sarà relativamente semplice confrontare andature di individui sotto le stesse condizioni, mentre lavorare con un set di andature eterogenee si dimostrerà più complicato.

Nella letteratura recente, gli approcci al problema rientrano principalmente in due ampie categorie: i metodi model-based (per esempio: [1],[24],[3],[17],[23]) e i metodi appearance-based (per esempio: [6], [9], [11], [2]).

I primi mirano a costruire un modello tridimensionale rappresentante l'in-

tero corpo umano dell'individuo, a partire dalle immagini della sua andatura. Sono in genere più complessi e computazionalmente costosi dei metodi appearance-based. Inoltre, per un buon funzionamento, richiedono immagini di buona qualità e libere da rumore, il che potrebbe non essere disponibile in uno scenario di applicazione reale, come la sorveglianza in luoghi pubblici. Gli approcci appearance-based utilizzano particolari informazioni estratte direttamente dalle sagome dell'andatura, senza considerare esplicitamente la profondità della struttura sottostante le silhouette bidimensionali. Spesso questi approcci, rispetto ai primi, mostrano una certa stabilità e risultano meno sensibili al rumore. Le recenti tendenze nella ricerca sul riconoscimento del cammino sembrano quindi favorire queste strategie.

In questo elaborato verranno analizzati, rispettivamente nei Capitoli 2 e 3, due differenti metodi, entrambi appearance-based, tratti rispettivamente da [2] e [11].

Il primo è finalizzato ad ottenere una rappresentazione dell'andatura mediante un'immagine in scala di grigi, studiando la frequenza delle posizioni delle parti del corpo in movimento. Alla base di quest'approccio c'è la cosiddetta GEI (Gait Entropy Image), che ha recentemente guadagnato molta popolarità, trovando impiego in molti algoritmi utilizzati per studiare il problema in questione. Generalmente questa strategia è caratterizzata da un utilizzo di un numero di sagome non troppo elevato (mai più di un ciclo completo di andatura); inoltre il metodo non va a considerare l'ordine temporale di apparizione delle sagome nella sequenza video da cui sono estratte: sono informazioni disperse, sebbene potenzialmente utili.

Al contrario, il secondo approccio è capace di acquisire tali dati. L'idea di fondo di questo sarà costruire un complesso simpliciale  $\partial K(I)$ , ottenuto dalla pila delle sagome bidimensionali, e studiarne le proprietà topologiche attraverso la persistenza. Inoltre viene studiata solo la parte inferiore della sagoma dell'individuo (i piedi e l'ultima parte delle gambe): in questo modo, nel momento in cui si lavora con andature eterogenee (dove l'insieme delle sagome che si studia è composto da individui sotto diverse condizioni che

influenzano l'andatura o la sagoma), l'efficacia di questo metodo aumenterà rispetto al primo proposto, assieme alla stabilità e alla robustezza al rumore. Inoltre con questo tipo di analisi siamo liberi di poter considerare anche più di un ciclo completo di andatura per volta.

Nel prossimo Capitolo, il n. 1, verranno introdotti gli strumenti matematici utili all'esposizione dei due metodi.

L'ultimo paragrafo dei Capitoli 2 e 3 è dedicato all'analisi degli esperimenti condotti per ogni metodo.

Le conclusioni sono finalmente presentate nell'ultimo capitolo.



# Capitolo 1

## Prerequisiti matematici

### 1.1 Prerequisiti per il capitolo 2

**Definizione 1** (Spazio misurabile). Uno *spazio misurabile* è una coppia  $(\Omega, \mathcal{F})$  dove:

- $\Omega$  è un insieme non vuoto;
- $\mathcal{F}$  è una  $\sigma$ -algebra su  $\Omega$ , ossia  $\mathcal{F}$  è una famiglia non vuota di sottoinsiemi di  $\Omega$  tali che:
  - se  $A \in \mathcal{F} \Rightarrow A^c := \Omega \setminus A \in \mathcal{F}$ ;
  - l'unione numerabile di elementi di  $\mathcal{F}$  appartiene ad  $\mathcal{F}$ .

**Definizione 2** (Misura). Una *misura* sullo spazio misurabile  $(\Omega, \mathcal{F})$  è una funzione  $\mu : \mathcal{F} \rightarrow [0, +\infty]$  tale che:

- $\mu(\emptyset) = 0$ ;
- $\mu$  è  $\sigma$ -additiva su  $\mathcal{F}$ , ossia per ogni successione  $(A_n)_{n \in \mathbb{N}}$  di elementi disgiunti di  $\mathcal{F}$  vale:

$$\mu\left(\bigcup_{n=1}^{+\infty} A_n\right) = \sum_{n=1}^{+\infty} \mu(A_n) \quad (1.1)$$

Diremo quindi che  $(\Omega, \mathcal{F}, \mu)$  è uno *spazio di misura* con misura  $\mu$ .

**Definizione 3** (Misura di probabilità). Una misura  $P$  su  $(\Omega, \mathcal{F})$  tale che  $P(\Omega) = 1$ , è detta *misura di probabilità* o semplicemente *probabilità*.

Diremo quindi che  $(\Omega, \mathcal{F}, P)$  è uno *spazio di probabilità*. In uno spazio di probabilità ogni elemento  $\omega \in \Omega$  è detto *evento elementare*; ogni  $A \in \mathcal{F}$  è chiamato *evento* e il valore  $P(A)$  è detto *probabilità di A*.

*Osservazione 4* (Probabilità complementare). Sia  $A$  un evento in uno spazio di probabilità  $(\Omega, \mathcal{F}, P)$ ; vale che  $P(A^c) = 1 - P(A)$ .

Ciò segue direttamente dal fatto che  $P(\Omega) = 1$ ,  $P(\emptyset) = 0$  e che si può usare l'equazione 1.1 prendendo ad esempio  $A_1 = A, A_2 = A^c, A_i = \emptyset \forall i \neq 1, 2$ , essendo  $A \cap A^c = \emptyset$ .

**Definizione 5** (Variabile aleatoria). Una *variabile aleatoria* a valori reali sullo spazio di probabilità  $(\Omega, \mathcal{F}, P)$  è una funzione

$$X : \Omega \rightarrow \mathbb{R}$$

tale che

$$X^{-1}(\mathcal{B}) \subseteq \mathcal{F}$$

dove con  $\mathcal{B}$  indichiamo i Boreliani di  $\mathbb{R}$ : la più piccola  $\sigma$ -algebra su di  $\mathbb{R}$  dotato della topologia euclidea, che sia compatibile con la topologia stessa, ovvero che contenga tutti gli aperti della topologia.

Diremo che la variabile aleatoria  $X$  è *discreta* se l'insieme dei suoi possibili valori è finito o numerabile.

Partendo dalla definizione data, si può generalizzare al caso di *variabili aleatorie multidimensionali* (a valori in  $\mathbb{R}^d$ ): sono vettori colonna le cui componenti sono variabili aleatorie a valori reali.

**Definizione 6** (Valore atteso). Data una variabile aleatoria  $X$  a valori reali su uno spazio di probabilità  $(\Omega, \mathcal{F}, P)$ , il suo *valore atteso* o *media* è l'integrale di  $X$  rispetto alla misura di probabilità  $P$ :

$$E[X] := \int_{\Omega} X(\omega) dP(\omega). \quad (1.2)$$

Quindi, nel caso di  $X$  variabile aleatoria discreta, si ha:

$$E[X] = \sum_{i=1}^{+\infty} x_i p_i \quad (1.3)$$

dove  $p_i = P(\{\omega \in \Omega : X(\omega) = x_i\}) \in [0, 1]$ .

Il valore atteso formalizza l'idea euristica di valore medio di un fenomeno aleatorio.

**Definizione 7** (Varianza). La *varianza* della variabile aleatoria  $X$  è definita come il valore atteso del quadrato della variabile aleatoria  $X - E[X]$ :

$$\sigma_X^2 := E[(X - E[X])^2]. \quad (1.4)$$

La varianza fornisce una misura della variabilità dei valori assunti dalla variabile aleatoria stessa; nello specifico, la misura di quanto essi si discostino quadraticamente dal valore atteso.

**Definizione 8** (Covarianza, Matrice di covarianza). La *covarianza* di due variabili aleatorie  $X$  e  $Y$  è il valore atteso dei prodotti delle loro distanze dalla media:

$$Cov(X, Y) := E[(X - E[X])(Y - E[Y])]. \quad (1.5)$$

La covarianza di due variabili aleatorie è un numero che fornisce una misura di quanto le due varino assieme, ovvero della loro dipendenza.

