

Analisi tempo-frequenza e multiscala - Homework 1

Michele Lupini s316676, Biagio Torsello s315050

September 21, 2023

0 Introduzione

La seguente relazione descrive il lavoro svolto per l'implementazione e l'analisi di un algoritmo per il riconoscimento facciale ("Eigenfaces"). L'ordine espositivo è il seguente:

- presentazione generale del problema affrontato;
- presentazione dell'algoritmo implementato: verranno descritte le due fasi di training e di test, discutendo i metodi implementati e la sensitività degli stessi al variare di determinati parametri.

1 Presentazione del problema

La Principal Component Analysis è un metodo largamente utilizzato per l'archiviazione delle immagini. Sia $D = \{f_1, f_2, \dots, f_L\}$ un dataset di immagini, nel nostro caso volti umani, aventi tutte la stessa dimensione ($m \times n$ pixels) ed in scala di grigio. Il presupposto fondamentale è considerare tali immagini come elementi di uno spazio vettoriale di dimensione $m \times n$:

$$f_i \in \mathbb{R}^{mn}, \quad \forall i = 1, \dots, L$$

La base canonica di tale spazio è quella costituita dagli elementi e_k , $k = 1, \dots, mn$, tali che:

$$(e_k)_i = \delta_{ik}, \quad \|e_k\| = 1$$

Possiamo pensare agli elementi e_k come a delle immagini in cui si "accende" un pixel alla volta. L'obiettivo della Principal Component Analysis è quello di rappresentare le immagini nel modo più sparso possibile, ovvero trovare una nuova base di vettori u_k , $k = 1, \dots, nm$ tali che l'energia di ogni segnale sia concentrata su poche componenti della base. In termini matematici, detti α_{ki} i coefficienti di f_i rispetto a tale base, allora:

$$\sum_{k=1}^{nm} |\alpha_{ki}|^2 \simeq \sum_{k=1}^P |\alpha_{ki}|^2, \quad P \ll mn, \quad \forall i = 1, \dots, L$$

Il chiaro vantaggio sta nel fatto che in questo modo riduciamo notevolmente la dimensionalità del nostro problema, sia in termini di archiviazione sia in termini di riconoscimento/confronto dei volti all'interno od all'esterno del dataset.

La determinazione della nuova base si effettua procedendo con il seguente metodo iterativo.

1.1 Metodo iterativo

- **Passo base:** si definisce la faccia media

$$\tilde{f} = \frac{1}{L} \sum_{i=1}^L f_i$$

Si calcolano le quantità:

$$\phi_i = f_i - \tilde{f}$$

Sia inoltre dato $j = 1$.

Parte iterativa:

1. si determina il sottospazio di dimensione j che approssima “meglio” le facce (ovvero le quantità ϕ , depurate dalla media), ovvero quello generato dai vettori $\{u_1, \dots, u_j\}$ tali che $\alpha_1, \dots, \alpha_j$ (i coefficienti che rappresentano le facce rispetto a questa nuova base; per semplicità è stato ommesso il doppio indice introdotto in precedenza) siano **mediamente** i più grandi possibili (e saranno in ordine decrescente, per costruzione). In riferimento al caso $j = 1$, con il termine “**mediamente**” si intende che, in teoria, il vettore u_1 dovrebbe rendere massima la seguente quantità:

$$\mathbb{E}[|\alpha_1|^2]$$

Questo equivale ad interpretare ogni faccia come una variabile aleatoria (di media nulla) la cui realizzazione è salvata nel nostro dataset. Tale impostazione è però poco utile dal punto di vista pratico; dunque, al valore atteso si sostituisce la media campionaria, ottenendo il seguente problema di massimizzazione:

$$\max_{\|x\|=1} \frac{1}{L} \sum_{i=1}^L |\langle \phi_i, x \rangle|^2 \quad (1)$$

Il vettore u_1 è quello risultante dalla risoluzione del problema variazionale ottenuto. Introduciamo ora l'operatore (matrice) Φ^* tale che

$$\Phi^*(i, :) = \phi_i, \quad \forall i = 1, \dots, L$$

$$\Phi^*(x) = (\langle x, \phi_1 \rangle, \dots, \langle x, \phi_L \rangle)$$

Si può dimostrare che la risoluzione di (1) equivale alla ricerca degli autovettori della **matrice di covarianza** definita da:

$$\frac{1}{L} \Phi \Phi^*$$

Nel caso $j = 1$, l'autovalore massimo, che indichiamo con λ_1 , rappresenta il valore che massimizza la media campionaria in (1) ed il corrispondente autovettore è proprio la soluzione dello stesso, ovvero u_1 .

2. $j = j + 1$, e si ritorna al punto 1, se $j \leq mn$.

