



DISCLAIMER

Il materiale contenuto nel drive è stato raccolto e richiesto tramite autorizzazione ai ragazzi frequentanti il corso di studi di Informatica dell'Università degli Studi di Salerno. Gli appunti e gli esercizi nascono da un uso e consumo degli autori che li hanno creati e risistemati per tanto non ci assumiamo la responsabilità di eventuali mancanze o difetti all'interno del materiale pubblicato.

Il materiale sarà modificato aggiungendo il logo dell'associazione, in tal caso questo possa recare problemi ad alcuni autori di materiale pubblicato, tale persona può contattarci in privato ed elimineremo o modificheremo il materiale in base alle sue preferenze.

Ringraziamo eventuali segnalazioni di errori così da poter modificare e fornire il miglior materiale possibile a supporto degli studenti.



CoScienze
Associazione

INTRODUZIONE

Image Intensity: energia luminosa emessa da una determinata porzione dell'immagine. Dipende dal dispositivo che si utilizza.

Image Brightness: come ci appare un'area di un'immagine (una soggettiva apparenza di un'area unitaria di un'immagine). Dipende dal contesto ed è soggettiva.

Contrasto: Il contrasto di un punto immagine è dato dalla differenza tra l'intensità del punto che si sta considerando e l'intensità dei punti vicini.

Istogramma: L'istogramma è una rappresentazione statistica dell'immagine.

Histogram Equalization: è il processo che serve per appiattire l'istogramma in modo tale da rendere i livelli di grigi tutti uguali. Questo corrisponde a rendere tutti i colori della stessa frequenza e lo si fa per avere un'immagine ideale cioè molto più leggibile e per migliorare il contrasto.

INTRODUZIONE BIOMETRIA

Biometria:

- Una caratteristica fisica: iride, impronte digitali, forma della mano, ...
- Una caratteristica comportamentale: firma, voce,

Registrazione di un soggetto (Enrollment)

- Si associa un insieme di caratteristiche all'identità di un soggetto
- Raccolta dei dati ed estrazione delle caratteristiche
- Il modello estratto (template) è memorizzato in un database o su un supporto portatile (smart card)
- Questo processo può essere effettuato singolarmente o in blocco (batch enrollment)

Verifica e Riconoscimento

Verifica: Confronto Uno a Uno. Conferma l'identità dichiarata da un individuo.

Riconoscimento: Confronto Uno a Molti. Stabilisce l'identità di un soggetto a partire da un insieme di persone registrate

Proprietà di una caratteristica biometrica

- **Universalità:** ogni persona deve possedere tale biometria.
- **Distintività (Unicità):** due persone devono essere sufficientemente distinguibili in base a tale biometria.
- **Permanenza (Stabilità):** la biometria deve rimanere invariante (rispetto ai termini di confronto) nel tempo.
- **Misurabilità:** essa deve essere misurabile quantitativamente.
- **Performance (Efficacia/Efficienza):** l'acquisizione, l'estrazione delle caratteristiche e il confronto delle entità non devono essere troppo costose in termini di tempo, memoria ed efficienza del sistema;
- **Accettabilità:** la procedura di acquisizione deve essere tollerata da una vasta parte della popolazione;

- **Eludibilità:** la probabilità che il sistema venga aggirato o indotto in errore deve essere minimizzata.

Biometrics	Univer- sality	Unique- ness	Perma- nence	Collect- ability	Perfor- mance	Accept- ability	Circum- vention
Face	H	L	M	H	L	H	L
Fingerprint	M	H	H	M	H	M	H
Hand Geometry	M	M	M	H	M	M	M
Keystroke Dynamics	L	L	L	M	L	M	M
Hand vein	M	M	M	M	M	M	H
Iris	H	H	H	M	H	L	H
Retina	H	H	M	L	H	L	H
Signature	L	L	L	H	L	H	L
Voice	M	L	L	M	L	H	L
Facial Thermogram	H	H	L	H	M	H	H
DNA	H	H	H	L	H	L	L

H=High, M=Medium, L=Low

Tipologia di Utente:

- **Cooperativo:** è interesse dell'utente che il sistema lo riconosca (un impostore cerca di farsi riconoscere).
- **Non-Cooperativo:** l'utente è indifferente al processo di riconoscimento (un impostore cerca di non farsi riconoscere).
- **Abituato/non abituato:** frequenza con cui gli utenti del sistema si prestano al riconoscimento biometrico (più volte al giorno, settimanalmente, mensilmente, ecc.).
- **Consapevole/inconsapevole:** se l'utente è o no al corrente del processo di riconoscimento in atto.

Tipologia di Applicazione

- **On-line:** l'utente è in attesa finché il sistema non risponde (tempi brevi di risposta).
- **Off-line:** l'utente viene acquisito ma non attende una risposta immediata.

Variabilità delle Biometrie

- **Intra_Classe:** variabilità delle biometrie per un individuo
- **Inter_Classe:** similarità delle biometrie tra individui diversi

FILTRI

Per semplificare la trattazione del problema lavoreremo solo su immagini a toni di grigio. Le medesime operazioni descritte per tali immagini si estendono alle immagini RGB operando separatamente sui tre canali (piani) R, G e B e trattando ciascuno di essi come una immagine a toni di grigio indipendente dagli altri canali.

L'obiettivo dell'operazione di filtraggio è quello di partire dall'immagine originale ed eseguire delle trasformazioni per eliminare o rendere evidenti alcune informazioni.

$G(x,y) = T[f(x,y)]$; T è lineare se è omogenea e associativa

Le trasformazioni si classificano in:

- tecniche operanti nel dominio delle frequenze
- tecniche operanti nel dominio dei pixel o anche detto dominio spaziale
- tecniche ibride

Le tecniche si classificano in:

- + **Point operation:** ogni pixel della nuova immagine dipende solo dal pixel corrispondente dell'immagine di partenza.
- + **Local operation:** ogni pixel della nuova immagine dipende da un intorno regolare dei pixel dell'immagine di partenza.
- + **Object operation:** ogni pixel della nuova immagine dipende da un intorno non regolare (definito come oggetto), dei pixel nell'immagine di partenza.
- + **global operation:** ogni pixel della nuova immagine dipende da tutti i pixel dell'immagine di partenza.

Tecniche operanti nel dominio spaziale

- *Point Processing: la trasformazione T è definita nel punto (x,y) :*
 - *Negativo (T monotona decrescente): inversione della scala di grigi, tutti quelli che stanno a 255 vanno a 0 e viceversa*
 - *Binarizzazione (T monotona crescente): si definisce una certa soglia e tutti i valori al di sopra di essa assumono valore 1 e tutti quelli al di sotto valore 0*
 - *Contrast Stretching (T è monotona crescente): tenta di migliorare un'immagine allungando la gamma di valori di intensità che contiene per sfruttare appieno i valori possibili.*

- *Masks Processing: la trasformazione T è definita in un intorno del punto (x,y) . La maschera è una matrice $n \times m$, solitamente dispari così da avere un unico punto centrale che viene sovrapposto al pixel da trasformare. In ogni cella della maschera c'è un coefficiente maggiore di zero che definisce quanto il pixel dell'immagine di partenza sottostante deve influenzare la trasformazione:*
 - LowPass (passa basso): attenua o elimina le alte frequenze» Contorni e dettagli
 - Highpass (passa alto): attenua o elimina le basse frequenze» Contrasto e intensità
 - Bandpass (passa banda): attenua frequenze in una banda predefinita.

Smoothing Filter (fase di pre-processing): riducono velatura e rumore

*LowPass (corrispondenza nel dominio delle frequenze)

+Lineari

» **Filtro Media (aritmetica):** i coefficienti della maschera vengono posti tutti a 1 e il pixel centrale viene sostituito con la media dei pixel circostanti.

» **Filtro Media Adattiva:** viene introdotta un valore soglia S . Se la somma delle differenze, tra il pixel centrale ed i pixel nel suo intorno è minore di questo valore soglia S viene applicata la media altrimenti si conserva il valore originale.