La *matrice delle covarianze* è una generalizzazione della covarianza al caso di dimensione maggiore di due. Essa è una matrice  $N \times N$  che rappresenta la variazione di ogni variabile aleatoria  $X_1, \dots, X_N$  rispetto alle altre (inclusa se stessa). In particolare l'elemento della matrice di posizione  $(i, j)$  esprime la covarianza tra le variabili aleatorie  $X_i$  e  $X_j$ . Quindi sulla diagonale, nel posto  $(i, i)$ , vi è la varianza di  $X_i$ :  $\sigma_{X_i}^2$ .

Quella delle covarianze è una matrice simmetrica.

D'ora in poi chiameremo *sorgente* una variabile aleatoria discreta su  $(\Omega, \mathcal{F}, P)$  a valori reali. Diremo che  $x$  è un *messaggio emesso dalla sorgente*  $X$  se  $x$  è tra le immagini di  $X$ . Con un piccolo abuso di notazione, diremo che  $x$

è un evento, riferendoci all'insieme  $\{\omega \in \Omega : X(\omega) = x\} \in \mathcal{F}$  che è un evento.

**Definizione 9** (Informazione contenuta in un messaggio). L'*autoinformazione* di  $x$  o *informazione contenuta nel messaggio*  $x$ , emesso da una sorgente  $X$ , si calcola come:

$$I(x) := -\log_b p(x) \quad (1.6)$$

dove  $p(x)$  è la probabilità che l'evento  $x$  accada, ossia  $p(x) = P(\{\omega \in \Omega : X(\omega) = x\}) = P(\{X = x\})$ , e  $b$  è la base di numerazione scelta; come si vedrà più avanti, noi useremo  $b = 2$ , con riferimento al sistema binario (in modo da misurare l'informazione in bit).

$I(x)$  esprime la quantità di informazione contenuta nell'evento  $x$  e con "informazione" stiamo intendendo semplicemente una misura della libertà di scelta di cui si dispone nello scegliere un messaggio dall'insieme di quelli disponibili. L'autoinformazione è definita in questo modo affinché:

- sia una grandezza non negativa:  $I \geq 0$ ;
- sia tanto più grande quanto più rara è l'occorrenza del messaggio:  $I$  dipende da  $p$ , ossia  $I = I(p)$ , ed è decrescente in  $p$ ;
- se  $x$  è un evento *certo* (ossia  $p(x) = 1$ ) l'informazione contenuta in  $x$ ,  $I(x)$ , è nulla.

**Definizione 10** (Entropia di Shannon). Considerando l'informazione contenuta in un messaggio come una variabile aleatoria, possiamo definire l'*entropia di Shannon* o *entropia di una sorgente*  $X$  come il valore atteso dell'autoinformazione, ovvero l'informazione media contenuta in ogni messaggio emesso:

$$H[X] := E[I[X]] = E[-\log_b p(X)]. \quad (1.7)$$

Avendo per sorgente  $X$  una variabile aleatoria discreta, questo valore atteso si riduce ad una media aritmetica dell'informazione di ogni messaggio  $x_i$  pesata

con la propria probabilità  $p(x_i)$ :

$$H[X] = - \sum_{i=1}^{+\infty} p(x_i) \log_b p(x_i). \quad (1.8)$$

L'entropia di Shannon esprime l'informazione media contenuta in tutti i valori possibili di  $X$ .

In statistica l'entropia di Shannon è utilizzata come indice di eterogeneità di un insieme di dati; si può infatti dimostrare che l'entropia è massima quando tutti i valori  $x_i$  ( $i = 1, 2, \dots, n$ ) che può assumere la  $X$ , sono equiprobabili:  $p(x_i) = \frac{1}{n} \forall i = 1, 2, \dots, n$ .

## 1.2 Prerequisiti per il capitolo 3

Le Definizioni 15,16, parte della Definizione 14 e parte della Definizione 17 sono tratte da [5], così come l'Esempio 18.

La Definizione 20, parte della Definizione 14, parte della Definizione 17 e l'Esempio 19 sono tratte da [11].

**Definizione 11** (Simpleso, Facce di un simpleso). Dati, in uno spazio euclideo  $\mathbb{E}^m$ ,  $n + 1$  punti in posizione generale  $x_1, \dots, x_{n+1}$  (ossia tali che i vettori  $x_2 - x_1, \dots, x_{n+1} - x_1$  siano linearmente indipendenti) si dirà *simpleso di dimensione  $n$*  o  *$n$ -simpleso* la chiusura convessa di  $\{x_1, \dots, x_{n+1}\}$ , cioè:

$$s^n = \left\{ x \in \mathbb{E}^n \left| x = \sum_{i=1}^{n+1} \lambda_i x_i, \text{ con } \lambda_i \in \mathbb{R}, \lambda_i \geq 0, \text{ t.c. } \sum_{i=1}^{n+1} \lambda_i = 1 \right. \right\} \quad (1.9)$$

Se  $x_1, \dots, x_{n+1}$  sono in posizione generale, anche  $k + 1$  di questi (con  $k$  arbitrario,  $k \leq n$ ) sono in posizione generale; il simpleso  $k$ -dimensionale da essi generato è chiamato *faccia  $k$ -dimensionale*.

**Definizione 12** (Complesso simpliciale). Un *complesso simpliciale*  $K$  è un insieme di simplessi di uno stesso spazio  $\mathbb{E}^m$ , tale che:

- ogni faccia di un simpleso è un elemento di  $K$ ;

- l'intersezione di due semplici è vuota o è una faccia di entrambi;
- l'insieme  $K$  è finito.

**Definizione 13** (Complesso cubico). Un insieme  $X \subseteq \mathbb{R}^d$  si dice *complesso cubico* se può essere scritto come unione di *cubi elementari*, dove un cubo elementare è un prodotto finito del tipo  $Q = I_1 \times \dots \times I_d \subset \mathbb{R}^d$ , dove gli  $I_k$  sono intervalli di  $\mathbb{R}$  del tipo  $I_k = [\lambda_k, \lambda_k + 1]$  o  $I_k = [\lambda_k, \lambda_k]$ , per qualche  $\lambda_k \in \mathbb{Z}$ .

**Definizione 14** (Filtrazione, Funzione filtrante, Insiemi di Sottolivello). Una *filtrazione*  $F$  di un complesso simpliciale  $K$  è un ordine parziale sul complesso dettato dai valori di una funzione continua  $f : K \rightarrow \mathbb{R}$  detta *funzione filtrante*, tale che se un semplice  $s \in K$  è una faccia di un altro semplice  $s' \in K$ , allora  $f(s) \leq f(s')$  (ossia  $s$  appare prima o nello stesso momento di  $s'$  nella filtrazione indotta da  $f$ ).

Dati un complesso simpliciale  $K$ , una funzione filtrante  $f$  e uno scalare  $u$ , l'*insieme di sottolivello* relativo a  $u$  è l'insieme  $K_u := \{s \in K \mid f(s) \leq u\}$ .

Diamo ora una breve introduzione tecnica dell' omologia persistente.

**Definizione 15** (Operatore di bordo, Cicli, Bordi, Classi di omologia). Sia  $K$  un complesso simpliciale. Una  $p$ -*catena* su  $K$  è una combinazione lineare a coefficienti in  $\mathbb{Z}_2$  di  $p$ -simplessi. Le  $p$ -catene formano uno  $\mathbb{Z}_2$ -spazio vettoriale, che chiameremo  $C_p(K)$ .

Definiamo ora una trasformazione lineare  $\partial_p : C_p \rightarrow C_{p-1}$ , chiamata *operatore di bordo  $p$ -esimo*,  $\forall p \in \mathbb{Z}$ . Essendo una trasformazione lineare tra spazi vettoriali possiamo definirla sui generatori (i  $p$ -simplessi) e poi estenderla per linearità a tutto lo spazio. Indichiamo con  $s = [x_1, \dots, x_{p+1}]$  il semplice generato dai vertici in  $\{x_1, \dots, x_{p+1}\}$ , e con  $[x_1, \dots, \hat{x}_j, \dots, x_{p+1}]$  la sua faccia generata dai vertici in  $\{x_1, \dots, x_{p+1}\} \setminus \{x_j\}$  ( $j = 1, \dots, p+1$ ). Allora definiamo:

$$\partial_p(s) := \sum_{j=1}^{p+1} [x_1, \dots, \hat{x}_j, \dots, x_{p+1}] \quad (1.10)$$

È possibile provare che  $\partial_p \partial_{p+1} \equiv 0$ ;  $B_p := \text{Im}(\partial_{p+1})$  è dunque contenuto in  $Z_p := \text{Ker}(\partial_p)$ . Chiameremo gli elementi di  $B_p$   $p$ -bordi e gli elementi di  $Z_p$   $p$ -cicli.

Lo spazio vettoriale dell'omologia  $p$ -esima è definito come il quoziente  $H_p(K) := Z_p/B_p$ . I suoi elementi non nulli, detti  $p$ -classi di omologia, sono rappresentati da  $p$ -cicli che non sono  $p$ -bordi. Diremo che due  $p$ -cicli sono omologhi se la loro differenza è un  $p$ -bordo (i.e. sono equivalenti nello spazio vettoriale quoziente della  $p$ -omologia: appartengono alla stessa classe di omologia).