Evidenziamo i seguenti aspetti:

- risultati di Analisi Funzionale (teorema spettrale per operatori autoaggiunti) garantiscono che: $\lambda_1 \geq \lambda_2 \geq \dots \geq 0$ e u_1, u_2, \dots costituiscono una base ortonormale di \mathbb{R}^{mn} ;
- $\text{rango}(\Phi \Phi^*) \leq L$, perciò si hanno al più L autovalori non nulli. Da ciò deduciamo che abbiamo effettivamente ridotto la dimensionalità del problema, in quanto proiettiamo ogni faccia nello spazio degli autovettori relativi a tali autovalori. Si può ridurre ancora la dimensionalità imponendo a livello pratico la seguente condizione:

$$\frac{\sum_{k=1}^{L'} \lambda_k}{\sum_{k=1}^L \lambda_k} \geq \text{var}, \quad 0 < \text{var} < 1, \quad L' < L$$

Tale condizione equivale a scegliere un nuovo sottospazio di autovettori $S_{L'} = \{u_1, \dots, u_{L'}\}$ che trattiene il $(100\text{var})\%$, della varianza, ovvero della variabilità attorno alla faccia media ([Nic23]);

- alla luce di quanto appena detto, si definisce **l'operatore di proiezione**:

$$A : \mathbb{R}^{nm} \rightarrow \mathbb{R}^{L'}$$

$$Ax = (\alpha_1, \dots, \alpha_{L'}), \quad \alpha_k = \langle x, u_k \rangle$$

ed ogni faccia f_i può essere rappresentata (dunque archiviata, confrontata, ...) in maniera approssimata attraverso i coefficienti $Af_i \in \mathbb{R}^{L'}$. Ad esempio, per confrontare due facce si calcola:

$$\|Ax - Ay\|^2 = \sum_{k=1}^{L'} |\alpha_k - \beta_k|^2$$

Definiamo la **distanza** ϵ tra una faccia x e lo spazio $S_{L'}$ (spazio delle *autofacce*) nel seguente modo:

$$\epsilon = \left\| (x - \tilde{f}) - \sum_{i=1}^{L'} (Ax)_i u_i \right\| \quad (2)$$

- Infine, osserviamo che $\Phi\Phi^*$ è una matrice di taglia $nm \times nm$, mentre $\Phi^*\Phi$ è di taglia $L \times L$; tuttavia, esse hanno gli stessi autovalori non nulli. Come vedremo dopo, questo consente di ridurre il costo computazionale del problema agli autovettori di $\Phi\Phi^*$ sfruttando adeguatamente la proprietà appena enunciata.

2 Discussione dell'algoritmo implementato e dei risultati ottenuti

L'algoritmo prevede due fasi: la prima è quella di **training**, in cui si definisce lo spazio delle facce, e si implementa la procedura descritta sopra per la proiezione delle facce sul sottospazio degli L' autovettori della matrice di covarianza; la seconda è quella di **test**, durante la quale si stabilisce se un data faccia appartiene o meno al database e, in caso positivo, a quale soggetto corrisponde.

2.1 Parametri specifici del problema

Riportiamo le caratteristiche specifiche del dataset utilizzato, definendo i valori delle quantità menzionate in (1):

- $m = 112$, $n = 92$, da cui $m \times n = 10304$;
- il dataset è costituito da 40 soggetti, per ognuno dei quali disponiamo di 10 foto; come caso base scegliamo 6 foto per ogni soggetto nella fase di training. Questo implica che $L = 240$ (nel seguito varieremo tale parametro). Le rimanenti facce sono poi utilizzate nella fase di test.

2.2 Fase di training

Riportiamo i metodi implementati per la definizione del processo iterativo descritto in (1.1). I valori di m, n, L sono quelli riportati in (2.1). Trasformiamo innanzitutto ogni immagine del dataset in un vettore di \mathbb{R}^{mn} , e salviamo i risultati in una matrice $f \in \mathbb{R}^{L \times mn}$. Essa corrisponde alla matrice Φ^* definita in (1.1). Definiamo la faccia media, le quantità ϕ_l , $l = 1, \dots, mn$ e la matrice di covarianza:

```

1 L = np.shape(f)[0]
2 meanface = np.sum(f, axis=0)/L
3
4 phi = np.zeros((L,mn))
5 phi = f - f_meanface
6
7 mn = np.shape(f)[1]
8 C = np.zeros((mn,mn))
9 C = np.dot(np.transpose(f), f)/L #matrice di covarianza

```

Si procede ora con il calcolo degli autovettori e degli autovalori della matrice di covarianza C . Come accennato in (1.1), anzichè risolvere il problema spettrale relativo a C (matrice di dimensione 10304×10304) possiamo procedere al calcolo degli autovettori della matrice:

$$C_{reduced} = \Phi^* \Phi \in \mathbb{R}^{240 \times 240}$$

A partire da essi è possibile risalire agli autovettori di C mediante la procedura riportata in [Mat91b, p.4-5] ed implementata nel nostro codice:

$$u_i = \Phi \cdot v_i,$$

dove v_i è un autovettore di $C_{reduced}$ ed u_i il corrispondente autovettore di C . Prima del calcolo degli autovettori di C , si procede alla riduzione della dimensionalità dello spazio degli autovettori di $C_{reduced}$ mediante il criterio basato sulla varianza descritto in (1.1): abbiamo dunque $L' < L$ autovettori significativi (l'indice i varia cioè tra 1 e L').

```

1 def eig_covariance_reduced(f, var):
2
3     L_red = -1 #L_red = L'
4
5     L = np.shape(f)[0]
6     phi = np.transpose(f)
7
8     C_reduced = np.zeros((L,L))
9     C_reduced = np.dot(f, phi)/L
10
11     eigenvalues, eigenvectors = np.linalg.eig(C_reduced)
12
13     # Ordina gli autovalori in ordine decrescente
14     # e gli autovettori corrispondenti
15     sorted_indices = np.argsort(eigenvalues)[::-1]
16     aval = eigenvalues[sorted_indices]
17     avec = eigenvectors[:, sorted_indices]
18
19     # implementiamo il criterio della varianza
20     sum_tot = sum(aval)
21     sum_aval = 0
22     for i in range(L):
23         sum_aval = sum_aval + aval[i]
24         rapp = sum_aval/sum_tot
25         if rapp > var:
26             L_red = i
27             break
28
29     aval_red = aval[0:L_red]
30     avec_red = avec[:, 0:L_red]
31
32     #a partire dagli autovettori di C_reduced otteniamo quelli di C
33     u = np.dot(phi, avec_red)
34     u = u/(linalg.norm(u)) #per normalizzare
35
36     return aval_red, u, L_red

```

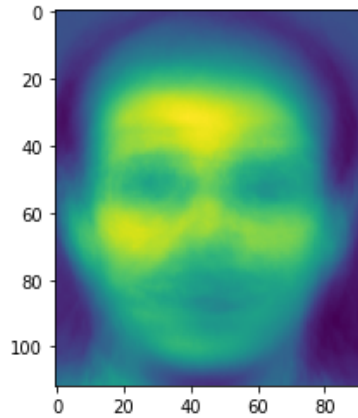


Figure 1: Faccia media.

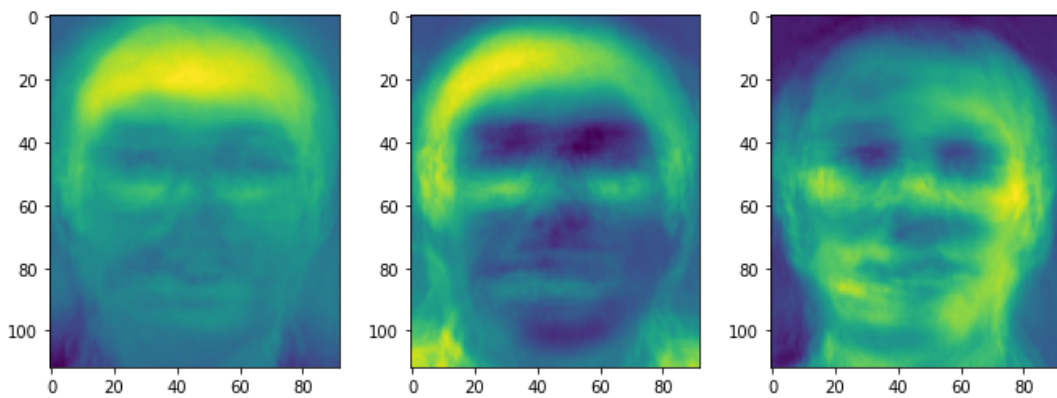


Figure 2: Tre autofacce calcolate a partire dal dataset considerato.

L'ultimo passo della fase di training consiste nel proiettare le facce sul sottospazio degli L' autovettori rimasti:

```
1 projection_training = np.zeros((L, L_reduced))
2 projection_training = np.dot(phi, avec_reduced)
```

2.3 Fase di test

In questa fase considereremo le facce non utilizzate nella fase di training come facce “nuove”, per testare la robustezza dell'algoritmo implementato nel riconoscere se la faccia è presente o meno nel database (e, se lo è, nell'associarla al soggetto corretto). Svolgeremo dunque uno studio parametrico per vedere se e quanto impattano i seguenti parametri nei risultati: L (dunque numero di facce usate per il training), var (e conseguentemente L'), ed infine le soglie θ e θ_k , che definiamo come segue:

- θ : è la soglia riferita alla distanza ϵ tra la nuova faccia e lo spazio delle autofacce. In altri termini, se $\epsilon > \theta$, allora il soggetto è “lontano” dallo spazio delle autofacce, dunque non riconosciuto all'interno del dataset:

```
1 def face_recognition(f_new, theta):
2
3     IsRecognized = 0
4
```

```

5      # calcoliamo lo scarto tra la nuova immagine f_new
6      # e la faccia media
7      phi_new = f_new - hm.f_meanface
8
9      # proiettiamo phi_new nello spazio delle autofacce
10     projection_new = np.dot(phi_new, hm.avec_reduced)
11
12     # calcoliamo la minima distanza euclidea tra phi_new
13     # e lo spazio generato dalle autofacce
14     eps = linalg.norm(phi_new - np.dot(hm.avec_reduced, projection_new))
15
16     if eps < theta:
17         IsRecognized = 1
18
19     return IsRecognized, eps, projection_new