» **Filtro Gaussiano (dominio della frequenza):** ogni pixel viene rimpiazzato dalla media pesata dei pixel in un suo intorno, secondo una funzione Gaussiana. Il peso dei coefficienti della maschera è inversamente proporzionale alla distanza dei pixel nell'intorno rispetto a quello centrale (i pixel a distanza maggiore di 3 moltiplicato la deviazione standard non hanno alcuna influenza sull'operazione di filtraggio). La somma dei pesi della matrice è uguale a 1 e il kernel assume un valore maggiore della metà.

+Non Lineari: i valori d'intensità nell'intorno di ogni punto vengano prima di tutto ordinati, quindi successivamente viene calcolata una statistica su di essi ed il pixel centrale è rimpiazzato con tale statistica.

» **Filtro Mediano:** i coefficienti della maschera vengono posti tutti a 1. I valori di intensità dei pixel della matrice vengono ordinati in modo crescente, il pixel di output è l'elemento di posizione centrale

» **Filtro Max:** Ogni valore di grigio, in questo caso, è invece sostituito dal massimo valore nella maschera. Le zone più chiare si espandono a danno di quelle scure, dettagli fini e scuri vengono cancellati. I coefficienti del filtro sono tali che l'unico coefficiente diverso da zero, e pari ad uno è quello relativo al massimo.

» **Filtro Min:** Ogni valore di grigio è sostituito dal minimo valore nella maschera. Dettagli fini e chiari vengono cancellati, mentre le zone più scure si espandono. I coefficienti del filtro sono tali che l'unico coefficiente diverso da zero, e pari ad uno è quello relativo al minimo.

» **Filtro Di Susan** (non lineare-adattivo): l'idea consiste nel calcolare la media dei valori di intensità utilizzando solo i pixel dell'area USAN, se un pixel appartiene alla stessa area del pixel centrale gli do un peso (i pesi vengono calcolati a runtime) più alto rispetto ad un altro pixel che non appartiene a questa area, indipendentemente dalla distanza (quanto il pixel centrale si trova sullo scalino peseremo tanto tutti i suoi intorni che si trovano sopra e peseremo poco tutti i suoi intorni che si trovano sotto, così da preservare lo scalino). In questo modo punti al di fuori di esso non vengono coinvolti e la struttura dell'immagine dovrebbe rimanere invariata. Per stabilire i pixel che fanno parte dell'USAN associata ad un nucleo si utilizza la funzione di similitudine di intensità.

$$c(r, r_0) = e^{-\left(\frac{I(r) - I(r_0)}{t}\right)^2}$$

Con: $I(r)$ valore di intensità dei pixel che stiamo considerando, $I(r_0)$ quello del nucleo, t è un valore di soglia sull'intensità, $c(r, r_0)$ rappresenta il peso che il pixel in esame ha nel calcolo della media. Se il valore di c è uguale a 1 allora i pixel sono simili, più questo valore tende a 0 meno saranno i pixel e quindi il peso sarà minore. Oltre al peso dovuto alla funzione $c(r, r_0)$ per effettuare il filtraggio, si utilizza una funzione di tipo gaussiano. Se il pixel a cui vogliamo assegnare il peso differisce molto dal kernel, la differenza è alta, l'esponenziale tende a 0, al contrario invece tenderà a 1. Se l'area dell'USAN è nulla il valore di intensità del pixel in esame è calcolato applicando il filtro mediano.

*HighPass (corrispondenza nel dominio delle frequenze)

» **Filtro Sharpening:** elaborazione digitale dell'immagine tesa a recuperare/aumentare la nitidezza di un'immagine. Evidenziano i dettagli fini

Media vs Mediana vs Gaussiano

- Media: Eccessiva velatura (blurring) al crescere della maschera e delle iterazioni con conseguente perdita dei dettagli. Preserva con la versione adattiva le forme scalini).
- Mediana: Riduzione del rumore (noise reduction) senza (eccessiva) perdita di dettagli (blurring limitato). Non preserva le forme (scalini).
- Gaussiano: preserva meglio i contorni e ha un effetto di velatura ridotto

Edge Detection

L'operazione di estrazione dei contorni di un'immagine risulta ancora oggi una operazione fondamentale per l'analisi delle immagini. A differenza dei filtri di smoothing che eliminano le strutture di altra frequenza presenti in un'immagine, i filtri di estrazione del contorno hanno lo scopo di esaltare le alte frequenze ed eliminare le basse.

Un edge è un'area di transizione significativa tra regioni omogenee (o con variazioni non significative).

- Punti Isolati: più semplice da individuare perché il valore di grigio è normalmente molto differente da quello dei suoi vicini
- Linee (Orizzontali, Verticali, Oblique)
- Contorni (insieme di punti con inizio e fine coincidenti): Il contorno di un oggetto rappresenta infatti la separazione tra l'oggetto e lo sfondo o tra l'oggetto ed altri oggetti, per cui la sua estrazione è molto spesso il primo passo verso l'individuazione dell'oggetto.

Inoltre, esistono diversi tipi di edge: a gradino, a rampa, a linea e a cuneo.

Un operatore matematico che più riesce ad enfatizzare la transizione di intensità del colore è la derivata. La derivata prima e la derivata seconda sono significativamente diverse da 0 soltanto in corrispondenza di queste transizioni. In particolare, la

derivata prima è positiva in corrispondenza di una transizione scuro chiaro, e negativa nel caso opposto, mentre è nulla nella zona con livello di grigio costante. Mentre, la derivata seconda è positiva in prossimità della parte scura del contorno, negativa nella parte chiara, nulla nella zona di livello di grigio costante ed effettua un passaggio per lo zero (zero crossing) esattamente in corrispondenza alle transizioni. Una variazione significativa significa un massimo reale, mentre non significativa indica un massimo locale che quindi deve essere trascurata. Nei filtri edge la somma i posi possono essere anche negativi, ma la loro somma deve essere uguale a 0, così abbiamo la garanzia che nelle zone uniformi il filtro dà come risultato 0 e quindi siamo in una zona classificabile come non edge.

Esistono varie tecniche per l'individuazione di un edge:

Operatore Gradiente: il modo in cui i bordi di un'immagine sono orientati, non necessariamente coincide con gli assi coordinati. Ne consegue la necessità di calcolare le derivate direzionali dell'immagine f per ogni punto (x,y) nella direzione di massima discontinuità dei livelli di grigio (massimo locale della derivata di f). L'operatore che presenta queste caratteristiche è il gradiente. Discretizzandone direttamente il valore, e considerando che nelle immagini la minima distanza è di un pixel, è possibile esprimere il gradiente mediante le due:

-1	1	-1	0
0	0	1	0

Operatore di Roberts: ottenuto ruotando le due maschere, di dimensione pari, in modo che siano dirette una a 45° e l'altra a 135° costo computazionale minore ma sensibile al rumore

0	1	1	0
-1	0	0	-1

Operatore di Sobel: convoluzione fra una funzione di smoothing lineare, in genere gaussiana, e la funzione di derivazione. Utilizza matrici 3×3 . Costo computazionale più alto, ma meno sensibile al rumore

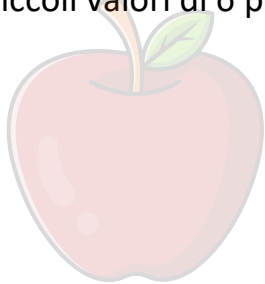
Operatore di Prewitt: variante dell'operatore di Sobel che al posto di un filtro gaussiano usa un box-filter. Più rumoroso rispetto a Sobel e non isotropo.