**Definizione 16** (Numeri di Betti). Il  $p$ -esimo numero di Betti relativo al complesso simpliciale  $K$ ,  $\beta_p(K)$ , è la dimensione dello spazio vettoriale dell'omologia  $p$ -esima  $H_p(K)$ ; è quindi il numero di generatori indipendenti di questo spazio vettoriale.

Intuitivamente  $\beta_0(K)$  conta il numero di componenti connesse di  $K$ , mentre  $\forall p > 0$ ,  $\beta_p(K)$  costituisce il numero di buchi o cavità  $p$ -dimensionali presenti in  $K$ .

Per ogni  $u, v \in \mathbb{R}$ ,  $u < v$ , la mappa di inclusione  $i^{u,v} : K_u \hookrightarrow K_v$  è continua ed induce, per ogni  $p$ , una trasformazione lineare  $i_*^{u,v} : H_p(K_u) \rightarrow H_p(K_v)$ . La funzione  $p$ -esimo numero di persistenza di Betti (funzione  $p$ -PBN) associa alla coppia  $(u, v)$  il numero  $\dim(\text{Im}(i_*^{u,v}))$ , ossia il numero di  $p$ -cicli di  $H_p(K_u)$  che 'sopravvivono' in  $H_p(K_v)$ .

**Definizione 17** (Diagramma di persistenza, Barcode di persistenza). Le funzioni  $p$ -PBN, che si presentano come funzioni a gradini, sono interamente determinate dalla posizione di alcuni punti e linee di discontinuità, che chiameremo *cornerpoints* e *cornerlines* (dette anche *cornerpoints all'infinito*, perchè rappresentano quelle classi che persistono fino all'infinito). Le coordinate  $(u, v)$  di un cornerpoint rappresentano rispettivamente i livelli di *nascita* e di *morte* di un generatore, una  $p$ -classe di omologia.

Intuitivamente, consideriamo una filtrazione  $F$  per il complesso simpliciale  $K$ , indotta dalla funzione filtrante  $f$ . Per semplificare la spiegazione, supponiamo che i simplessi della filtrazione siano totalmente ordinati (i.e. vengono aggiunti uno per volta); sia quindi  $F = (s_1, \dots, s_t)$ . Se il simpleso  $s_i$ , quando

viene aggiunto al sottoinsieme ordinato  $F_i = (s_1, \dots, s_i)$  in accordo con la funzione filtrante, completa un  $p$ -ciclo con gli elementi di  $F_i$ , una  $p$ -classe di omologia  $\alpha$  è nata al tempo  $f(s_i)$ ; invece, una  $(p-1)$ -classe di omologia  $\alpha$  è morta al tempo  $f(s_i)$ , se l'aggiunta di  $s_i$  a  $F_i$  rende un  $p$ -ciclo rappresentante di un  $p$ -bordo.

L'indice di persistenza, o semplicemente *persistenza*, di un cornerpoint è la differenza delle sue coordinate  $v - u$  ed esprime la durata della vita di una particolare classe di omologia.

Inserendo i cornerpoints e le cornerlines in un sistema di due assi cartesiani, otterremo una rappresentazione della persistenza che è chiamata  *$p$ -esimo diagramma di persistenza* relativo a una filtrazione  $F$  (indicato con  $p - dgm(F)$ ). Si veda l'Esempio 18.

Una rappresentazione equivalente, nota come  *$p$ -esimo barcode di persistenza* (relativo alla filtrazione  $F$  indotta da  $f$ ), è ottenuta inserendo in un sistema di assi cartesiani (dove sull'asse delle ascisse vi sono i tempi dettati dalla funzione filtrante  $f$ ) una barra orizzontale  $[\tau, \tau')$  se una classe di  $p$ -omologia è nata al tempo  $\tau$  ed è morta al tempo  $\tau'$ . Si veda l'Esempio 18 e l'Esempio 19.

*Esempio 18.* Si vogliono studiare le 0-classi di omologia (cioè le componenti connesse) dei grafemi rappresentanti le lettere 'M' e 'W' visibili nella Figura 1.1 (a) (dove le ordinate giocano il ruolo di funzione filtrante).

I due grafemi sono omeomorfi: uno è l'altro a meno di una rotazione di angolo  $\pi$ ; sono quindi topologicamente indistinguibili. Usando la topologia persistente invece riusciamo a caratterizzare i grafemi (tramite lo 0-esimo numero di persistenza di Betti, ovvero sia la persistenza delle componenti connesse) in modo da differenziarli. In particolare il grafico che descrive la funzione 0-PBN di 'M' visibile in Figura 1.1(b) è diverso da quello di 'W' in 1.1(c), lo 0-dgm( $F$ ) di 'M' in 1.1(d) è diverso da quello di 'W' in 1.1(e), il barcode 0-esimo di 'M' in 1.1(f) differisce da quello di 'W' in 1.1(g).

*Esempio 19.* Nella Figura 1.2 possiamo osservare in basso un esempio di barcode 0- e 1-esimo del complesso simpliciale  $K$  in alto. Le barre colo-

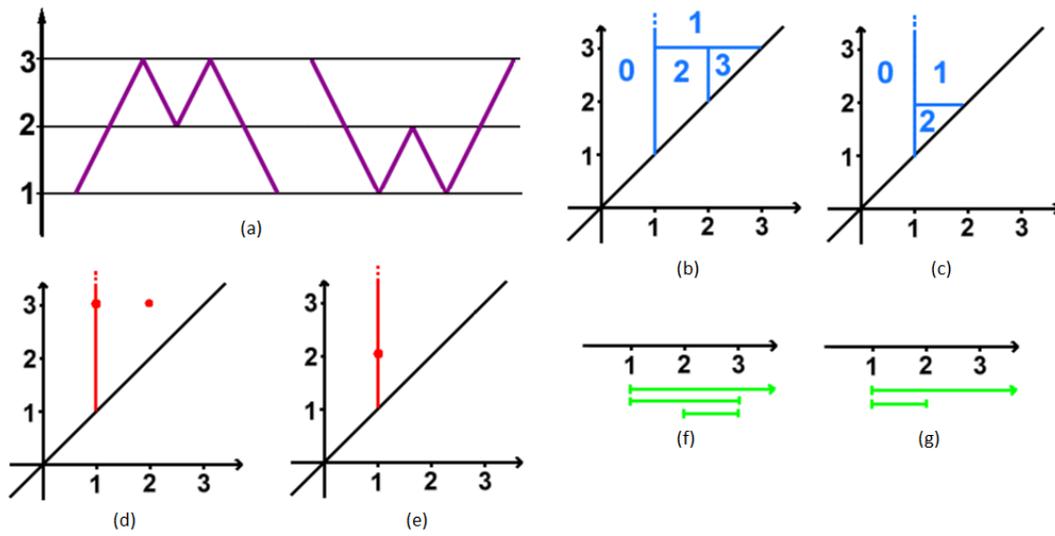


Figura 1.1: (a) I grafemi di ‘M’ e di ‘W’. (b), (c) Le funzioni 0–PBN di ‘M’, ‘W’. (d), (e) 0–diagrammi di persistenza (0–dgm(F)) di ‘M’, ‘W’. (f), (g) Gli 0–barcodes di ‘M’, ‘W’. Immagine tratta da [5].

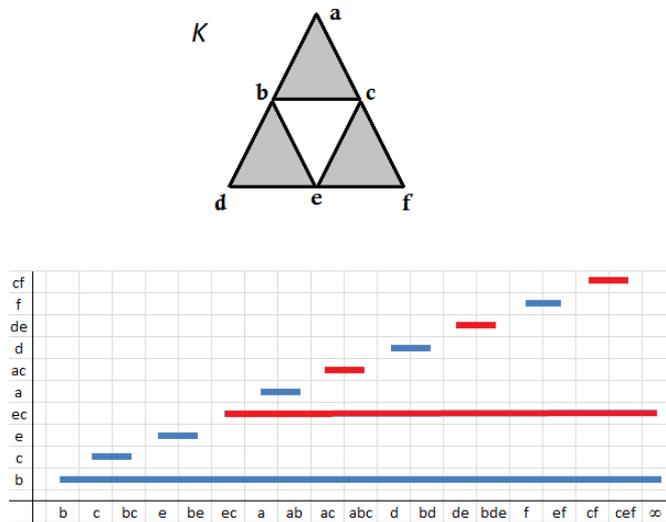


Figura 1.2: Lo 0– e l’ 1–esimo barcode di persistenza di un complesso simpliciale: le barre blu indicano la persistenza delle 0–classi di omologia, le rosse le 1–classi di omologia. Immagine tratta da [11].

rate in blu corrispondono alla persistenza delle 0–classi di omologia (i.e. la persistenza delle componenti connesse), le barre colorate in rosso le 1–classi di omologia (i.e. la persistenza dei buchi 1–dimensionali). La filtrazione sul complesso in questione è  $F = (b, c, bc, e, be, ec, a, ab, ac, abc, d, bd, de, bde, f, ef, cf, cef)$  e può essere letta sull’asse delle ascisse del barcode. In questo caso è un ordine totale.