```

- θ_k : è la soglia riferita al controllo successivo a quello menzionato in precedenza. Si procede con un metodo iterativo volto al calcolo della distanza minima tra la faccia e le proiezioni delle facce di training sullo spazio delle autofacce. Se la distanza minima ϵ_{min} è minore di θ_k , allora si considera l'associazione che l'algoritmo ha eseguito e successivamente si verifica se essa è corretta o meno (potremmo infatti essere in presenza di un “falso positivo”, si veda (2.3.1)):

```

1     def which_subject(projection_new, theta_k):
2         subject = 0
3         flag = 1
4         eps_min = linalg.norm(projection_new - hm.projection_training[0,:])
5
6         k = 1
7         for proj in hm.projection_training[1:]:
8             eps_temp = linalg.norm(projection_new - proj)
9             if eps_temp < eps_min:
10                 eps_min = eps_temp
11                 subject = k // hm.n_faces_training
12             k = k + 1
13
14         if eps_min > theta_k:
15             flag = 0
16
17         return subject, flag, eps_min

```

Prima di procedere con uno studio parametrico, nel quale verranno analizzati i risultati ottenuti facendo variare uno alla volta i parametri descritti in precedenza, chiariamo alcuni aspetti propedeutici alla comprensione dei risultati.

2.3.1 Precisazioni propedeutiche allo studio parametrico

- **Caso studio:** il caso studio “base” dal quale partiremo è il seguente: $var = 0.99$, $L = 240$, $\theta = 5000$, $\theta_k = 5000$. Ciò vuol dire che in ogni studio parametrico ciascuna quantità assumerà il valore appena specificato, ad eccezione del parametro che varia nel caso in esame.
- Il controllo effettuato sulle soglie θ e θ_k può portare a 4 scenari diversi:
 - $\epsilon < \theta$ e $\epsilon_{min} < \theta_k$: la faccia è riconosciuta come tale ed associata all'individuo corretto;
 - $\epsilon < \theta$ ma $\epsilon_{min} > \theta_k$: la faccia è riconosciuta come tale, ma l'algoritmo non la associa a nessun individuo e si potrebbe trattare di un individuo sconosciuto;

- $\epsilon > \theta$ e $\epsilon_{min} < \theta_k$: la faccia non è riconosciuta come tale, ma è “vicina” (nel senso di distanza introdotto in (2)) ad una classe di facce nel dataset (e dunque erroneamente associata ad essa): in questo caso si parla di **falso positivo**;
- $\epsilon > \theta$ e $\epsilon_{min} > \theta_k$: la faccia è “lontana” sia dallo spazio delle autofacce che dalle classi di facce nel nostro dataset.

Quanto appena esposto riprende le casistiche possibili descritte in [Mat91b], [Mat91a].

- Precisiamo la terminologia che useremo nell’esposizione dei risultati nelle sezioni successive. Con l’espressione “effettivi falsi positivi” indichiamo tutte quelle facce che superano la seconda soglia ($\epsilon_{min} < \theta_k$) e non la prima ($\epsilon > \theta$), ma, a differenza dei falsi positivi, sono state associate correttamente all’individuo esatto. Ciò vuol dire che se abbassassimo la soglia θ esse sarebbero considerate facce correttamente riconosciute come tali ed associate al soggetto corrispondente.
- Riportiamo qui alcuni valori di distanze ottenute nel caso studio base (per le facce di test):

```
1787.6834875865757 distanza media dallo spazio delle facce
2570.6067670716193 distanza massima dallo spazio delle facce
1045.9991720143887 distanza minima dallo spazio delle facce
2350.0716815143014 distanza media dalle class face
3733.783082140672 distanza massima dalle class face
919.946068709828 distanza minima dalle class face
```

Figure 3: Distanze ottenute nel caso studio base.

Questi risultati suggeriscono e motivano i valori di θ e θ_k che verranno scelti nelle sezioni (2.3.3) e (2.3.4) per lo studio parametrico.

2.3.2 Studio al variare di var ($L = 240$, $\theta = 5000$, $\theta_k = 5000$)

Analizziamo i risultati al variare del parametro riguardante la varianza: minore è il valore di var (compreso tra 0 e 1), minore è il numero di autovettori “trattenuti” (dunque L'). Considereremo in particolare tre casi: $var = 0.99$ ($L' = 201$), $var = 0.95$ ($L' = 127$), $var = 0.76$ ($L' = 26$), $var = 0.50$ ($L' = 5$).

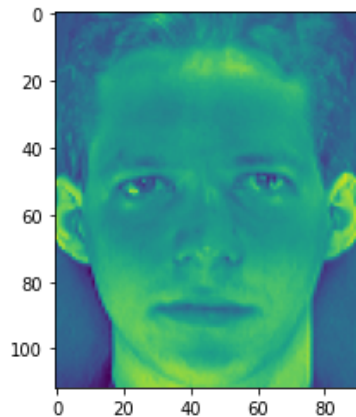


Figure 4: Faccia di training.

