

LEZIONE1

TEMPLATE

modello biometrico estratto da un soggetto e composto da un Feature Vector, permette di ricostruire i tratti biometrici originali. Il Template Aging è l'invecchiamento o il cambiamento in generale del soggetto, e di conseguenza del modello biometrico.

PROBE

è il Template di un soggetto sottomesso alla Recognition.

GALLERY

l'insieme dei template appartenenti ai soggetti registrati nel sistema.

ENROLLMENT

è il processo di cattura dei dati biometrici di un utente i quali verranno registrati sotto forma di template nel template archive.

RECOGNITION

Un probe esterno viene sottomesso al sistema il quale determinerà tramite dei matching se il probe appartiene ad un soggetto nella gallery.

Ci sono due modalità di funzionamento per quanto riguarda il recognition:

- **VERIFICATION (1:1)**

Il soggetto dichiara la sua identità e il sistema determinerà se effettivamente è lui o meno

- **IDENTIFICATION (1:N)**

Il sistema deve verificare l'identità di un soggetto in un insieme di template e l'utente non dichiara la propria identità.

L'identification può avvenire utilizzando due tipi di set:

- 1) **OPEN-SET**

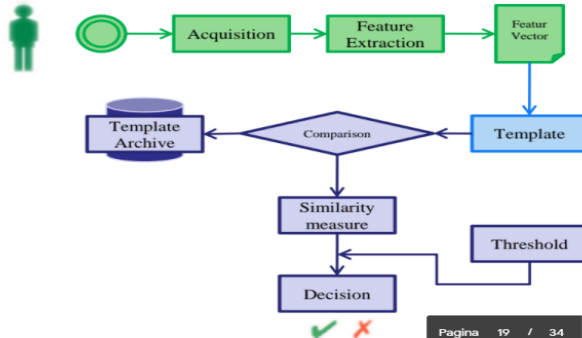
E' possibile che molti probes sottomessi al sistema non appartengono a nessun soggetto presente nel gallery. Il sistema deve prima capire se il soggetto è uno sconosciuto o no e poi nel caso verificarne l'identità.

- 2) **CLOSED-SET**

Tutti i probes sottomessi al sistema devono appartenere ad un soggetto nel gallery

Enrollment:

Capture and processing of user biometric data for use by system in subsequent authentication operations (gallery).



Recognition:

Capture and processing of user biometric data in order to render an authentication decision based on the outcome of a matching process of the stored to current template. (verification 1:1 identification 1:N)

Probe: each template which is submitted for recognition.

Gallery: the set of templates pertaining to enrolled subjects

PROPRIETÀ BIOMETRICHE

Acquisibile: deve essere facilmente acquisibile da un sistema biometrico artificiale;

Accettabilità: le persone non devono avere problemi durante l'acquisizione. Non deve essere un sistema invasivo.

Circumversion (Elusione): Misura la facilità con cui un tratto può essere imitato usando un artefatto o un sostituto.

Permanenza: non deve cambiare nel tempo (altrimenti non sarebbe un buon tratto);

Universalità: tutti devono avere questo tratto;

Unicità: ogni coppia di persone deve avercelo differente;

LEZIONE2



WIDE INTRA-CLASS VARIATION

il tratto biometrico potrebbe variare troppo all'interno di uno stesso soggetto; (pernacchia)

SMALL INTER-CLASS VARIATION

il tratto biometrico potrebbe essere troppo simile tra due differenti soggetti; (gemelli)

Possibili problemi:

Non-Universalità: Violazione delle caratteristiche per un buon tratto; (impronte digitali delle persone anziane)

Attacchi al sistema: da altri soggetti (Spoofing) o da sé stessi (Camouflage).

POSSIBILI ERRORI

Un soggetto viene accettato se il similarity score ottenuto dal matching con i template nel gallery supera un certo threshold.