Operatore Laplaciano(isotropico): sfrutta il fatto che in corrispondenza dei massimi e minimi del gradiente la derivata seconda attraversa lo zero. Il problema è che l'operatore di derivazione in sé amplifica il rumore, è ciò è ancora più accentuato nella derivata seconda. Ha il vantaggio che lo spessore dei bordi rilevati tipicamente è di un solo pixel, inoltre tende a formare contorni chiusi, cosa che avviene con difficoltà con i metodi basati sul gradiente, in quanto non vengono considerati i massimi assoluti, ma quelli locali

Operatore Canny: consiste in tre fasi:

1. Esaltazione dei contorni: rimuovere il rumore mediante un filtro gaussiano, utilizzando una maschera di dimensioni opportune.
2. Soppressione dei non massimi: vengono soppressi i punti che non sono edge. Vengono eliminati i pixel che non sono massimi locali rispetto all'orientazione del gradiente.
3. Selezione finale degli edge: al fine di selezionare solo gli edge significativi, evitando allo stesso tempo la frammentazione, si fissano due valori di soglia $t_1 > t_2$ per ciascun punto avente valore superiore a t_1 , sono considerati edge, tutti i punti compresi tra t_1 e t_2 . sono considerati edge solo se sono contigui ad un edge. Grazie all'isteresi la soglia superiore serve a localizzare solo le strutture significative dell'immagine, mentre quella inferiore ha lo scopo di connetterle

Tempi di calcolo molto elevati dovuti principalmente alla seconda fase. Accentua le alte frequenze. Alti valori di σ permettono di trovare edge a scale più grandi piccoli valori di σ permettono di scovare i dettagli più fini.



CoScienze
Associazione

FURIER

-TRASFORMATA

In alcune situazioni può essere conveniente pagare prezzo per trasformare il segnale in un altro dominio e operare poi in questo dominio, dove l'analisi risulta semplificata. Un sistema lineare T prende in input (un segnale) $f(t)$ e produce in output $g(t)$, può essere vista come una black box. Un sistema lineare deve soddisfare due condizioni:

Homogeneity: $T\{a f[n]\} = a T\{f[n]\}$ **Additivity:** $T\{f_1[n] + f_2[n]\} = T\{f_1[n]\} + T\{f_2[n]\}$

In oltre un sistema lineare è definito *shift invariante* (invarianza al tempo) quando il suo comportamento nel tempo rimane immutato a meno di una traslazione temporale (la risposta è identica ma traslata di un alfa nel tempo). La *convoluzione* è un'operazione tra due funzioni di una variabile che consiste nell'integrare il prodotto tra la prima e la seconda traslata di un certo valore x . L'integrale è fondamentalmente un'area e questo integrale rappresenta l'intersezione tra le due funzioni al variare di x . Il *Delta* di *Dirac* o impulso unitario è una funzione che vale uno in un unico punto e zero in tutti gli altri [detta $\delta(n-k)$, che vale uno quando $n=k$, e zero in ciascun altro punto]. È importante per rappresentare qualsiasi funzione $f(n)$ come somma lineare di impulsi scalati e shiftati nel tempo, perché qualsiasi funzione $f(n)$ può essere descritta come la sommatoria tra $-\infty$ e $+\infty$ di j con $f(j)$ per delta di $n-j$, dove scegliamo delta $n-j = 1$ solo dove la funzione esiste. Vogliamo porci l'obiettivo di costruire un filtro standardizzato, il più possibile indipendente dalla funzione di input f , quindi invece di costruire T sulla funzione di input lo costruisco sul delta, mi prendo la risposta (di T con il delta in input) e la metto a convoluzione con f , per le proprietà di linearità e shift invarianza del nostro sistema, ottengo così lo stesso risultato se avessi costruito T sulla base di f , così ho il vantaggio di aver progettato T non su f ma su delta che risulta molto più semplice.

Vogliamo fare ancora meglio, in termini di standardizzazione e semplicità, progettando un T che sia il più possibile indipendente dall'input e che si comporti sempre allo stesso modo, vogliamo che sia lineare e shift invariante. Lo possiamo costruire nel dominio della frequenza (trigonometria). Questo è possibile perché c'è corrispondenza biunivoca tra i due domini senza perdita di informazione, se abbiamo un segnale discreto con n punti avremo un segnale nel dominio delle frequenze con n frequenza. Fourier dice che tutte le funzioni periodiche possono essere descritte come somme pesate di seni e coseni a frequenze variabili. La

combinazione lineare di un funzione seno e coseno inoltre, può essere espressa attraverso la semplice funzione seno shiftata e di cui abbiamo cambiato l'ampiezza. Quando applico Furier quello che devo fare è trovare le ampiezze che dobbiamo assegnare alle frequenze che andrò a selezionare. Individuo prima la portante, che avrà bassa frequenza e un'ampiezza grande, poi aggiungo frequenze sempre più alte, con ampiezza bassa per rappresentare i dettagli. Ne consegue che ogni funzione può essere espressa come combinazione lineare di funzioni sinusoidali scalate (modificando l'ampiezza), e shiftate (modificando la fase) che equivale alla formula

$$f(x) = \sum_{\omega} F(\omega) e^{i2\pi\omega x} \quad F(\omega) = \int_x f(x) e^{-i2\pi\omega x} dx$$

Per tornare nuovamente al dominio dei pixel invece applico la anti trasformata

$$f(x) = \int_{\omega} F(\omega) e^{i2\pi\omega x} d\omega$$

Possiamo avere una rappresentazione visiva della trasformata di Furier, tramite lo spettro di Furier, il quale rappresenta le ampiezze delle frequenze scelte. Notiamo che le ampiezze sono alte vicino allo 0, e servono per approssimare le caratteristiche generali del segnale, man mano che le frequenze aumentano le ampiezze diminuiscono, il segnale deve andare asintoticamente verso zero. In caso di rumore non è così. Notiamo dei bagliori anche delle zone periferiche, che rappresentano il rumore.

A seconda delle informazioni che voglio conservare, stabilisco che filtro usare. Fisso D_0 come *soglia* e uso:

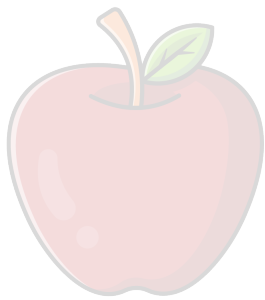
- **Passa basso:** per mantenere le basse frequenze ed eliminare le alte (ed il rumore). Conseguenza: sfocatura → blurring dell'immagine. Prendo ciò che è sotto la soglia.
- **Passa alto:** per mantenere le alte frequenze ed eliminare le basse. Conseguenza: risalto i dettagli (frequenze più alte) dell'immagine. Prendo ciò che è sopra la soglia.
- **Passa banda:** escludo le frequenze al di sotto e al di sopra di D_0

$$H(u,v) = \begin{cases} 1 & D(u,v) \leq D_0 \\ 0 & D(u,v) > D_0 \end{cases}$$

$$H(u,v) = \begin{cases} 0 & D(u,v) \leq D_0 \\ 1 & D(u,v) > D_0 \end{cases}$$

$$H(u,v) = \begin{cases} 0 & D(u,v) \leq D_0 - \frac{W}{2} \\ 1 & D_0 - \frac{W}{2} \leq D(u,v) \leq D_0 + \frac{W}{2} \\ 0 & D(u,v) > D_0 + \frac{W}{2} \end{cases}$$

Data un'immagine, eventualmente facciamo un pre-processing, poi facciamo la trasformata, ci costa $n \log(n)$, che ci restituisce le ampiezze per le frequenze selezionate, 0 per quelle non selezionate, ottengo l'insieme delle ampiezze poi effettuo l'operazione di filtro, che avrà anche esso un certo costo, ci sarà un filtro $H(u,v)$ molto banale sia perché è facile costruire sia perché l'operazione che andrò implementare è molto vantaggiosa, cioè una moltiplicazione tra H e F , effettuo l'anti trasformata, ci costa $n \log(n)$, e tengo così l'immagine filtrata. Conviene lavorare nel filtro delle frequenze perché: la trasformazione T è molto semplice, si tratta di una moltiplicazione tra la funzione $F(u,v)$ (F è un vettore dove ogni elemento è il valore moltiplicativo della frequenza) e, il filtro $H(u,v)$ che è anche esso un vettore di zeri in corrispondenza delle ampiezze che vogliamo eliminare e di uni in corrispondenza delle ampiezze che vogliamo mantenere. Inoltre, creare un filtro nel dominio spazio-tempo per un'immagine molto rumorosa è davvero molto complesso. Lo spettro di Fourier è shift invariante, se modifico l'immagine lo spettro rimane uguale, per apprezzare questo cambiamento dovrei sfruttare anche la componente fase.