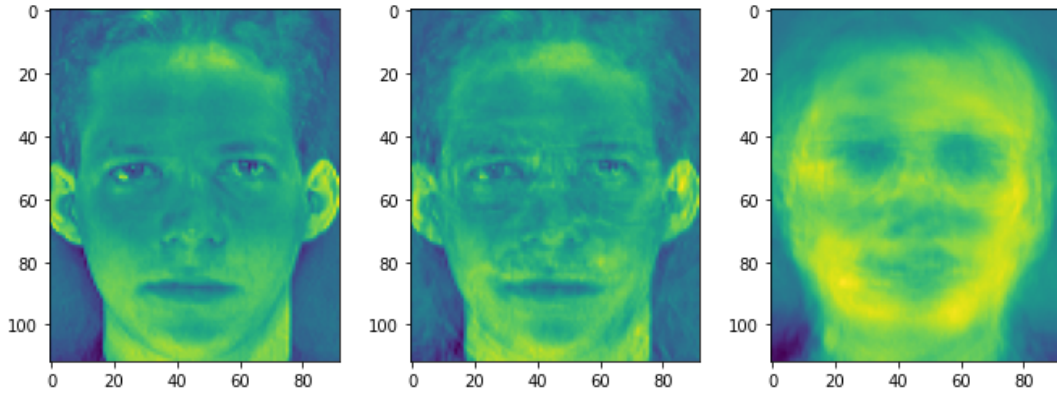


Figure 5: Le tre proiezioni della faccia considerata sullo spazio delle autofacce, al variare di var (da sinistra verso destra: $var = 0.99$, $var = 0.95$, $var = 0.50$).

Nella figura (5) sono riportate le tre diverse proiezioni della faccia rappresentata in (4) sullo spazio delle autofacce, al variare di var . Dalla figura (5) notiamo come al diminuire di var , diminuisca anche la qualità dell'approssimazione della faccia considerata nello spazio delle autofacce.

Riportiamo ora i risultati riguardanti il riconoscimento delle facce di test (che ci aspettiamo essere meno accurato al diminuire di var) e di quelle di training (che invece ci aspettiamo essere perfetto indipendentemente da var).

```
**** Testitamo le immagini di test ****

Sono stati riconosciuti con successo 160 volti, su 160 test
La frazione di volti riconosciuti con successo è: 1.0

Sono state determinate con successo 149 associazioni faccia-individuo, su 160 immagini riconosciute come facce
La frazione di volti correttamente associati, rispetto alle immagini riconosciute come facce è: 0.93125

Sono state determinate con successo 149 associazioni faccia-individuo, su 160 immagini riconosciute
La frazione di volti correttamente associati, rispetto alle immagini riconosciute è: 0.93125

Sono stati inoltre trovati:
  0 potenziali falsi positivi (su 0 effettivi falsi positivi),
  11 errori nelle associazioni
  0 immagini fuori dalla soglia di vicinanza agli individui noti
```

Figure 6: Riconoscimento ed associazioni delle facce di test nel caso $var = 0.99$.

```
**** Testitamo le immagini di test ****

Sono stati riconosciuti con successo 160 volti, su 160 test
La frazione di volti riconosciuti con successo è: 1.0

Sono state determinate con successo 147 associazioni faccia-individuo, su 160 immagini
riconosciute come facce
La frazione di volti correttamente associati, rispetto alle immagini riconosciute come facce è:
0.91875

Sono state determinate con successo 147 associazioni faccia-individuo, su 160 immagini
riconosciute
La frazione di volti correttamente associati, rispetto alle immagini riconosciute è: 0.91875

Sono stati inoltre trovati:
  0 potenziali falsi positivi (su 0 effettivi falsi positivi),
  13 errori nelle associazioni
  0 immagini fuori dalla soglia di vicinanza agli individui noti
```

Figure 7: Riconoscimento ed associazioni delle facce di test nel caso $var = 0.76$.


```

**** Testitamo le immagini di test ****

Sono stati riconosciuti con successo 160 volti, su 160 test
La frazione di volti riconosciuti con successo è: 1.0

Sono state determinate con successo 117 associazioni faccia-individuo, su 160 immagini riconosciute come facce
La frazione di volti correttamente associati, rispetto alle immagini riconosciute come facce è: 0.73125

Sono state determinate con successo 117 associazioni faccia-individuo, su 160 immagini riconosciute
La frazione di volti correttamente associati, rispetto alle immagini riconosciute è: 0.73125

Sono stati inoltre trovati:
  0 potenziali falsi positivi (su 0 effettivi falsi positivi),
  43 errori nelle associazioni
  0 immagini fuori dalla soglia di vicinanza agli individui noti

```

Figure 8: Riconoscimento ed associazioni delle facce di test nel caso $var = 0.50$.

Le figure (6), (7), (8) mostrano come al diminuire di var l'algoritmo riconosca comunque le facce come tali (dovuto al fatto che la soglia θ è sufficientemente elevata), ma nel caso $var = 0.50$ diminuisce il grado di accuratezza nelle associazioni tra faccia ed individuo. Sono stati omessi i risultati riguardanti $var = 0.95$ in quanto analoghi a quelli con $var = 0.99$. In definitiva, al diminuire della dimensionalità dello spazio delle autofacce, otteniamo un'approssimazione peggiore delle facce ed inoltre un minor grado di accuratezza nell'associazione tra le facce nuove ed il soggetto al quale corrispondono.

```

**** Testiamo le immagini usate nella fase di training ****

Sono stati riconosciuti con successo 240 volti su 240 test
La frazione di volti riconosciuti con successo è: 1.0

Sono state determinate con successo 240 associazioni faccia-individuo, su 240 immagini riconosciute come facce
La frazione di volti correttamente associati, rispetto alle immagini riconosciute come facce è: 1.0

Sono state determinate con successo 240 associazioni faccia-individuo, su 240 immagini riconosciute
La frazione di volti correttamente associati, rispetto alle immagini riconosciute è: 1.0

Sono stati inoltre trovati:
  0 potenziali falsi positivi (su 0 effettivi falsi positivi),
  0 errori nelle associazioni
  0 immagini fuori dalla soglia di vicinanza agli individui noti