Possiamo notare che solo una barra rossa ed una barra blu persistono fino alla fine: il complesso è connesso ed ha un solo buco unidimensionale.

**Definizione 20** (Distanza bottleneck). Siano  $F$  e  $F'$  due filtrazioni di, rispettivamente, due complessi simpliciali  $K$  e  $K'$ .

Allora la *distanza bottleneck* tra i rispettivi diagrammi di persistenza  $\text{dgm}(F)$  e  $\text{dgm}(F')$  è definita come:

$$d_b(\text{dgm}(F), \text{dgm}(F')) := \inf_{\gamma} \left\{ \sup_a \{ \|a - \gamma(a)\|_{\infty} \} \right\} \quad (1.11)$$

dove  $\gamma$  varia tra le biezioni che associano a un punto del diagramma  $\text{dgm}(F)$ , o un punto su  $\text{dgm}(F')$  o la sua proiezione sulla diagonale, e la norma infinito  $L_{\infty}$ , dati i punti  $a = (x, y) \in F \cup D$ ,  $\gamma(a) = (x', y') \in F' \cup D$  (sia  $D = \{(x, x) | x \in \mathbb{R}\}$  la diagonale), è definita come  $\|a - \gamma(a)\|_{\infty} := \max\{|x - x'|, |y - y'|\}$ .

# Capitolo 2

## Riconoscimento tramite la GEI

In questa sezione analizzeremo in toto il metodo proposto in [2].

### 2.1 Rappresentazione del cammino con la Gait Energy Image

In primis si selezionano dei frame dal video di un ciclo completo di andatura, quindi si estrae la sagoma dell'individuo da ogni frame, la si allinea in orizzontale e se ne normalizzano le dimensioni usando il metodo proposto in [19]. A questo punto si ottiene una serie di immagini binarie (in bianco e nero) dalla segmentazione (ancora proposta in [19]) delle immagini di cui sopra. Possiamo quindi definire la *Gait Energy Image* (GEI) nel pixel  $(x, y)$  (visto come un punto di una griglia individuato da 2 coordinate) come

$$\text{GEI} = G(x, y) = \frac{1}{T} \sum_{t=1}^T \text{Im}(x, y, t) \quad (2.1)$$

dove:

- $T$  è il numero totale di frame estratti da un ciclo completo di andatura;
- $t$  è il numero del frame nel ciclo;

- $Im(x, y, t)$  è il colore 1 (bianco) o 0 (nero) del pixel di coordinate  $(x, y)$  del  $t$ -esimo frame del ciclo.

Si può pensare alla GEI come un'immagine media (a scala di grigi) delle silhouette binarie estratte da un ciclo di cammino.

Nella figura 2.1 osserviamo in basso la GEI ottenuta dalle sagome sopra.

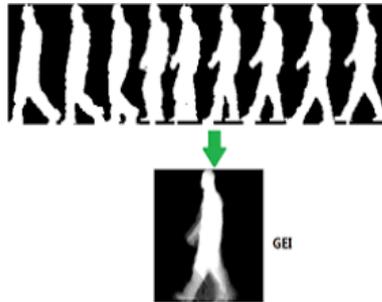


Figura 2.1: Un esempio di generazione della GEI. Immagine tratta da [4].

Si nota che nella GEI ci sono parti più luminose di altre: più una zona ha un alto livello di luminosità, più questa è occupata dal corpo dell'individuo durante il cammino; in sostanza le aree con un'alta intensità corrispondono alle parti del corpo che meno si muovono lungo un ciclo di cammino (chiameremo queste zone *aree statiche*) come ad esempio il busto o la testa, mentre quelle con bassa intensità individuano parti del corpo che si muovono costantemente, come braccia e gambe (chiameremo queste zone *aree dinamiche*).

Le aree dinamiche non sono sensibili a particolari variabili che modificano la silhouette dell'individuo, come per esempio portare un grosso zaino o indossare un cappotto lungo (chiameremo queste situazioni *condizioni covariate*); sono quindi le zone della GEI che possono conferire informazioni più utili.

Anche le aree statiche contengono informazioni utili (es. taglio di capelli), ma purtroppo sono sensibili alle condizioni covariate (es. indossare un cappello).

Partendo da quest'osservazione viene sviluppato nel paragrafo successivo un metodo per selezionare informazioni dalla GEI.

## 2.2 Uso della Gait Entropy Image

Si vogliono distinguere le aree statiche da quelle dinamiche della GEI. Per farlo si calcola di ogni pixel la sua entropia di Shannon, chiamata brevemente GENI (che sta per Gait Entropy Image).

Nello specifico si consideri un ciclo di andatura da cui si ottengono  $T$  sagome. Considerando il valore dell'intensità di un pixel di una sagoma come una variabile aleatoria discreta, l'entropia di Shannon misura il valore atteso dell'informazione contenuta nel pixel, ossia la quantità media di informazione che si ottiene osservando il valore assunto dal pixel.

Essa è definita come:

$$\text{GENI} = H(x, y) = - \sum_{j=1}^K p_j(x, y) \log_2 p_j(x, y) \quad (2.2)$$

dove  $p_j(x, y)$  è la probabilità che il pixel di coordinate  $(x, y)$  assuma il  $j$ -esimo valore. Essendo le sagome immagini binarie,  $K = 2$ ; inoltre si ha che  $p_1(x, y) = \frac{1}{T} \sum_{t=1}^T Im(x, y, t)$  (la GEI in  $(x, y)$ ) e  $p_2 = 1 - p_1$ , per la probabilità complementare.

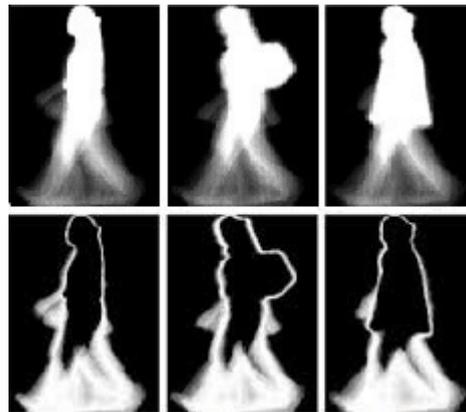


Figura 2.2: Confronto tra GEI e GENI della sagoma di una stessa persona sotto condizioni covariate. Immagine tratta da [2].

Nella figura 2.2 si possono osservare in basso le GENI e in alto le rispettive

GEI. È evidente che le aree dinamiche della GEI sono enfatizzate nella GENI corrispondente; in altri termini, le zone dinamiche sono quelle con maggior entropia, maggiori informazioni, a differenza di quelle statiche.

Purtroppo però la GENI evidenzia anche le condizioni covariate: le caratteristiche selezionate con questo processo contengono ancora informazioni inutili o fuorvianti. Ciò ha indotto gli autori di [2] a formulare un ulteriore metodo di selezione dei dati.

### 2.3 Selezione delle caratteristiche a coppie

Da ora in poi una *gallery* sarà l'insieme delle sagome (nello specifico GEI) per il riconoscimento dell'andatura, delle quali si conosce l'identità dell'individuo corrispondente; invece un *probe* sarà una raccolta di sagome delle quali si ignora l'identità dell'individuo corrispondente.

Quindi, per ogni GEI nel probe, si vuole trovare una corrispondenza nella gallery per risalire all'identità del soggetto. Ci si propone di selezionare delle caratteristiche dell'andatura per ogni coppia gallery-probe in modo che siano invarianti alle condizioni covariate e rilevanti per entrambi. A tal fine si costruisce una funzione di selezione dei pixel per ogni GEI, usando la rispettiva GENI; chiamata *maschera di selezione* o *selection mask* di parametro  $\theta$  (detto *soglia*), essa viene definita come:

$$M_{\theta}(x, y) = \begin{cases} 1, & H(x, y) > \theta, \\ 0, & \text{altrimenti} \end{cases} \quad (2.3)$$

dove  $(x, y)$  sono le coordinate del pixel. Il valore 1 indicherà "pixel selezionato", accendendolo in bianco, 0 "pixel scartato", spegnendolo.