CoScienze
Associazione

-CAMPIONAMENTO

Il campionamento è una tecnica per trasformare un segnale dal dominio analogico al dominio digitale. Durante la trasformazione sicuramente ci sarà perdita di informazioni perché stiamo passando da un dominio continuo nel tempo in un dominio discreto. Il problema principale sta nello scegliere un intervallo di campionamento T opportuno, perché se prendessi un T molto piccolo potrei catturare molte informazioni inutili, se lo prendessi troppo grande potrei perdere informazioni importanti, e inoltre risulta complicato scegliere anche l'istante esatto in cui campionare il segnale. Potrei scegliere un passo di campionamento adattivo che utilizza frequenze di campionamento diverse, ma dovrei comunque conoscere a priori come è fatto il segnale, cosa non possibile.

Nel dominio dello spazio-tempo per digitalizzare un segnale, mando un treno di impulsi, che intercettano il segnale, a distanza T e li moltiplico.

Se faccio la trasformata di Fourier di questo segnale spostandomi nel dominio delle frequenze ottengo invece un treno di impulsi ma a distanza $1/T$ che andrò a convolvere con lo spettro di Fourier del segnale (perché se nel dominio dei pixel faccio la moltiplicazione del dominio delle frequenze faccio la convoluzione e viceversa) ottenendo così una gaussiana che si replicherà tante volte in base a $1/T$. Più piccolo sarà T nel dominio dello spazio-tempo, tanto più distanti saranno le gaussiane nel dominio delle frequenze, più sarà grande T invece, tanto più le gaussiane saranno vicine. Se le gaussiane arrivano a sovrapporsi (nelle code ci sono le alte frequenze quindi i dettagli delle immagini) perdo delle informazioni perché ho due valori per un'unica frequenza generando **aliasing**.

La soluzione sarebbe scegliere T in modo che le gaussiane siano contigue. Quindi come scegliere T ? se sono in presenza di un segnale limitato in banda, evento raro, devo scegliere la frequenza di campionamento deve essere il doppio della frequenza massima del segnale ($1/T > 2W_{\max}$), in questo modo sono sicuro che non perdo nessun informazioni dalla trasformazione da digitale ad analogico (teorema di Nyquist). Laddove abbiamo un segnale che non ha banda limitata, non abbiamo ω_{\max} , ma ω tende a ∞ io scelgo comunque un ω_{\max} oltre il quale non andare, perdendo una parte di informazioni, cioè quelle che risiedono nei dettagli molto fini.

PARSEVAL

Indipendentemente dal contesto, quando il recupero avviene sulla base di caratteristiche visuali, prendiamo l'immagine estraiano le caratteristiche e creiamo il vettore che rappresenta la base della nostra query. Quando si vuole fare indicizzazione di immagini lo spazio delle caratteristiche, è più tosto ampio.

Quando la dimensionalità cresce una delle problematiche è la riduzione della stessa. Quando è necessario fare interrogazioni massive all'interno di un database il match one-to-many è un overdispensioso. Una soluzione sarebbe costruire un indice per fare mappature in alberi di ricerca (b tree, b+ tree, alberi binari), cioè alberi in cui ad ogni livello si esclude una parte consistente del database. Altre volte anche questa scelta è poco vantaggiosa, per motivi di manutenzione su database che vengono aggiornati ripetutamente. L'idea è provare a ridurre la dimensionalità.

Una soluzione è quella di operare in un dominio in cui il taglio della dimensionalità è controllato, cioè posso prevedere quello che può succedere, che sono: la trasformata di Fourier, o la analisi delle componenti principali.

Nel caso di Fourier, applicando la trasformata sul vettore delle caratteristiche, se abbiamo un vettore n dimensionale nel dominio spazio-tempo, la dimensionalità non cambia (la sommatoria dei quadrati è uguale) e l'energia del vettore rimane la stessa. Se abbiamo due vettori delle caratteristiche nello spazio-tempo x e y , e X e Y sono i corrispettivi nel dominio delle frequenze la distanza dei vettori nel dominio spazio-tempo rimane invariata.

Il teorema del *Parseval* ci dice: se noi decidiamo nel dominio delle frequenze, di non considerare la distanza su tutte le caratteristiche e l'applichiamo solo su certe caratteristiche, le più significative che sono le prime k frequenze, la distanza tra x e y calcolata su k è sicuramente la minore uguale di quella calcolata su N .

Cosa succede? Non aggiungo *false dismissal* cioè fondamentalmente se nella dimensionalità piena io avevo delle immagini significative, queste non andranno perse perché la riduzione della dimensionalità ha fatto sì che tutte quelle che erano presenti precedentemente continuino ad esserlo. Questo perché non si sono allontanate, e quindi non ho perdita di informazioni. Per contro troveremo anche alcune che prima non c'erano, perché avendo ridotto la dimensionalità possono essersi avvicinate, quindi ho dei *false alarms* in più. Cioè, potrei trovare tutte le immagini di prima più la spazzatura, che può essere eliminata

FINGER-PRINTING

Le impronte digitali sono uniche per ciascuna persona. Ogni impronta possiede:

- **Alcuni tratti caratterizzanti** (global classification): archi, cerchi (cappi) e spire/spirali
- **Molte altre dette** (local classification) minuzie (singolarità)

La massima diversità fra le impronte digitali è quella riscontrabile tra individui appartenenti a razze diverse. Seguono in ordine decrescente di diversità le impronte di:

- soggetti della stessa razza ma senza alcuna parentela,
- padre e figlio (che hanno in comune metà dei geni),
- fratelli e/o sorelle,
- gemelli omozigoti.

Esistono due modalità fondamentali di acquisizione delle impronte digitali :

off-line: l'acquisizione off-line avviene in due fasi, i polpastrelli vengono prima passati su un tampone inchiostro e poi trasferiscono l'immagine delle impronte digitali tramite pressione su carta. Una successiva digitalizzazione dell'impronta su carta tramite scansione ottica o telecamera ad alta risoluzione conclude il processo di acquisizione.

live-scan: l'immagine digitale dell'impronta digitale è acquisita direttamente tramite contatto del polpastrello con un apposito sensore.

Esistono vari tipi di impronte:

visibili: formato dal contatto con materiale colorato.

plastica: impresse un materiale morbido.

latenti o invisibili: Causato dal trasferimento corporeo di oli e / o traspirazione presente sulle creste delle dita.

I principali parametri che caratterizzano l'immagine digitale delle impronte sono: risoluzione, area di acquisizione, numero di pixel, accuratezza geometrica, contrasto, distorsione geometrica.

I sensori di impronte digitali possono utilizzare le seguenti tecnologie:

- **FTIR Frustrated Total Internal Reflection (Ottica)**, in grado di rilevare le immagini di creste e vallate con due diversi livelli di messa a fuoco;
- **Capacitiva**, basata sul rilevamento delle microvariazioni di capacità (accumulo di carica elettrica) originate dalla struttura tridimensionale del polpastrello;
- **Piezoelettrica**, basata sul rilevamento dell'impronta tramite materiali in grado di convertire la variazione di pressione locale in una differenza di potenziale elettrico;

- **Termica**, basata sull'utilizzo di sensori termici in grado di rilevare l'immagine del polpastrello non in lunghezze d'onda visibili ma nella gamma dell'infrarosso.

Abbiamo tre livelli di riconoscimento di un'impronta digitale:

#Caratteristiche Globali (condizione necessaria ma non sufficiente): sono di natura generica, vengono utilizzate solo per classificare. Sono di tre tipi:

Loop: hanno due punti focali, i core o centro del ciclo, che corrisponde alla terminazione o apice dell'anello, e i delta che corrispondono ad una triangolazione o divisione delle creste. Sono circa il 65% di un'impronta.

Archi: non hanno né delta né core, sono circa il 5% di un'impronta.

Spirali: hanno due o più delta e un core, sono circa il 35% di un'impronta.