```

Figure 9: Riconoscimento ed associazioni delle facce di training nel caso $var = 0.99$.

```

1.7720757509850835e-12: distanza media dalle class face
366.2969307746731      : distanza media dallo spazio delle facce

```

Figure 10: Distanze delle facce di training calcolate con $var = 0.99$.

```

1.6790599230422156e-12: distanza media dalle class face
889.809166368685       : distanza media dallo spazio delle facce

```

Figure 11: Distanze delle facce di training calcolate con $var = 0.95$.

Passando all'analisi del riconoscimento e dell'associazione delle facce di training, la figura (9) mostra come il riconoscimento e l'associazione tra le facce di training ed i soggetti nel dataset sia perfetta (come ci aspettavamo); inoltre, dalle figure (10) e (11) notiamo come l'associazione delle facce all'individuo corrispondente sia praticamente esente da errore (dell'ordine di 10^{-12}). Questo implica la possibilità di porre una soglia molto bassa θ_k nel caso delle facce di training; inoltre, confrontando (10) con (3) (i parametri sono i medesimi) notiamo come anche la distanza media dallo spazio delle facce sia molto minore nel caso delle facce di training (366 quella delle facce di training, 1788 quella delle facce di test).

2.3.3 Studio al variare di θ ($var = 0.99$, $L = 240$, $\theta_k = 5000$)

Analizziamo i risultati al variare del parametro θ : consideriamo i casi $\theta = 2000$, $\theta = 1500$ (si veda (2.3.1)).

```
**** Testitamo le immagini di test ****

Sono stati riconosciuti con successo 119 volti, su 160 test
La frazione di volti riconosciuti con successo è: 0.74375

Sono state determinate con successo 115 associazioni faccia-individuo, su 119 immagini riconosciute come facce
La frazione di volti correttamente associati, rispetto alle immagini riconosciute come facce è: 0.9663865546218487

Sono state determinate con successo 115 associazioni faccia-individuo, su 119 immagini riconosciute
La frazione di volti correttamente associati, rispetto alle immagini riconosciute è: 0.9663865546218487

Sono stati inoltre trovati:
  41 potenziali falsi positivi (su 7 effettivi falsi positivi),
  4 errori nelle associazioni
  0 immagini fuori dalla soglia di vicinanza agli individui noti
```

Figure 12: Riconoscimento delle facce di test nel caso $\theta = 2000$, $var = 0.99$.

```
**** Testitamo le immagini di test ****

Sono stati riconosciuti con successo 29 volti, su 160 test
La frazione di volti riconosciuti con successo è: 0.18125

Sono state determinate con successo 29 associazioni faccia-individuo, su 29 immagini riconosciute come facce
La frazione di volti correttamente associati, rispetto alle immagini riconosciute come facce è: 1.0

Sono state determinate con successo 29 associazioni faccia-individuo, su 29 immagini riconosciute
La frazione di volti correttamente associati, rispetto alle immagini riconosciute è: 1.0

Sono stati inoltre trovati:
  131 potenziali falsi positivi (su 11 effettivi falsi positivi),
  0 errori nelle associazioni
  0 immagini fuori dalla soglia di vicinanza agli individui noti
```

Figure 13: Riconoscimento delle facce di test nel caso $\theta = 1500$, $var = 0.99$.

I risultati mostrati in figura (12) e (13) rispecchiano quanto ci aspettiamo: al diminuire di θ , diminuisce la frazione di volti riconosciuti. Notiamo, inoltre, che la frazione di volti correttamente associati, tra quelli riconosciuti come tali, non si discosta molto da quella mostrata in figura (6). Questo conferma il fatto che il parametro θ impatta maggiormente sul riconoscimento dei volti e meno sulla corretta associazione, dove invece il parametro var gioca un ruolo maggiormente importante. Osserviamo comunque che la riduzione del parametro θ riduce il numero di errori nelle associazioni, dunque la scelta di un valore opportunamente “basso” garantisce maggiore sicurezza nelle associazioni, al costo di maggiore difficoltà nel riconoscere le immagini.

Osservazione (studio simultaneo al variare di var e θ): nella sezione (2.3.2) abbiamo evidenziato come la soglia θ sufficientemente elevata consentisse di riconoscere comunque tutti i volti al diminuire di var .

```

**** Testitamo le immagini di test ****

Sono stati riconosciuti con successo 110 volti, su 160 test
La frazione di volti riconosciuti con successo è: 0.6875

Sono state determinate con successo 106 associazioni faccia-individuo, su 110 immagini riconosciute
come facce
La frazione di volti correttamente associati, rispetto alle immagini riconosciute come facce è:
0.9636363636363636

Sono state determinate con successo 106 associazioni faccia-individuo, su 110 immagini riconosciute
La frazione di volti correttamente associati, rispetto alle immagini riconosciute è:
0.9636363636363636

Sono stati inoltre trovati:
  50 potenziali falsi positivi (su 7 effettivi falsi positivi),
  4 errori nelle associazioni
  0 immagini fuori dalla soglia di vicinanza agli individui noti

```

Figure 14: Riconoscimento delle facce di test nel caso $var = 0.95$, $\theta = 2000$, $\theta_k = 5000$.