In particolare, data una gallery, se per l' $i$ -esima GEI dell'insieme si usa una certa selection mask  $M_{\theta_i}(x, y)$  e similmente, dato un probe, si usa sulla  $j$ -esima una  $M_{\theta_j}(x, y)$ , per selezionare le caratteristiche rilevanti per entrambi gli elementi della coppia si combinano le due maschere di selezione, andando

ad applicare alle due GEI la nuova funzione:

$$M^{i,j}(x, y) = M_{\theta_i}(x, y) \&\& M_{\theta_j}(x, y) \quad (2.4)$$

dove  $\&\&$  è l'operatore booleano 'AND'. Possiamo considerare  $M^{i,j}$  come l'intersezione delle due selection mask.

La figura 2.3 mostra un esempio di applicazione del metodo di selezione a coppie sopra spiegato. In 2.3(a) e 2.3(b) rispettivamente le GEI dalla gallery e dal probe; in 2.3(c) ciò che viene evidenziato dalla  $M^{i,j}$ ; in 2.3(d) la GEI dalla gallery con  $M^{i,j}$  applicata; in 2.3(e) la GEI dal probe con  $M^{i,j}$  applicata. Si può notare che nell'applicare  $M^{i,j}$  gli effetti delle condizioni covariate diminuiscono notevolmente. In particolare 2.3(d) e 2.3(e) indicano fortemente che 2.3(a) e 2.3(b) appartengono alla stessa persona.

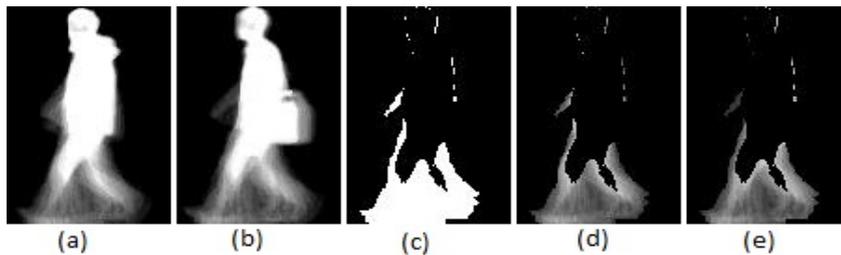


Figura 2.3: Rappresentazione di un esempio di selezione delle caratteristiche a coppie. Immagine tratta da [2].

## 2.4 Cenni su ulteriori metodi di riduzione della dimensione del problema

Dopo aver applicato una  $M^{i,j}$  ad ogni coppia gallery-probe  $(i, j)$ -esima, il problema del riconoscimento dell'andatura può essere risolto osservando quali coppie sono formate da elementi più simili; per fare ciò si possono definire e misurare distanze tra GEI.

Se infatti consideriamo ogni sagoma proveniente da gallery o probe come un vettore colonna (ottenuto concatenando i valori dei pixel illuminati delle righe della rispettiva GEI), ci siamo ricondotti ad un problema vettoriale, dove si possono definire diverse distanze note.

Purtroppo, sebbene la spiegazione dell'essenza del metodo può considerarsi conclusa, c'è bisogno di un'ulteriore riduzione del problema. È stato infatti dimostrato in [7] e [16] che questo tipo di confronto diretto risulta essere troppo sensibile a rumore o a piccole distorsioni della silhouette. Ciò è dovuto al fatto che, nonostante i meccanismi di selezione spiegati, la dimensione dello spazio delle caratteristiche delle GEI resta troppo grande (siamo nell'ordine delle migliaia). Per ovviare al problema, ci si riduce a considerare particolari sottospazi ottenuti usando PCA (Principal Component Analysis) e MDA (Multiple Discriminant Analysis).

La PCA è una particolare trasformazione lineare ortogonale che trasforma i dati, scritti come vettori di una certa dimensione, in vettori appartenenti a sottospazi di dimensione minore. Il sottospazio ottenuto è scelto in modo da mantenere le più grandi varianze dei dati dalla loro proiezione (quindi si conservano le maggiori informazioni) ed in modo da minimizzare un certo errore di ricostruzione.

In particolare: supponiamo di avere le  $N$  GEI di una gallery, indicate con  $x_1, \dots, x_N$ , dove ogni  $x_i$  è un vettore  $d$ -dimensionale. Per calcolare la distanza tra l' $i$ -esima GEI della gallery e la  $j$ -esima del probe, applichiamo a ogni sagoma della gallery la  $M^{i,j}$ . Otteniamo quindi le GEI della gallery scritte come vettori  $d_{i,j}$ -dimensionali:  $x_1^{i,j}, \dots, x_N^{i,j}$ .

La trasformazione PCA  $(i, j)$ -esima individua un sottospazio vettoriale di dimensione  $d'_{i,j}$ , descritto in coordinate rispetto ad una base ortonormale  $\{e_1^{i,j}, \dots, e_{d'_{i,j}}^{i,j}\}$ , tale da minimizzare il seguente errore:

$$J_{d'_{i,j}} = \sum_{n=1}^N \left\| \left( m + \sum_{k=1}^{d'_{i,j}} a_{n,k} e_k^{i,j} \right) - x_n^{i,j} \right\|^2, \quad (2.5)$$

dove  $n$  è l'indice di scorrimento tra le  $N$  GEI della gallery,  $k$  quello tra le coordinate di ogni GEI  $n$ -esima della gallery,  $m$  è la media delle coordinate delle GEI della gallery,  $a_{n,k}$  è la proiezione dell' $n$ -esima GEI della gallery sul vettore  $e_k^{i,j}$ .

Si può vedere che  $J_{d'_{i,j}}$  è minimizzato quando  $\{e_1^{i,j}, \dots, e_{d'_{i,j}}^{i,j}\}$  sono i  $d'_{i,j}$  autovettori dei più grandi autovalori (in ordine decrescente) della matrice delle covarianze (che essendo simmetrica ammette autovettori ortonormali per il teorema spettrale).

Per approfondimenti sulla PCA: [2], [20], [21].

La PCA è seguita dalla MDA, che ha come obiettivo quello di individuare un sottospazio in cui i dati provenienti da diverse *classi* (che nel nostro caso sono gli individui da cui si estraggono le GEI) sono separati al meglio; matematicamente si richiede che i nuovi dati risolvano un certo problema ai minimi quadrati. A differenza della PCA, la MDA è un metodo di apprendimento supervisionato (tecnica di apprendimento automatico che mira a istruire un sistema informatico in modo da consentirgli di elaborare automaticamente previsioni sui valori di uscita di un sistema rispetto ad un input sulla base di una serie di esempi ideali, costituiti da coppie di input e di output, che gli vengono inizialmente forniti) e richiede che i dati degli elementi della gallery siano etichettati in classi. Per approfondimenti a riguardo: [2].

## 2.5 Esperimenti

Sempre in [2], sono stati proposti due diversi esperimenti per testare il metodo proposto.

L'unico parametro libero nel metodo spiegato è  $\theta$ , la soglia delle selection mask, ed è quindi l'unica variabile da fissare negli esperimenti. Si è visto empiricamente in [2] che per  $\theta \in [0, 0.95]$  il tasso di riconoscimento resta utile e non si discosta troppo da un valore medio. Avendo però osservato un picco per  $\theta = 0.75$ , si è scelto quel valore per la soglia.

In entrambi gli esperimenti la gallery e il probe sono stati ottenuti prenden-

do sagome da due database: il *CASIA Gait Database* ed il *Human ID at a distance gait database (SOTON database)*.

Nel primo esperimento la gallery contiene sequenze di andatura di soggetti liberi da condizioni covariate, o sotto le medesime condizioni covariate note in precedenza. Questo primo esperimento è lo stesso fatto nella maggior parte degli altri lavori simili che trattano il problema dell'andatura.

I risultati ottenuti sono soddisfacenti: il metodo ha funzionato nel 100% dei casi di gallery formate da sagome provenienti da CASIA e SOTON database in cui l'individuo era libero da condizioni covariate; nel 78.3% e nell' 81.8% dei casi di gallery (rispettivamente per CASIA e SOTON database) in cui l'individuo trasportava borse o zaini; nel 44.0% e nell' 83.3% dei casi (rispettivamente per CASIA e SOTON database) in cui l'individuo indossava lunghi cappotti.

Confrontando questi dati con quelli in [22], in cui viene fatto un confronto diretto tra GEI senza alcuna selezione delle caratteristiche (come GEnI, selection mask, PCA, MDA), si nota che questo metodo è molto più performante.

Il secondo esperimento, atto a rappresentare uno scenario proveniente dal mondo reale, è simile al precedente ma differisce nella composizione della gallery: sono insiemi di sagome eterogenei, dove gli individui sono sotto diverse condizioni covariate, non note a priori. Questo tipo di esperimento è stato una novità tra i lavori svolti fino a quel momento nell'ambito del riconoscimento dell'andatura; tale sfida però ha messo a dura prova il metodo, che ha mostrato un notevole decremento dell'efficacia.