Possono però verificarsi 4 possibili casi:

- 1) **GA - GENUINE ACCEPTANCE:** L'identità è vera e il soggetto viene accettato
- 2) **FR - FALSE REJECTION:** L'identità è vera ma il soggetto viene rifiutato
- 3) **GR - GENUINE REJECT:** L'identità di un impostore viene rifiutata
- 4) **FA - FALSE ACCEPTANCE:** Un impostore viene accettato

Come comparare dei sistemi:

- **FAR - FALSE ACCEPTANCE RATE**

E' la percentuale di volte in cui durante l'identificazione avviene una False Acceptance.

$$FAR(t) = \frac{|\{p_j : s_{xj} \geq t \wedge id(g_x) \neq id(p_j)\}|}{|\{p_j : id(g_x) \neq id(p_j)\}|}$$

$$g_x = topMatch(p_j, i)$$

$$s_{xj} = s(g_x, p_j)$$

$$Scenario\ 1 : \forall p_j \in P_G \cup P_N \quad \forall i \in I$$

$$Scenario\ 2 : \forall p_j \in P_N \quad \forall i \in I'$$

Numeratore: numero di volte in cui la claimed identity viene accettata nonostante l'identità sia falsa

Denominatore: numero totale di volte in cui il probe aveva l'identità dichiarata diversa da quella reale. / è la cardinalità dell'insieme di persone che non dobbiamo considerare correttamente.

- **FRR - FALSE REJECTION RATE**

E' la percentuale di volte in cui durante l'identificazione avviene una False Rejection.

$$FRR(t) = \frac{|\{p_j : s_{xj} \leq t, id(g_x) = id(p_j)\}|}{|\{p_j : id(g_x) = id(p_j)\}|}$$

$$\begin{aligned} g_x &= topMatch(p_j, id(p_j)) \\ s_{xj} &= s(g_x, p_j) \\ \forall p_j &\in P_G \end{aligned}$$

Numeratore: numero di volte in cui avviene la false reject.

Denominatore: numero totale di volte in cui il probe aveva l'identità dichiarata uguale a quella reale. / è la cardinalità dell'insieme di persone che dobbiamo considerare correttamente.

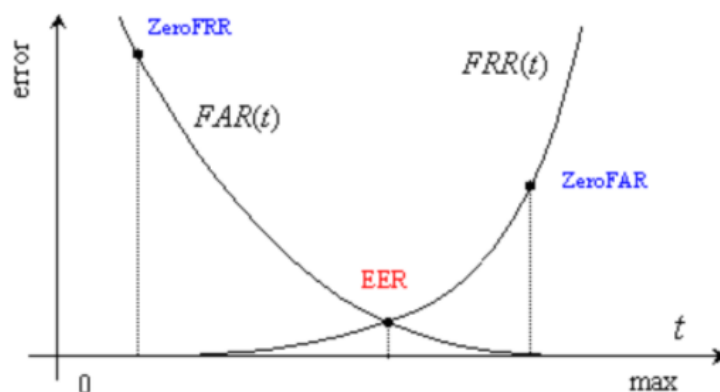
THRESHOLD

entrambe le misure sono dipendenti dal threshold adottato perché più è alto più aumenta il False Rejection e diminuisce il False Acceptance, e viceversa. Solitamente il threshold si sceglie uguale all'Equal Error Rate. Infatti, tracciando il FAR e il FRR troviamo tre punti d'interesse:

Equal Error Rate (EER): punto d'intersezione tra FAR e FRR;

Zero FRR: proiezione sul FAR del punto dove il FRR vale zero;

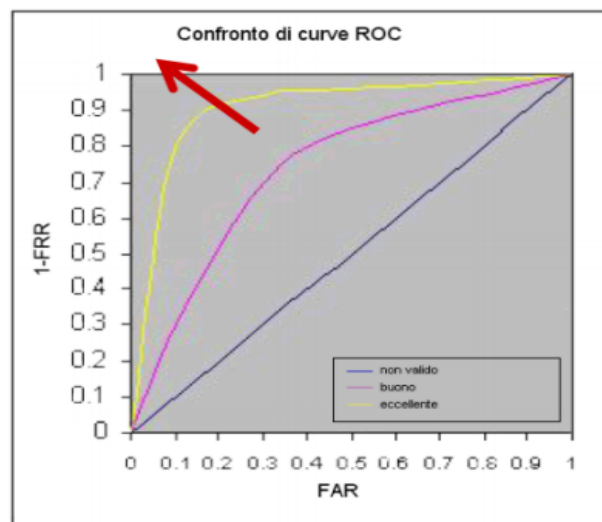
Zero FAR: proiezione sul FRR del punto dove il FAR vale zero.