Se confrontiamo le figure (14) e (12), notiamo come, a parità di θ , al diminuire del parametro var diminuisca la frazione di volti riconosciuti (nel caso specifico, si passa da $var = 0.99$ in (12) a $var = 0.95$ in (14)).

2.3.4 Studio al variare di θ_k ($var = 0.99$, $L = 240$, $\theta = 5000$)

Analizziamo i risultati al variare del parametro θ_k : consideriamo i casi $\theta_k = 3000$, $\theta_k = 2000$ (si veda (2.3.1)).

```

**** Testitamo le immagini di test ****

Sono stati riconosciuti con successo 160 volti, su 160 test
La frazione di volti riconosciuti con successo è: 1.0

Sono state determinate con successo 127 associazioni faccia-individuo, su 160 immagini riconosciute come facce
La frazione di volti correttamente associati, rispetto alle immagini riconosciute come facce è: 0.79375

Sono state determinate con successo 127 associazioni faccia-individuo, su 130 immagini riconosciute
La frazione di volti correttamente associati, rispetto alle immagini riconosciute è: 0.9769230769230769

Sono stati inoltre trovati:
  0 potenziali falsi positivi (su 0 effettivi falsi positivi),
  3 errori nelle associazioni
  30 immagini fuori dalla soglia di vicinanza agli individui noti

```

Figure 15: Riconoscimento delle facce di test nel caso $\theta_k = 3000$.

```

**** Testitamo le immagini di test ****

Sono stati riconosciuti con successo 160 volti, su 160 test
La frazione di volti riconosciuti con successo è: 1.0

Sono state determinate con successo 54 associazioni faccia-individuo, su
160 immagini riconosciute come facce
La frazione di volti correttamente associati, rispetto alle immagini
riconosciute come facce è: 0.3375

Sono state determinate con successo 54 associazioni faccia-individuo, su
54 immagini riconosciute
La frazione di volti correttamente associati, rispetto alle immagini
riconosciute è: 1.0

Sono stati inoltre trovati:
  0 potenziali falsi positivi (su 0 effettivi falsi positivi),
  0 errori nelle associazioni
  106 immagini fuori dalla soglia di vicinanza agli individui noti

```

Figure 16: Riconoscimento delle facce di test nel caso $\theta_k = 2000$

Le figure (15), (16) sono da confrontare con (6): notiamo come al diminuire di θ_k diminuisca sensibilmente la numero di volti correttamente associati, ma aumenti la frazione di volti correttamente associati rispetto a quelli che superano entrambe le soglie, come ci aspettavamo.

Osservazione (studio simultaneo al variare di θ e θ_k , commento sui falsi positivi): consideriamo nuovamente i risultati mostrati in (14) ($var = 0.95$, $\theta = 2000$, $\theta_k = 5000$) e confrontiamoli con quelli riportati nella seguente figura:

```

**** Testitamo le immagini di test ****

Sono stati riconosciuti con successo 110 volti, su 160 test
La frazione di volti riconosciuti con successo è: 0.6875

Sono state determinate con successo 105 associazioni faccia-individuo, su 110 immagini riconosciute come facce
La frazione di volti correttamente associati, rispetto alle immagini riconosciute come facce è: 0.9545454545454546

Sono state determinate con successo 105 associazioni faccia-individuo, su 108 immagini riconosciute
La frazione di volti correttamente associati, rispetto alle immagini riconosciute è: 0.9722222222222222

Sono stati inoltre trovati:
  30 potenziali falsi positivi (su 1 effettivi falsi positivi),
  3 errori nelle associazioni
  22 immagini fuori dalla soglia di vicinanza agli individui noti

```

Figure 17: $var = 0.95$, $\theta = 2000$, $\theta_k = 3000$.

Dal confronto tra (14) e (17), notiamo come l'abbassamento della soglia θ_k ci garantisca l'abbassamento del numero di falsi positivi. D'altra parte, in tal modo aumenta il numero di immagini lontane dalla classe di facce alla quale appartengono.

2.3.5 Studio al variare di L ($var = 0.99$, $\theta = 5000$, $\theta_k = 5000$)

Analizziamo i risultati al variare di L , ovvero del numero di facce di training (e, conseguentemente, del numero di facce di test). Considereremo i casi $L = 200$ (5 facce di training per ogni soggetto) e $L = 160$ (4 facce di training per ogni soggetto).

**** Testitamo le immagini di test ****

Sono stati riconosciuti con successo 200 volti, su 200 test
La frazione di volti riconosciuti con successo è: 1.0

Sono state determinate con successo 184 associazioni faccia-individuo,
su 200 immagini riconosciute come facce
La frazione di volti correttamente associati, rispetto alle immagini
riconosciute come facce è: 0.92

Sono state determinate con successo 184 associazioni faccia-individuo,
su 200 immagini riconosciute
La frazione di volti correttamente associati, rispetto alle immagini
riconosciute è: 0.92

Sono stati inoltre trovati:
0 potenziali falsi positivi (su 0 effettivi falsi positivi),
16 errori nelle associazioni
0 immagini fuori dalla soglia di vicinanza agli individui noti

Figure 18: Riconoscimento delle facce di test nel caso $L = 200$, $var = 0.99$.