Le percentuali di riuscita sono state: 53.1% di media per tre diverse gallery formate da sagome provenienti dal CASIA database, 54.5% di media per tre diverse gallery dal SOTON database. In particolare è stato notato che l'utilizzo delle selection mask è stato di cruciale importanza per quest'esperimento, poichè le maschere di selezione riescono a nascondere abbastanza bene le condizioni covariate. Saltare la selezione a coppie avrebbe significato perdere importanti punti percentuale di riuscita.

È importante sottolineare che il metodo descritto funziona solo per sagome riprese da uno stesso punto di vista, o da punti di vista che differiscono di piccoli angoli di visione, mentre in una situazione più realistica l'angolazione delle riprese di una telecamera può essere qualunque. Diversi approcci sono stati proposti per affrontare il problema dell'angolazione, come ad esempio [18] e [8]. Questi metodi possono essere integrati a quello descritto in questo capitolo.



# Capitolo 3

## Riconoscimento tramite la persistenza

In questo capitolo esamineremo il metodo proposto in [11], quale evoluzione finale dei precedenti lavori ([12], [14], [10], [15], [13]) svolti dal medesimo team di ricerca. L'obiettivo è utilizzare la persistenza per ottenere la cosiddetta *firma topologica dell'andatura*, uno strumento topologico utile al riconoscimento dell'identità dell'individuo dal suo modo di camminare, e fornire degli strumenti per il confronto di queste.

### 3.1 Il complesso simpliciale $\partial K(I)$

In questa sezione viene introdotta la costruzione del complesso simpliciale  $\partial K(I)$  associato ad uno o più cicli d'andatura di un individuo.

La procedura inizia con l'estrazione di una serie di silhouette da una sequenza d'andatura: il primo piano è ricavato dallo sfondo con tecniche di segmentazione e sogliatura.

A questo punto si costruisce un'immagine 3D binaria  $I = (\mathbb{Z}^3, B)$  impilando a distanza di 1 le sagome trovate ed allineandone i baricentri (un'immagine 3D binaria è una coppia  $(\mathbb{Z}^3, B)$  dove  $B \subset \mathbb{Z}^3$  è un insieme finito detto primo piano, e  $B^c$  è lo sfondo).

Quindi  $I = (\mathbb{Z}^3, B)$  è usata per costruire un complesso cubico  $Q(I)$ .

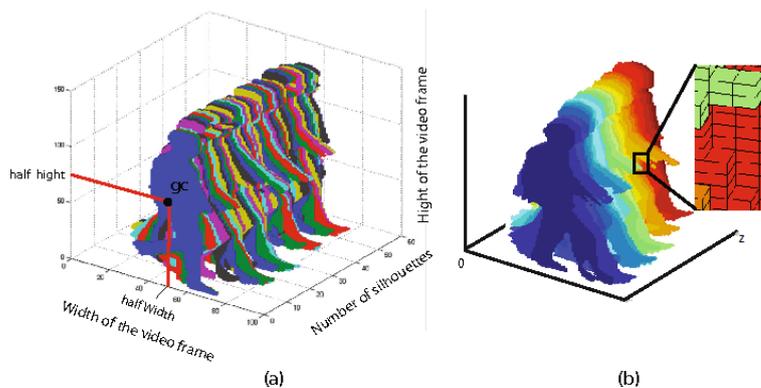


Figura 3.1: (a) Sagome allineate che danno vita a  $I = (\mathbb{Z}^3, B)$ . (b) Il complesso simpliciale  $\partial K(I)$ ; per semplificare la rappresentazione è stata omessa la divisione dei quadrati in triangoli. Immagine tratta da [12].

Nel nostro caso sarà semplicemente una struttura combinatoria costituita da un insieme di cubi unitari aventi i vertici in  $\mathbb{Z}^3$  e le facce parallele ai tre piani generati da ogni coppia di assi cartesiani. Vediamo come si costruisce  $Q(I)$  nel dettaglio.

Un cubo ha per 0-facce i suoi otto vertici, per 1-facce i dodici spigoli, per 2-facce i sei quadrati e l'unica 3-faccia è il cubo stesso. Per ogni  $v = (i, j, k) \in B \subset \mathbb{Z}^3$ , il cubo di vertici  $V = \{(i, j, k), (i + 1, j, k), (i, j + 1, k), (i, j, k + 1), (i + 1, j + 1, k), (i + 1, j, k + 1), (i, j + 1, k + 1), (i + 1, j + 1, k + 1)\}$  è aggiunto a  $Q(I)$  (con tutte le sue facce) se e solo se  $V \subseteq B$ .

Il complesso simpliciale  $\partial K(I)$  è finalmente ottenuto selezionando tutte le 2-facce (quadrati) di  $Q(I)$ , assieme ai loro vertici e spigoli, che siano facce di esattamente un cubo nel complesso cubico (in questo modo escludiamo tutte le facce che non siano esterne); in particolare questi quadrati vengono divisi lungo una diagonale in due triangoli rettangoli.

Si noti che la coordinata  $z$  del complesso simpliciale esprime il numero di sagome impilate.

Per diminuire gli effetti negativi sulla firma topologica dell'andatura dovuti alle variazioni della parte superiore del corpo (legate per esempio al gestic-

lare delle mani durante una telefonata, all'indossare un particolare abbigliamento, al portare uno zaino o una borsa...) si procede come proposto in [10] dagli stessi autori, ossia ci si riduce a considerare per  $\partial K(I)$  solo la quarta parte delle sagome, quella più bassa relativa ai piedi e alla sezione inferiore delle gambe. L'utilità di questa selezione è confermata dai risultati ottenuti in [2], dove viene evidenziato che è la zona del corpo che fornisce la maggior parte delle informazioni utili alla classificazione.

Infine si normalizzano tutte e tre le coordinate  $(x, y, z)$  dei punti di  $\partial K(I)$  in modo che abbiano valori nell'intervallo  $[0, 1]$ .

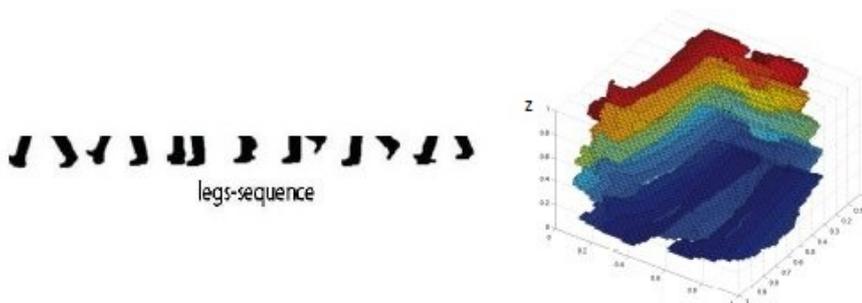


Figura 3.2: Un esempio di  $\partial K(I)$  ottenuto dalla quarta parte inferiore delle sagome. Immagine tratta da [11].

### 3.2 Le filtrazioni sul complesso simpliciale $\partial K(I)$

Il prossimo passo del processo è ordinare i semplici di  $\partial K(I)$ , tramite funzioni filtranti. In particolare saranno ben otto le filtrazioni utilizzate.

Per definirle si considerano otto piani: due verticali, due orizzontali e quattro obliqui, osservabili nella figura 3.3.

Quindi per ogni piano  $\pi$  si definisce la rispettiva funzione filtrante  $f_\pi : \partial K(I) \rightarrow \mathbb{R}$  che assegna ad ogni semplice del complesso simpliciale  $\partial K(I)$  la più grande distanza dei suoi vertici da  $\pi$ .

Indicheremo con  $\partial K_\pi(I)$  la filtrazione ottenuta ordinando i semplici secondo

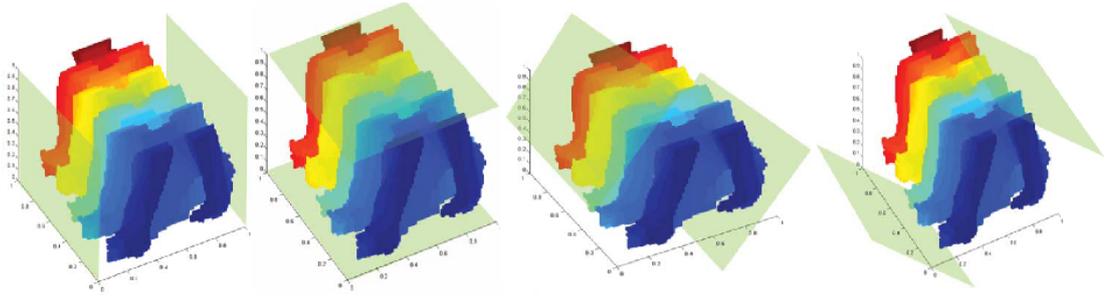


Figura 3.3: Gli otto piani usati per ottenere le otto filtrazioni per  $\partial K(I)$ . Immagine tratta da [11].

il valore di  $f_\pi$ .