#Caratteristiche Locali: anche dette minuzie sono molto discriminatorie, vengono utilizzate solo per il riconoscimento all'interno della classe prodotta dalle caratteristiche globali. È necessaria una risoluzione di almeno 500dpi per individuarle e ne bastano 14-20 per riconoscere un individuo. Sono situate principalmente dentro le creste, è possibile identificarne circa 150. Vengono rappresentate tramite un grafo, la verifica (confronto tra grafi) è un problema NP per questo vengono utilizzate solo dopo la classificazione

Caratteristiche ultra-fini: sono dettagli intra cresta quali essenzialmente i pori per la sudorazione, la cui posizione e forma sono considerate estremamente distintive, ma è necessaria una risoluzione di almeno 1000dpi.

Vantaggi e Svantaggi

Vantaggi: universalità, unicità, permanenza, misurabilità, accettabilità, acquisizione attiva e passiva, con basso costo dei sensori

Svantaggi: efficienza, stabilità (abrasione, rughe, sporco, sudore, disidratazione), insidia (facile da duplicare)

FACE DETECTION AND RECOGNITION

Il riconoscimento del volto garantisce affidabilità e accettabilità. Può essere applicato in ambito forense, controllo degli imbarchi, identificazione di persone nella folla, per accedere a dispositivi personali.

Presenta alcuni problemi come:

- Variabilità intra-persona: dovuti alla postura, alle condizioni di luminosità, all'espressione (PIE), all'avanzamento dell'età del soggetto (A-PIE).
- Variabilità inter-persona: dovuta principalmente alla somiglianza tra due soggetti (sosia, gemelli)

Verifica: confronto **uno-a-uno**, per confermare l'identità dichiarata da un individuo

Riconoscimento: confronto **uno a molti**, per stabilire l'identità di un soggetto a partire da un insieme di persone registrate.

Gli step per il riconoscimento facciale sono i seguenti:

- a partire da un'immagine viene agganciato la testa
- Si estrare la regione contenente solamente il viso, non i capelli
- la regione estratta viene segmentata associando significato semantico alle regioni di interesse: occhi, naso, bocca.

FACE DETECTION

Il primo passo è localizzare un volto in un immagine. Data una singola immagine o un video, individuare la presenza di uno o più volti e localizzarne la posizione. È importante che la localizzazione sia indipendente rispetto la posizione, orientamento, grandezza, espressione, illuminazione. La **face localization** può essere:

Feature-based: tecnica basata sull'extrapolazione di caratteristiche note della struttura del volto come il colore o la forma delle labbra e degli occhi, il colore della pelle, oppure confrontando il volto con un modello, si differenziano in:

- **Low-level analysis: proprietà dei pixel:** edges, colore della pelle
- **Feature analysis: informazioni sulla geometria del volto:** forma delle labbra o degli occhi
- **Template matching:** l'immagine viene confrontata con un modello di un volto

#FACE COUNTER è un algoritmo feature based che analizza tutti i triangoli composti da una probabile bocca e due probabili occhi candidati. Ogni triangolo viene verificato controllando:

1. Variazioni di luminosità e media delle orientazioni del gradiente delle aree contenenti occhi e bocca
 2. Geometria e orientazione dei triangoli
 3. Presenza del contorno di un volto attorno al triangolo
 4. Viene assegnato uno score a ogni triangolo che soddisfa queste condizioni, considerando maggiormente orientazione e simmetria
 5. Viene selezionato il triangolo con lo score più alto
- **Image-based techniques:** tecnica che si riconduce al riconoscimento generico di pattern. L'obiettivo è di imparare a riconoscere un volto sulla base di numerosi esempi (*Subspace Method, Neural Networks, Hidden Markov Model, Support Vector Machine*).

VIOLA-JONES è un algoritmo Image-based per l'object localization di molto successo, e può essere applicato al face detection. L'algoritmo richiede di creare un classificatore che deve essere allenato usando molteplici esempi positivi, dell'oggetto da individuare, e alcuni esempi negativi che potrebbero causare errore. Viene sfruttato il fatto che tutti i volti umani condividono alcune proprietà simili, come:

- La regione degli occhi è più scura delle guance superiori.
- La regione del ponte del naso è più luminosa degli occhi.

Queste caratteristiche possono essere utilizzate per la localizzazione del naso e degli occhi. Per l'individuazione di queste aree discriminanti vengono utilizzate le Haar Features, maschere rettangolari, composti una o più zone bianche o zone nere. Al passaggio di queste maschere di dimensioni diverse sull'immagine viene calcolato il valore della feature come la differenza tra l'intensità dei pixel dell'area scura, e l'intensità dei pixel dell'area chiara, se questo valore supera una certa soglia, possiamo dire di essere in presenza di un probabile volto. Per calcolare in modo efficiente i valori delle Haar feature, l'algoritmo utilizza un'immagine ausiliaria chiamata immagine integrale. L'immagine integrale ha le stesse dimensioni dell'immagine originale, ma i valori d'intensità di ogni pixel sono calcolati come la somma del pixel in questione e di tutti i pixel alla sinistra e sopra di esso. Siccome non tutte le feature sono utili allo scopo, viene utilizzata una variante di AdaBoost sia per selezionare le feature sia per allenare il classificatore. L'idea base è quella di

costruire un classificatore più complesso come combinazione lineare di tanti classificatori più deboli così che ogni uno utilizzi una feature diversa, così da eliminare dall'immagine le porzioni che sicuramente non contengono l'oggetto che si sta cercando. L'algoritmo di Viola e Jones, può essere utilizzato per la detection real-time, ma i classificatori hanno bisogno di molto tempo per essere allenati

FACE RECOGNITION

Le fasi del riconoscimento di un volto si dividono in:

#ENROLLMENT

- Raccolta dei dati ed estrazione delle caratteristiche: si associa un insieme di caratteristiche all'identità di un soggetto
- Il modello estratto (template) è *memorizzato* in un database o su un supporto portatile (smart card)

Questo processo può essere effettuato singolarmente o in blocco (batch enrollment)

#TESTING

- Raccolta dei dati ed estrazione delle caratteristiche: si estrae il template dal volto del soggetto che richiede di essere autenticato
- Il modello estratto (template) è *cercato* in un database o su un supporto portatile (smart card)
- Un criterio di matching stabilisce se il soggetto deve essere autorizzato o no.

Esistono vari sistemi in letteratura

- Image Based: ICA, Neural Networks, Eigenfaces; (a partire dall'immagine viene effettuata la trasformazione in un dominio più efficiente e in quel dominio provano ad avere una rappresentazione che consenta una rappresentazione più efficiente ed efficace)
- Feature based: Elastic Graph Matching; (vengono estratte le caratteristiche direttamente dalle immagini,)
- Metodologia 3D: 3D Morphable Models
- Tecniche ibride: Fractals, Wavelets;

Una delle tecniche image based più utilizzate è **eigenfaces** o **analisi delle componenti principali (PCA)**. Lo scopo della tecnica è quello di ridurre il numero di variabili che descrivono un insieme di dati, limitando il più possibile la perdita di informazioni. Il problema è come di legare le immagini ad un sistema lineare di