**** Testitamo le immagini di test ****

Sono stati riconosciuti con successo 200 volti, su 200 test
La frazione di volti riconosciuti con successo è: 1.0

Sono state determinate con successo 143 associazioni faccia-individuo,
su 200 immagini riconosciute come facce
La frazione di volti correttamente associati, rispetto alle immagini
riconosciute come facce è: 0.715

Sono state determinate con successo 143 associazioni faccia-individuo,
su 200 immagini riconosciute
La frazione di volti correttamente associati, rispetto alle immagini
riconosciute è: 0.715

Sono stati inoltre trovati:
0 potenziali falsi positivi (su 0 effettivi falsi positivi),
57 errori nelle associazioni
0 immagini fuori dalla soglia di vicinanza agli individui noti

Figure 19: Riconoscimento delle facce di test nel caso $L = 200$, $var = 0.50$.

Le figure (18) e (19) sono da confrontarsi rispettivamente con (6) e (8): notiamo che in entrambi i casi la diminuzione delle facce di training porta ad una diminuzione dell'accuratezza nell'associazione tra le facce e gli individui.

**** Testitamo le immagini di test ****

Sono stati riconosciuti con successo 240 volti, su 240 test
La frazione di volti riconosciuti con successo è: 1.0

Sono state determinate con successo 209 associazioni faccia-individuo,
su 240 immagini riconosciute come facce
La frazione di volti correttamente associati, rispetto alle immagini
riconosciute come facce è: 0.8708333333333333

Sono state determinate con successo 209 associazioni faccia-individuo,
su 240 immagini riconosciute
La frazione di volti correttamente associati, rispetto alle immagini
riconosciute è: 0.8708333333333333

Sono stati inoltre trovati:
0 potenziali falsi positivi (su 0 effettivi falsi positivi),
31 errori nelle associazioni
0 immagini fuori dalla soglia di vicinanza agli individui noti

Figure 20: Riconoscimento delle facce di test nel caso $L = 160$, $var = 0.99$.

**** Testitamo le immagini di test ****

Sono stati riconosciuti con successo 240 volti, su 240 test
La frazione di volti riconosciuti con successo è: 1.0

Sono state determinate con successo 163 associazioni faccia-individuo,
su 240 immagini riconosciute come facce
La frazione di volti correttamente associati, rispetto alle immagini
riconosciute come facce è: 0.6791666666666667

Sono state determinate con successo 163 associazioni faccia-individuo,
su 240 immagini riconosciute
La frazione di volti correttamente associati, rispetto alle immagini
riconosciute è: 0.6791666666666667

Sono stati inoltre trovati:
0 potenziali falsi positivi (su 0 effettivi falsi positivi),
77 errori nelle associazioni
0 immagini fuori dalla soglia di vicinanza agli individui noti

Figure 21: Riconoscimento delle facce di test nel caso $L = 160$, $var = 0.95$.

Dalle figure (20), (21) (da confrontare rispettivamente con (6)-(18) e (8)-(19)) possiamo confermare il trend esposto in precedenza: al diminuire delle facce di training corrisponde una diminuzione nella frazione di volti correttamente associati.

2.4 Conclusioni

Dallo studio parametrico condotto, deduciamo le seguenti conclusioni:

- il numero delle facce di training ed il parametro var influiscono sulla corretta associazione in fase di test: al loro diminuire decresce l'accuratezza nelle associazioni; ciò è giustificabile dal punto di vista teorico ricordando che il numero di facce di training e var sono legati alla dimensionalità dello spazio delle autofacce, e minore è quest'ultima, peggiore è l'approssimazione delle facce mediante la loro proiezione su tale spazio;

- la soglia θ ricopre invece un ruolo maggiormente importante sul riconoscimento delle nuove facce come tali; essa è legata alla soglia θ_k dal fenomeno dei falsi positivi e della presenza di eventuali facce non riconosciute come tali ma comunque correttamente associate. In definitiva, possiamo affermare che questi due parametri sono strettamente legati, dunque interagiscono reciprocamente influenzando i risultati forniti dall'algoritmo;
- nella presente relazione è stato perlopiù condotto uno studio parametrico basato sulla variazione di un singolo valore alla volta; tuttavia, come abbiamo visto per $var-\theta$ e $\theta-\theta_k$, i parametri interagiscono tra di loro e la variazione dell'uno può portare a risultati diversi a seconda del valore “relativo” degli altri rispetto a tale parametro.

References

- [Mat91a] Alex Pentland Matthew Turk. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (1991) (cit. on p. 7).
- [Mat91b] Alex Pentland Matthew Turk. In: *Journal of Cognitive Neuroscience* 3.1 (1991), pp. 71–86. URL: <https://doi.org/10.1162/jocn.1991.3.1.71> (cit. on pp. 4, 7).
- [Nic23] Prof. Fabio Nicola. *Note del corso “Analisi tempo-frequenza e multiscala”*. 2023 (cit. on p. 2).