Venendo aggiunti più di un semplice per volta (tutti quelli che hanno la stessa distanza dal piano  $\pi$ ), la relazione di ordine che si ottiene tra i semplici non è totale.

Questo modo di ordinare i semplici rende il metodo robusto alla variazione della quantità di semplici presenti in  $\partial K(I)$ , robusto al rumore e più stabile. Prima di [11] e [10] gli autori proponevano funzioni filtranti definite diversamente, che inducevano un ordine totale sui semplici, ma così facendo il metodo risultava essere troppo sensibile al rumore e al numero di semplici, e poco stabile.

### 3.3 Omologia persistente per la firma topologica

L'ultimo step del processo per ottenere una firma topologica di una sequenza di andatura è calcolare l'omologia persistente per ognuna delle otto filtrazioni.

Per ogni  $\partial K_\pi(I)$  indotta da  $f_\pi$ , costruiamo i rispettivi 0- e 1-barcode di persistenza.

Osserviamo che, fissato un piano di riferimento  $\pi$ , la lunghezza di ogni intervallo nel barcode per  $\partial K_\pi(I)$  è:

- $\leq 1$  se  $\pi$  non è obliquo;
- $\leq \sqrt{2}$  se  $\pi$  è obliquo.

Quindi in generale ha lunghezza  $\leq \sqrt{2}$ .

Per la firma topologica considereremo solo le barre di lunghezza  $> 0$ , ossia vengono esclusi gli eventi (classi di omologia) nati e morti nello stesso momento, alla stessa distanza dal piano di riferimento. Ciò non è un problema: gli stessi eventi verranno catturati considerando un piano diverso.

Ora, per ogni piano  $\pi$ , vengono studiati lo 0– e l'1– barcode di persistenza (le persistenze delle componenti connesse e delle cavità 1–dimensionali) relativo alla filtrazione  $\partial K_\pi(I)$ . Più precisamente, dato un intero  $n$  (negli esperimenti è stato usato  $n = 24$ ) che esprime il numero di ‘finestre’ usate per valutare il barcode, si calcola il valore  $h = \frac{k}{n}$  che esprime la larghezza di queste, dove  $k$  è la più grande distanza di un vertice di  $\partial K(I)$  dal piano  $\pi$  fissato: quindi  $k \leq \sqrt{2}$  e  $h \leq \frac{\sqrt{2}}{n}$ .

La firma topologica per una sequenza di andatura è finalmente ottenuta: sarà l'insieme dei sedici vettori  $2n$ –dimensionali  $\{\mathcal{V}_{\pi_i}^0, \mathcal{V}_{\pi_i}^1\}_{i=1,\dots,8}$ , dove il vettore  $\mathcal{V}_{\pi_i}^0$  (risp.  $\mathcal{V}_{\pi_i}^1$ ) di  $\mathbb{R}^{2n}$  è costruito come segue:

per  $s = 0, \dots, n - 1$ :

- a) la componente  $2s$ –esima è il numero di 0–classi di omologia (risp. 1–classi di omologia) visibili nel barcode di  $\partial K_\pi(I)$  nel punto  $s \cdot h$ , che sono nate prima di  $s \cdot h$ ;
- b) la componente  $(2s+1)$ –esima è il numero di 0–classi di omologia (risp. 1–classi di omologia) visibili nel barcode che sono nate dall'istante  $s \cdot h$  in poi, fino a prima di  $(s+1) \cdot h$ .

Osserviamo che dividendo le entrate dei vettori nelle due categorie a) e b), vengono evidenziati più dettagli, il che è cruciale per distinguere due diverse andature.

Per esempio supponiamo di avere  $m$  0–classi di omologia nate in  $[s \cdot h, (s+1)h)$  che persistono o muoiono in  $(s+2)h$ , alla fine di  $[(s+1)h, (s+2)h)$ , e che

nessun'altra classe di 0-omologia sia nata, persiste, o muore in questi intervalli. Allora si pone 0 nell'entrata  $2s$ -esima e nell'entrata  $(2s + 3)$ -esima di  $\mathcal{V}_{\pi_i}^0$ , e  $m$  nelle entrate  $2s + 1$  e  $2s + 2$  di  $\mathcal{V}_{\pi_i}^0$ .

Oppure, supponiamo di avere  $m$  classi di 0-omologia nate e morte nell'intervallo  $[s \cdot h, (s + 1)h)$  ed  $m$  nell'intervallo  $[(s + 1)h, (s + 2)h)$ , e nessun'altra classe di 0-omologia è nata, morta o persiste in questi intervalli. Allora si pone 0 nelle entrate  $2s$  e  $2s + 2$  di  $\mathcal{V}_{\pi_i}^0$  e  $m$  nelle entrate  $2s + 1$  e  $2s + 3$  dello stesso vettore.

Per confrontare due firme topologiche  $\{\mathcal{V}_{\pi_i}^0, \mathcal{V}_{\pi_i}^1\}_{i=1, \dots, 8}$  e  $\{\mathcal{W}_{\pi_i}^0, \mathcal{W}_{\pi_i}^1\}_{i=1, \dots, 8}$  viene calcolata la somma degli angoli fra i vettori di ogni coppia  $(\mathcal{V}_{\pi_i}^j, \mathcal{W}_{\pi_i}^j)$  con  $j = 0, 1$ ,  $i = 1, \dots, n$ . Essendo la firma costituita di sedici vettori aventi tutte le coordinate non negative, due firme saranno tanto più simili quanto la somma si avvicina a  $0^\circ$  e tanto più diverse quanto la somma si avvicina a  $90^\circ \cdot 16 = 1440^\circ$ .

Nei precedenti lavori svolti dagli stessi autori di [11] non venivano valutati gli angoli tra vettori, ma i coseni di questi. Si è notato però che usando direttamente gli angoli la precisione subiva un incremento del 5%.

### 3.4 Stabilità della firma topologica

In [11] è stato dimostrato che il metodo proposto in questo capitolo è stabile. In particolare considerando due immagini binarie 3D  $I, I'$  simili abbastanza, la distanza bottleneck tra i due diagrammi di persistenza relativi alle filtrazioni  $\partial K_\pi(I), \partial K_\pi(I')$  (indotte dalle funzioni filtranti relative allo stesso piano  $\pi$ ) resta piccola a piacere: ossia per piccole variazioni dell'immagine binaria 3D iniziale, non si hanno grossi cambiamenti nella firma. Ciò rende il metodo robusto al rumore.

È stato inoltre mostrato che se la sequenza di andatura contiene più di un ciclo completo, il modulo dei vettori di  $\{\mathcal{V}_{\pi_i}^0, \mathcal{V}_{\pi_i}^1\}_{i=1, \dots, 8}$  aumenta, seppur la direzione di questi, in modo approssimativo, non cambi.

Questo perchè i diagrammi di persistenza risultano avere più punti. Se infatti

ti una sequenza contiene due cicli completi di andatura, il numero di punti presenti nel diagramma di persistenza sarà circa il doppio di quelli di un diagramma di persistenza di una sequenza contenente solo un ciclo completo; conseguentemente, per come sono stati definiti, anche il modulo dei vettori della firma all'incirca raddoppierà.

Uno strumento classicamente usato per confrontare due diagrammi di persistenza è la distanza bottleneck. Comparando due diagrammi relativi a due sequenze di andatura con la distanza bottleneck, questa tenderà a zero all'assomigliarsi dei due, ma la velocità con cui lo farà, si è visto in [11], dipende purtroppo dal numero di cicli di andatura: la distanza resterà tanto più grande quanti sono i cicli completi nella sequenza.

Invece un confronto fatto misurando i coseni degli angoli tra i vettori della firma, o misurando direttamente gli angoli, non risente di questo fattore. Per avere un tipo di stima il più generico possibile, è stato pertanto preferito dagli autori l'uso del confronto di firme tramite angoli o coseni.

### 3.5 Esperimenti

Gli esperimenti condotti in [11] sono stati fatti sulle sequenze di andatura presenti nel *CASIA-B gait dataset*.

Sono stati eseguiti gli stessi due esperimenti fatti in [2], di cui ricordiamo una metà essere stati svolti sul CASIA gait dataset, come quelli in questa sezione 3.5.

Ciò ci dà l'opportunità di comparare in maniera oggettiva i risultati ottenuti dai due metodi.

Raffrontando le percentuali di riuscita del primo esperimento, ricordiamo che il metodo [2] funzionò rispettivamente nel 78.3%, 44.0%, 100% dei casi (a seconda delle situazioni in cui veniva immortalato il soggetto della gallery), con una media del 74.1%, mentre il metodo [11] spiegato in questo capitolo si è visto avere successo rispettivamente nel 80.5%, 81.7%, 92.4% dei casi, con una media del 84.9%. Notiamo dunque un miglioramento in due test su tre,

e una comunque considerevole percentuale di riuscita per il terzo test.