variabili, come prima cosa bisogna capire se sono correlate, o meglio se tutte le variabili mi servono e se qualcuna è riconducibile ad altre, così da eliminare quelle ridondanti ottenendo un set minimo. A partire da questo insieme lo posso utilizzare così come è, oppure posso cancellare un altro po' di roba, perdendo un po' di informazione (l'informazione persa può essere quantificata), cerco cioè di individuare tra le variabili, indipendenti, scarsamente correlate, le più significative utilizzando una matrice di covarianza (*evidenzia la correlazione tra un variabile e tutte le altre*), le entry di questa matrice possono essere valori maggiore di 0 questo indica che le due variabili sono dipendenti più questo valore è alto più sono correlate, se il valore è minore 0 le due variabili sono inversamente proporzionate, mentre se il valore è uguale a 0 sono tra di loro indipendenti (ortogonali). Passando dal dominio dell'algebra lineare a quello delle immagini l'obiettivo è stabilire, a partire da un database con M immagini, quali sono le meno correlate, quelle cioè che fondamentalmente posso rappresentare tutte le altre. Partendo da un insieme M di immagini ci accontentiamo di scegliere le prime immagini che presentano il più alto grado di varianza cioè le meno correlate, se non riusciamo a trovare immagini con indice di correlazione 0, prendiamo le immagini che hanno un valore di correlazione più vicino allo 0. Per estrarre la parte comune, calcoliamo la media di tutte le immagini del database e la sottraiamo a tutte le immagini così che ci rimane la parte spuria, che ha più potere caratterizzante. Successivamente dobbiamo trovare un set di immagini che in combinazione lineare mi permettono di generare, con errore quadratico minimo, tutte le immagini della mia base dati (principio PCA). Ogni immagine (di dimensioni $n \times n$) viene rappresentata come vettore unidimensionale di dimensioni n^2 , da questo vettore sottraiamo la media, che rappresenta la parte comune, calcolato sull'intero data base ottenendo così i potenziali auto vettori su cui andremo ad agire. Proiettiamo questi vettori nella matrice di covarianza e iniziamo la scelta delle componenti principali partendo dal vettore meno correlato dagli altri. Partendo da questo vettore, prendo k vettori meno correlati dove k è una soglia che indica quanti vettore voglio prendere e alla fine ottengo che ogni immagine della base dati può essere rappresentata dalla combinazione lineare dei vettori che ho scelto. L'approssimazione dipende dal grado di correlazione intrinseco delle immagini, e dalle k immagini che ho deciso di utilizzare per rappresentare il mio data base, quante più immagini utilizzo tanto più aumenterà la mia precisione. Ma la precisione dipende anche delle immagini stesse, quanto più ho immagini fortemente scorrelate, tanto più aumenta la capacità descrittiva.

VANTAGGI

Fase di identificazione veloce

Se si conservano gli auto-vettori è possibile ricostruire l'informazione originaria.

SVANTAGGI

Fase di training lenta, se si aggiungono nuovi soggetti è necessario il re training

Le immagini del data base devono avere tutte la stessa dimensione e i soggetti devono avere tutti la stessa posa

Gli autovettori (che da ora chiameremo anche autofacce) associati agli autovalori più grandi, saranno quelli che caratterizzano in maggior parte la varianza del training set.

#RETI NEURALI

Un altro approccio è basato sulle reti neurali, si addestrano un certo numero di neuroni a riconoscere alcune parti del volto, ciascuno di questi neuroni fornisce un risultato parziale che viene passato ad un livello superiore di neuroni che effettua il processo di autenticazione.

VANTAGGI

Riducono l'ambiguità fra soggetti appartenenti a classi simili.

Con opportuni accorgimenti sono robuste alle occlusioni

SVANTAGGI

Richiedono molta immagine per il training, possibilmente tutti quelli dell'universo con cui mi sto confrontando

Overfitting: la rete ha la stessa dimensione dell'input.

Overtraining: se le immagini sono grandi il metodo converge lentamente

Dtabase size: quando il numero di soggetti cresce diventano inefficienti.

#SISTEMI BASATI SU GRAFI

L'idea è quella di associare un volto ad un grafo. Si individuano alcuni punti di riferimento e si costruisce un grafo, tenendo conto della posizione dei nodi e degli archi

VANTAGGI

Sono robusti rispetto alle variazioni di posa e di illuminazione

Non richiedono il retraining del sistema

SVANTAGGI

Il processo di estrazione delle caratteristiche è lento

Il confronto one-to-many(riconoscimento) è lento (confronto tra grafi è NP-Hard)

#TERMOGRAMMMA

L'immagine viene catturata con particolari telecamere che sono in grado di rilevare la temperatura superficiale del volto, successivamente si effettua una rappresentazione topologica dei vari colori tramite un albero o un grafo.

VANTAGGI

Sono robusti rispetto alle variazioni di illuminazione e tempo, ma non di calore

Utilizzabile con opportune precauzioni anche in ambienti esterni

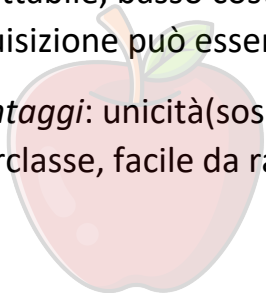
SVANTAGGI

Richiedono dispositivi di acquisizione costosi e sono sensibili al movimento del soggetto

Vantaggi e Svantaggi

Vantaggi: universalità, efficiente, efficace nella verifica(one-to-one), misurabile, accettabile, basso costo dei dispositivi di acquisizione nella maggior parte dei casi, acquisizione può essere attiva e passiva.

Svantaggi: unicità(sosia), non efficace nel riconoscimento(one-to-many), variabilità interclasse, facile da aggirare



CoScienze
Associazione

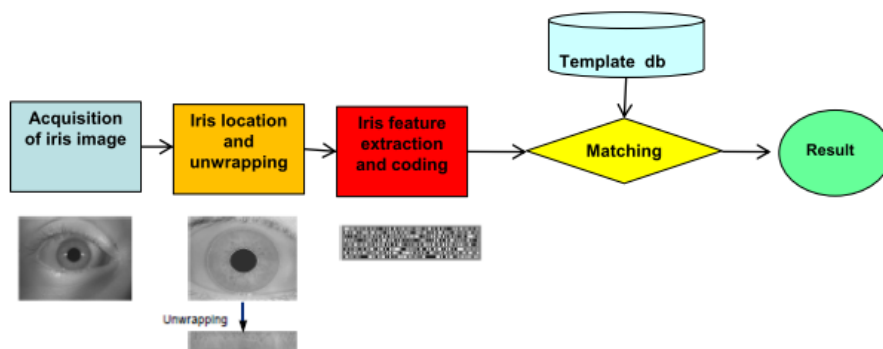
IRIS DETECTION

Le genesi dell'iride si forma quando si formano i capillari. Possiamo identificare fino a 300 minuzie, ma per certificare l'equivalenza tra due iridi ne servono circa 20-25. Alcune delle caratteristiche che vengono estratte sono i cosiddetti blobs (punti neri), ma anche il numero di striature la loro dimensione e posizione spaziale.

L'algoritmo di Daugman è stato il primo noto scientificamente in contesti non militari, che nel 2002 venne usato per attestare l'identità della monna lisa afghana. L'algoritmo di Daugman può essere utilizzato soltanto quando l'acquisizione dell'iride viene fatta ad una distanza molto limitata (50/60cm), tramite telecamere infrarossi (che annullano determinati particolari non necessari e permette di ottenere le informazioni significative per la classificazione utilizzando tecniche che lavorano in frequenza come gabor e wavlet). Il problema dell'iride è che si trova in un punto piuttosto sensibile, perché ci sono una serie di elementi di disturbo come la pupilla, perché può essere più o meno dilatata, le palpebre e le ciglia, perché possono coprire una parte dell'iride, e anche il disallineamento con la camera può generare qualche problema.

Il processo di analisi di un'iride è diviso in 4 parti:

- detection o segmentazione: cioè l'individuazione della parte utile su cui lavorare, fase molto complessa in contesti non controllati.
- normalizzazione: bisogna mappare la superficie utile della corona circolare che rappresenta effettivamente l'iride, in un vettore sempre della stessa dimensione, indipendentemente, appunto, dalla dimensione della corona circolare che può essere influenzata dalla distanza dal sensore e dalla presenza di fattori di noise.
- coding: estrazione dal vettore delle informazioni utili per caratterizzare
- matching



Daugman ha lavorato specialmente sulla parte **dell'iris detection** (location).

L'operatore è essenzialmente un edge detector circolare, che cerca un percorso circolare in cui sia massimizzata la variazione del valore dei pixel, variando il raggio r e il centro (x_0, y_0) , presentando un picco quando un cerchio candidato ha lo stesso centro e raggio dell'iride. L'operatore è usato sia per trovare il contorno della pupilla sia per rilevare il bordo esterno dell'iride. Per la localizzazione della linea superiore e inferiore delle palpebre viene adottato un procedimento simile a quello usato per l'iride solo che l'operatore è in questo caso ad arco e non circolare.