Ma è nel secondo esperimento, quello più ambizioso e più simile ad un potenziale scenario reale, che il riconoscimento tramite la persistenza ha manifestato la sua potenza: la percentuale di riuscita è stata un notevole 93.8% di media a fronte del 53.1% registrato in [2].

Precisiamo che anche il metodo descritto in questo capitolo è stato ideato e testato per confronti tra sagome riprese da uno stesso punto di vista. I procedimenti proposti in [18] e [8] possono dunque essere integrati per provare ad ovviare al problema.

# Conclusioni

In questo elaborato, dopo aver analizzato brevemente lo stato dell'arte, abbiamo esaminato due diversi metodi per il riconoscimento dell'andatura, con particolare interesse ad un utilizzo nella videosorveglianza.

Per i due approcci, entrambi appearance-based, ci siamo serviti rispettivamente della teoria dell'informazione e della topologia persistente, i cui prerequisiti sono stati forniti nel Capitolo 1.

Si è visto negli esperimenti che i due metodi (ideati per confronti tra sagome prese da uno stesso punto di vista) funzionano bene se usati su individui liberi da condizioni covariate, o su individui sotto le medesime condizioni covariate note in precedenza. Mentre nel caso di gallery eterogenee solo il secondo approccio dà buoni risultati, grazie soprattutto alla restrizione dello studio alla parte inferiore delle gambe.

In più usando la persistenza otteniamo un metodo stabile, robusto al rumore, capace di tener conto dell'ordine temporale delle sagome in un ciclo di andatura e adatto a trattare anche più cicli completi per individuo (grazie al confronto fatto con angoli o loro coseni).

I futuri lavori sul problema in questione potranno essere finalizzati ad ottenere metodi appearance-based capaci di confrontare sagome provenienti da diverse angolazioni e in grado di trattare un più ampio range di condizioni covariate, così da poter pensare ad una applicazione reale.



# Bibliografia

- [1] Gunawan Ariyanto and Mark S Nixon. Model-based 3d gait biometrics. 2011.
- [2] Khalid Bashir, Tao Xiang, and Shaogang Gong. Gait recognition without subject cooperation. *Pattern Recognition Letters*, 31(13):2052–2060, 2010.
- [3] Robert Bodor, Andrew Drenner, Duc Fehr, Osama Masoud, and Nikolaos Papanikolopoulos. View-independent human motion classification using image-based reconstruction. *Image and Vision Computing*, 27(8):1194–1206, 2009.
- [4] Yan Chen, Yawei Yang, and Jeongkyu Lee. Gait based gender classification using kinect sensor. 2014.
- [5] Massimo Ferri. Persistent topology for natural data analysis—a survey. In *Towards Integrative Machine Learning and Knowledge Extraction*, pages 117–133. Springer, 2017.
- [6] Michela Goffredo, Imed Bouchrika, John N Carter, and Mark S Nixon. Self-calibrating view-invariant gait biometrics. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 40(4):997–1008, 2010.
- [7] Jinguang Han and Bir Bhanu. Individual recognition using gait energy image. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (2):316–322, 2006.

- 
- [8] Worapan Kusakunniran, Qiang Wu, Hongdong Li, and Jian Zhang. Multiple views gait recognition using view transformation model based on optimized gait energy image. In *Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on*, pages 1058–1064. IEEE, 2009.
- [9] Worapan Kusakunniran, Qiang Wu, Jian Zhang, Hongdong Li, and Liang Wang. Recognizing gaits across views through correlated motion co-clustering. *IEEE Transactions on Image Processing*, 23(2):696–709, 2014.
- [10] Javier Lamar-Leon, Raul Alonso-Baryolo, Edel Garcia-Reyes, and Rocio Gonzalez-Diaz. Persistent homology-based gait recognition robust to upper body variations. In *Pattern Recognition (ICPR), 2016 23rd International Conference on*, pages 1083–1088. IEEE, 2016.
- [11] Javier Lamar-Leon, Raul Alonso-Baryolo, Edel Garcia-Reyes, and Rocio Gonzalez-Diaz. Persistent-homology-based gait recognition. *arXiv preprint arXiv:1707.06982*, 2017.
- [12] Javier Lamar-Leon, Raul Alonso Baryolo, Edel Garcia-Reyes, and Rocio Gonzalez-Diaz. Gait-based carried object detection using persistent homology. In *Iberoamerican Congress on Pattern Recognition*, pages 836–843. Springer, 2014.
- [13] Javier Lamar-León, Edel B Garcia-Reyes, and Rocio Gonzalez-Diaz. Human gait identification using persistent homology. In *Iberoamerican Congress on Pattern Recognition*, pages 244–251. Springer, 2012.
- [14] Javier Lamar Leon, Raúl Alonso, Edel Garcia Reyes, and Rocio Gonzalez Diaz. Topological features for monitoring human activities at distance. In *International Workshop on Activity Monitoring by Multiple Distributed Sensing*, pages 40–51. Springer, 2014.

- 
- [15] Javier Lamar Leon, Andrea Cerri, Edel Garcia Reyes, and Rocio Gonzalez Diaz. Gait-based gender classification using persistent homology. In *Iberoamerican Congress on Pattern Recognition*, pages 366–373. Springer, 2013.
- [16] Zongyi Liu and Sudeep Sarkar. Outdoor recognition at a distance by fusing gait and face. *Image and Vision Computing*, 25(6):817–832, 2007.
- [17] Haiping Lu, Konstantinos N Plataniotis, and Anastasios N Venetsanopoulos. A layered deformable model for gait analysis. In *Automatic Face and Gesture Recognition, 2006. FGR 2006. 7th International Conference on*, pages 249–254. IEEE, 2006.
- [18] Yasushi Makihara, Ryusuke Sagawa, Yasuhiro Mukaigawa, Tomio Echigo, and Yasushi Yagi. Gait recognition using a view transformation model in the frequency domain. In *European Conference on Computer Vision*, pages 151–163. Springer, 2006.
- [19] Sudeep Sarkar, P Jonathon Phillips, Zongyi Liu, Isidro Robledo Vega, Patrick Grother, and Kevin W Bowyer. The humanoid gait challenge problem: Data sets, performance, and analysis. *IEEE transactions on pattern analysis and machine intelligence*, 27(2):162–177, 2005.
- [20] Jonathon Shlens. A tutorial on principal component analysis. *arXiv preprint arXiv:1404.1100*, 2014.
- [21] Lindsay I Smith. A tutorial on principal components analysis. Technical report, 2002.
- [22] Shiqi Yu, Daoliang Tan, and Tieniu Tan. A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition. In *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, volume 4, pages 441–444. IEEE, 2006.

- [23] Rong Zhang, Christian Vogler, and Dimitris Metaxas. Human gait recognition at sagittal plane. *Image and vision computing*, 25(3):321–330, 2007.
- [24] Guoying Zhao, Guoyi Liu, Hua Li, and Matti Pietikainen. 3d gait recognition using multiple cameras. In *Automatic Face and Gesture Recognition, 2006. FGR 2006. 7th International Conference on*, pages 529–534. IEEE, 2006.

# Elenco delle figure

1.1	(a) I grafemi di ‘M’ e di ‘W’. (b), (c) Le funzioni 0–PBN di ‘M’, ‘W’. (d), (e) 0–diagrammi di persistenza ( $0\text{-dgm}(F)$ ) di ‘M’, ‘W’. (f), (g) Gli 0–barcodes di ‘M’, ‘W’. Immagine tratta da [5]. . . . .	17
1.2	Lo 0– e l’ 1–esimo barcode di persistenza di un complesso simpliciale: le barre blu indicano la persistenza delle 0–classi di omologia, le rosse le 1–classi di omologia. Immagine tratta da [11]. . . . .	17
2.1	Un esempio di generazione della GEI. Immagine tratta da [4].	20
2.2	Confronto tra GEI e GENI della sagoma di una stessa persona sotto condizioni covariate. Immagine tratta da [2]. . . . .	21
2.3	Rappresentazione di un esempio di selezione delle caratteristiche a coppie. Immagine tratta da [2]. . . . .	23
3.1	(a) Sagome allineate che danno vita a $I = (\mathbb{Z}^3, B)$ . (b) Il complesso simpliciale $\partial K(I)$ ; per semplificare la rappresentazione è stata omessa la divisione dei quadrati in triangoli. Immagine tratta da [12]. . . . .	30
3.2	Un esempio di $\partial K(I)$ ottenuto dalla quarta parte inferiore delle sagome. Immagine tratta da [11]. . . . .	31
3.3	Gli otto piani usati per ottenere le otto filtrazioni per $\partial K(I)$ . Immagine tratta da [11]. . . . .	32