"L'algoritmo di Daugman consiste in un Operatore locale che combina in convoluzione, una gaussiana che fa ogni volta uno smoothing, con un operatore che cerca, a partire da un centro il cerchio più grande dove si posiziona la variazione massima dei pixel. Fondamentalmente Daugman si posiziona sulla parte utile di iride, tirata fuori dal segnale, e comincia a cercare, a partire da un punto il cerchio massimo che si ottiene a partire da quel punto. Quindi a partire da un punto di un'immagine di un'iride incomincio a creare dei cerchi concentrici fino a che non trovo la variazione massima dei pixel e la distanza massima dal centro. Il punto che massimizza questa situazione è quello centrato all'interno della pupilla, quindi io troverò prima il cerchio che identifica la pupilla e successivamente il cerchio che identifica l'iride cioè i due cerchi che massimizzano la variazione. Nel passaggio tra l'iride e la pupilla troverò la variazione massima tra il nero e il colore dell'iride successivamente trovo la variazione massima o con la parte bianca, detta sclera. Quindi attraverso una scansione sequenziale da sinistra verso destra e dall'alto verso il basso vado ogni volta, a partire da un punto, a generare il cerchio massimo fino a che non trovo chiaramente quello che viene massimizzato. Che dovrebbe essere massimizzato se l'immagine risulta essere conforme prima con la determina della pupilla e poi con la determina dell'iride e prenderò poi il massimo di tutti i massimi. Lo smoothing gaussiano conserva ad ogni iterazione solo le variazioni massime, e non le piccole, rendendo più facile il processo di identificazione dei due cerchi in cui si concentrano maggiormente le variazioni. L'operatore di smoothing gaussiano reiterato facilita l'identificazione delle due circonferenze massime che sono prodotte in maniera iterativa. Ogni volta si allarga, quando ha trovato in quel punto il cerchio massimo passa al punto successivo. Alla fine, io avrò una serie di candidati e prendo i candidati che hanno massimizzato il raggio e che hanno massimizzato il gradiente, cioè la variazione tra la zona sinistra e quella destra. I fattori di disturbo sono le riflessioni dovute alla presenza di luce esterna perché rappresentano variazioni molte alte bianche che vengono confuse con la sclera facendo fallire la localizzazione precisa dell'iride e della pupilla beccandone comunque una parte, quindi la probabilità che noi in quel caso riusciamo ad estrarre tutte le caratteristiche significative diventa più bassa." Prof. Nappi

Esiste una grande variabilità intra-classe dovuta alla dilatazione della pupilla (che può variare in base ai cambiamenti di luce), alla dimensione dell'iride (che può variare in base alla distanza di acquisizione dell'immagine), e alla rotazione dell'occhio. Per l'**unwrapping dell'iride** viene utilizzato il Rubber Sheet Model. Il modello mappa ogni punto dell'iride in un punto di coordinate polari (r, θ) , dove $r \in [0, 1]$ e $\theta \in [0, 2\pi]$. Il modello compensa la dilatazione della pupilla e la variabilità nella dimensione producendo una rappresentazione invariante rispetto a dimensione e traslazione nel sistema di coordinate polari, ma non gestisce la rotazione che è invece considerata in fase di matching traslando il template dell'iride fino a ottenere l'allineamento delle due iridi da confrontare.

Per l'**estrazione delle feature** vengono utilizzati i filtri di Gabor (nel sistema di coordinate polari), variante della trasformata di Fourier che agisce nel dominio dello

spazio anziché nel dominio del tempo. Per ogni elemento di coordinate (r_0, θ_0) è calcolata una coppia di bit (h_{Re}, h_{Im}) . Dall'intera immagine dell'iride vengono estratti un totale di 256 byte di informazione. In implementazioni successive sono stati aggiunti altrettanti byte di mascheratura che identificano le aree "valide" dell'iride.

Il confronto (**matching**) tra due iris code si basa sulla distanza di Hamming. La distanza di Hamming tra due stringhe di ugual lunghezza è il numero di posizioni nelle quali i simboli corrispondenti sono diversi. *In altri termini, la distanza di Hamming misura il numero di sostituzioni necessarie per convertire una stringa nell'altra, o, vista in altro modo, il numero minimo di errori che possono aver portato alla trasformazione di una stringa nell'altra. Se la distanza di Hamming è < 0.33 la possibilità che due codici appartengano a iridi differiti è 1 su 2.9million.*

Alcuni metodi a confronto

Matching: Hamming distance

$$HD = \frac{1}{N} \sum_{j=1}^N A_j \otimes B_j$$

Matching: Hamming distance with mask

$$HD = \frac{\| (codeA \otimes codeB) \cap maskA \cap maskB \|}{\| maskA \cap maskB \|}$$

	Lim et al.	Daugman	Tisse et al.	Wildes	Sanchez-Avila et al.
Coordinate	Polari	Polari	Polari	Polari	Cartesiane
Pre-processing	Rilevazione dei contorni; metodo di bisezione	Operatore integrale-differenziale	Rilevazione dei contorni; operatori integro-differenziali	Rilevazione dei contorni; trasformata di Hough	Operatori di massimo
Trasformata	Wavelet di Haar	Filtri bidimensionali di Gabor	Fase istantanea, fase emergente e trasformata di Hilbert	Decomposizione isotropica passa banda; filtri laplaciani della cascata	Wavelet discreta diadica
Dim. Vettore	87 bit	2048 bit	6144 bit	2048 bit	256 bit
Distanza	Euclidea	Hamming	Hamming	Correlazione normalizzata	Hamming, Euclidea, funzione di diversità
Brevetto	SI (Senex, Evermedia)	SI (Iritech, Iridian)	NO	NO	NO

Vantaggi e Svantaggi

Vantaggi: l'iride è visibile ma ben protetto, è time invariant ed è un tratto estremamente distintivo, la sua immagine può essere acquisita senza diretto contatto, acquisizione: near infrared e visible wavelengths.

Svantaggi: superficie molto limitata, una buona acquisizione richiede una distanza inferiore ad un metro per garantire una risoluzione sufficiente

SISTEMI MULTIMODALI

Una caratteristica fisica o comportamentale di un individuo per poter assumere un ruolo di chiave biometrica deve soddisfare quattro requisiti fondamentali:

- Universalità: parte dell'utenza può essere priva di questa caratteristica dalla nascita o in seguito ad incidenti.
- Permanenza: alcune caratteristiche biometriche non sono stabili nel tempo.
- Unicità e Distinguibilità: alcuni soggetti possono avere caratteristiche biometriche molto simili, o facilmente replicabili.
- Collezioneabilità: esistono, inoltre, i cosiddetti errori d'acquisizione, errori generati durante il processo di rilevazione della biometria che rendono difficoltosa l'acquisizione.

Esistono casi in cui una più di questi requisiti non sono rispettati da una biometria. Quindi sono state studiate tecniche di integrazione nel tentativo di sopperire ad alcune delle limitazioni imposte da un sistema unimodale. Così sono nati i sistemi multimodali, nei quali si cerca di fondere insieme più biometrie.

Scegliamo i sistemi multi-biometrici fondamentalmente per:

- ridurre la possibilità che il sistema viene ingannato da un eventuale usurpatore.
- avere maggior sicurezza sull'identità di una persona a meno che il risultato tra le due biometriche sia discordante.
- in contesti non controllati la possibilità di acquisire più biometriche nella fase di riconoscimento ci consente di acquisire più dati di una persona magari anche spuri però nell'aggregazione di questi dati ho maggior possibilità di riconoscerla.

Nello sviluppo di un sistema multimodale bisogna considerare diversi aspetti progettuali:

- La scelta delle biometrie
- Il livello del sistema a cui fare l'integrazione
- La metodologia utilizzata per l'integrazione
- Requisiti funzionali

Ci sono più tipi di sistemi multi-biometrici:

- Multiple algorithm: una biometrica processata con algoritmi diversi.
- Replaced instances: una biometria acquisita più volte nelle stesse condizioni
- Multiple sensor: Stessa biometrica acquisita con sensori diversi.
- Multimodal: più biometriche combinate insieme
- Multiple istances: estrarre caratteristiche diverse da una stessa biometrica

Un'altra classificazione che può essere:

- multimodale: sensori multipli per la stessa biometrica
- multi-biometrici: se utilizzano biometriche diverse
- multi-expert: se utilizzano rappresentazioni multiple o algoritmi differenziati per analizzare le biometriche

Il processo di autenticazione di una biometria prevede 4 step: acquisizione mediante sensori, estrazione delle feature, matching con il template a disposizione, e decisione. Nei sistemi multi-biometrici bisogna effettuare un ulteriore passaggio: la **fusione** delle varie biometrie acquisite.

La fusione può essere effettuata prima del matching a livello di:

- sensore: I dati acquisiti da sensori diversi possono essere elaborati e integrati per generare nuovi dati dai quali si estraggono poi le feature e utilizzate dall'intero sistema.
- feature: i template generati dalle singole biometrie sono combinati in un'unica entità di caratteristiche che è utilizzata nell'intero sistema

oppure dopo il matching a livello di:

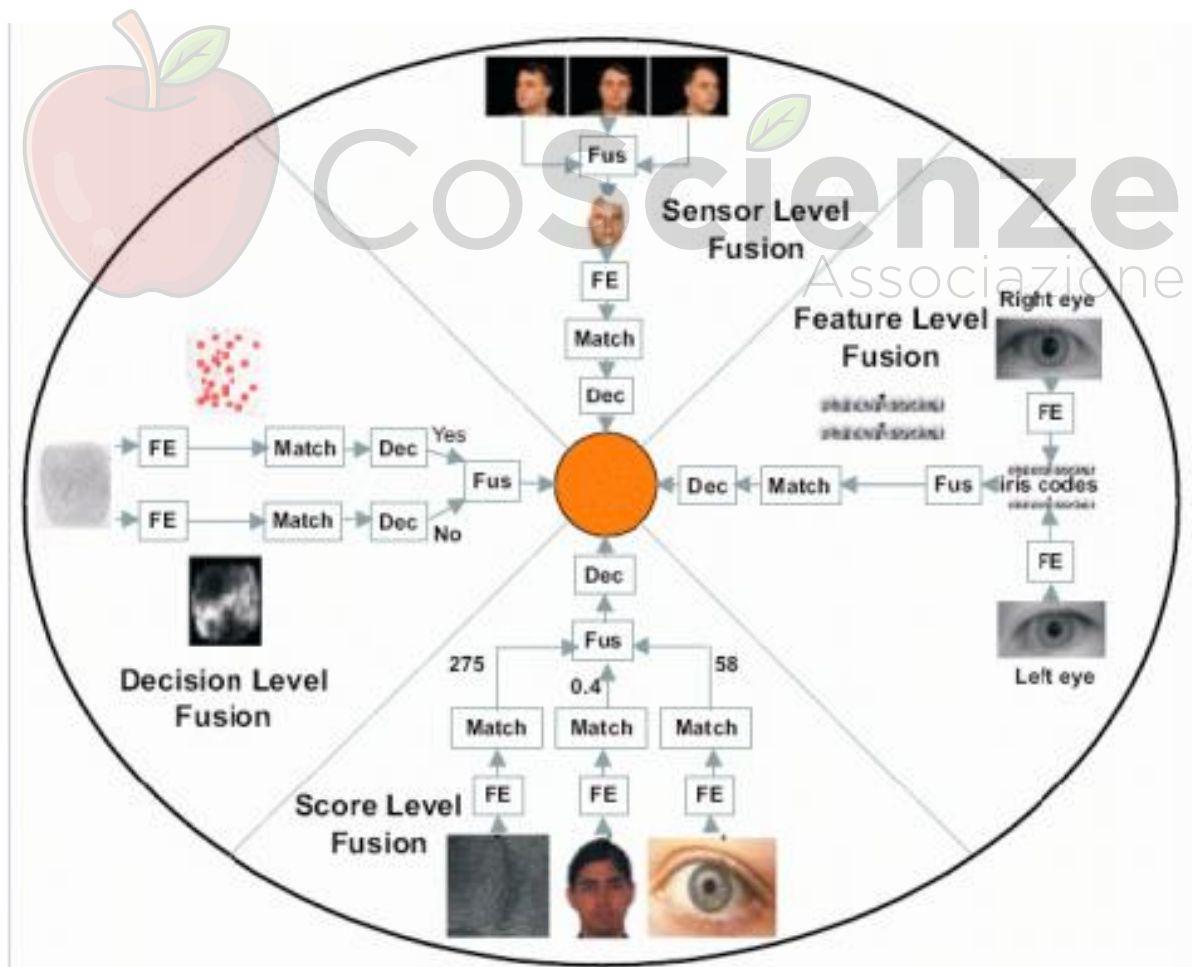
- decisione: le singole biometrie sono valutate indipendentemente dal sistema, su ognuna si prende una decisione ed il modulo di fusione esprime il responso finale valutando la congruenza dei risultati
- score: le singole biometrie vengono valutate singolarmente dal sistema ed ad ogni una di essa viene associato un score, questi vengono fusi per generare un unico valore finale su cui prendere poi una decisione.

Ritardare la fusione risulta sconveniente, perché il modulo decisionale riceve solo una minima parte delle informazioni acquisite e conosce solo l'esito del confronto delle singole biometrie, inoltre, comporta l'implementazione e quindi l'elaborazione di due sottosistemi ognuno dedicato alla specifica biometria.

Conviene fare uso di una strategia di fusione il prima possibile poiché si conserva una ricchezza di informazioni maggiore fornita dai dati acquisiti.

Esistono 5 tipi di scenari di fusione:

- Unica biometria e Multiplo sensore: sono combinate le informazioni dello stesso tratto biometrico acquisito però con due tipi di sensori diversi.
- Multipla biometria: sono combinate le caratteristiche di due biometrie diverse, un volto ed un'impronta.
- Unica biometria e Multipla unità: si combinano due elementi diversi appartenenti alla stessa classe biometrica (l'impronta di due dita diverse).
- Unica biometria e Multipla acquisizione: si combinano più sessioni di acquisizione dello stesso tratto biometrico (più impressioni della stessa impronta).
- Unica biometria e Multiplo algoritmo di confronto: la stessa acquisizione della biometria è valutata con due approcci diversi, due modalità di confronto delle impronte



PAROLE chiave

- Wavelet o ondicelle, analisi wavelet, e trasformata wavelet si riferiscono alla rappresentazione di un segnale mediante l'uso di una forma d'onda oscillante di lunghezza finita o a decadimento rapido (nota come wavelet madre). Questa forma d'onda è scalata e traslata per adattarsi al segnale in ingresso.
- Il filtro di Gabor è un filtro lineare (passa banda) la cui risposta all'impulso è definita da una funzione armonica moltiplicata per una funzione Gaussiana. Adatto per la texture analysis, edge detection, feature extraction, disparity estimation
- In matematica, il sistema di coordinate polari è un sistema di coordinate bidimensionale nel quale ogni punto del piano è identificato da un angolo e da una distanza da un punto fisso detto polo. Un sistema di coordinate polari è in corrispondenza biunivoca con un sistema di coordinate cartesiane, ossia ad un vettore di coordinate cartesiane ne corrisponde uno e uno solo in coordinate polari.
- I filtri adattativi sono una classe di filtri digitali in cui i coefficienti del filtro vengono fatti variare nel tempo in modo da ottimizzare le prestazioni del filtro stesso.
- isotropico, cioè invariante per rotazione, e quindi ha le caratteristiche del sistema visivo umano e risponde in egual modo alle variazioni di intensità in ogni direzione della maschera, senza quindi la necessità di dover utilizzare maschere multiple per calcolare la risposta più forte
- A zero-crossing is a point where the sign of a mathematical function changes
- Gradiente: Il gradiente è il vettore le cui componenti sono le derivate parziali nelle diverse direzioni. La direzione in cui l'intensità o il colore dell'immagine cambia
- autovettore di una funzione tra spazi vettoriali è un vettore non nullo la cui immagine è il vettore stesso moltiplicato per un numero (reale o complesso) detto autovalore.^[1] Se la funzione è lineare, gli autovettori aventi in comune lo stesso autovalore, insieme con il vettore nullo, formano uno spazio vettoriale, detto autospazio
- numero complesso è rappresentabile come una parte reale e una immaginaria, che hanno una corrispondenza importante con la trigonometria coseno è la parte reale il seno quella immaginaria
- esponenziale complesso è la sintesi lineare tra una componente seno e una coseno alla stessa frequenza.