(ROC) Receiver Operating Characteristic

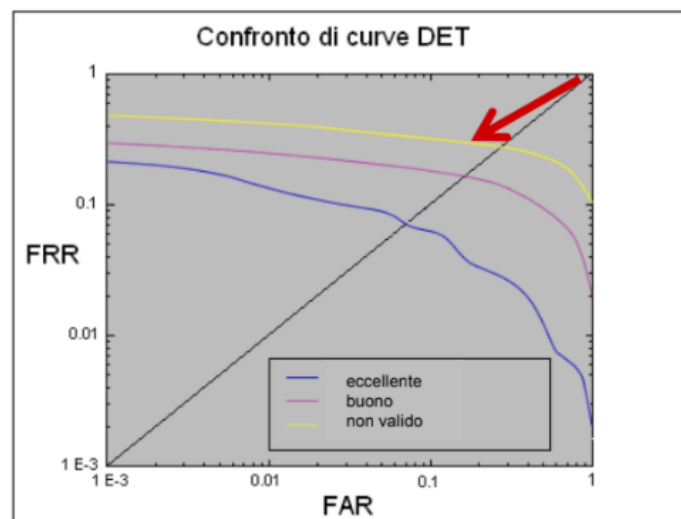
curva tracciata su un piano cartesiano dove gli assi sono il Genuine Acceptance Rate e il False Acceptance Rate. Rappresenta quindi la probabilità che un soggetto venga correttamente identificato al variare del FAR.

Sull'asse delle x sono quindi presenti il numero di volte in cui accetto erroneamente un utente e sull'asse delle y il numero di volte in cui accetto correttamente un utente nel sistema. Per questo motivo più la curva si trova in alto a sinistra del grafico meglio è (GA > FA).



(DET) Detection Error Trade-Off

curva logaritmica tracciata su un piano cartesiano dove gli assi sono il False Rejection Rate e il False Acceptance Rate. Rappresenta quindi la probabilità che un soggetto venga erroneamente non identificato al variare del FAR. Trattandosi di due indici negativi per il nostro sistema più la curva nel grafico è bassa meglio è.



rango(p) : posizione nella lista dei similarity score dove è presente il primo template dell'identità corretta.

FORMULE OPEN SET

DIR - DETECTION AND IDENTIFICATION RATE (AT RANK K)

Misura la probabilità di avere un'identificazione corretta al rank k (il soggetto corretto è presente nella posizione k). Quindi il miglior risultato che si può ottenere è il seguente:

$$- \text{DIR}(t, 1) = 1$$

$$\text{DIR}(t, k) = \frac{|\{p_j : \text{rango}(p_j) \leq k, s_{ij} \geq t, id(g_i) = id(p_j)\}|}{|P_G|} \quad \forall p_j \in P_G$$

FRR - FALSE REJECT RATE

E' il complemento del DIR

$$\text{FRR}(t) = 1 - \text{DIR}(t, 1)$$

FAR (False Acceptance Rate o False Alarm Rate (Watch List)): the probability of false acceptance/alarm

The rate between the number of impostor recognized by error and the total number of impostors in P_N

$$\text{FAR}(t) = \frac{|\{p_j : \max_i s_{ij} \geq t\}|}{|P_N|} \quad \forall p_j \in P_N \quad \forall g_i \in G$$

EER (Equal Error Rate): the point where the two probability errors are equal, i.e., $\text{FRR} = \text{FAR}$

$$\text{EER} = \{x : \text{FRR}(t) = x \wedge \text{FAR}(t) = x\}$$

FORMULE CLOSED SET

CMS - CUMULATIVE MATCH SCORE

Probabilità di avere una corretta identificazione entro la posizione k

CMC - CUMULATIVE MATCH CHARACTERISTICS

Mostra il valore assunto dal CMS al variare di k.

Rappresenta quindi la probabilità che un soggetto venga correttamente identificato al variare del Rank k.

In questo caso se il numero del rank che consideriamo è uguale alla dimensione del gallery avremo sicuramente probabilità 1 di avere un'identificazione corretta (closed set).

LEZIONE 2bis

OFFLINE PERFORMANCE EVALUATION

Per studiare statisticamente il nostro sistema biometrico dobbiamo testarlo offline.

Tutti i soggetti presentano una label e noi conosciamo la loro vera identità, cosa non possibile per operazioni reali.

Il dataset può essere diviso nei seguenti modi:

- **Training Set e Test Set**

Possibili configurazioni di TS E TR:

- 1) Ogni soggetto appare in entrambi i set ma con samples diversi
- 2) Il TR presenta un subset di soggetti mentre il TS contiene tutti i soggetti del TR + altri soggetti che non sono nel TR

- **Probe Set e Gallery Set**

Possibili configurazioni Probe e Gallery:

- 1) Per emulare la realtà vengono messe le immagini migliori nel gallery set e le restanti nel probe set
- 2) Si può avere sia un probe set che presenta i soggetti della gallery (closed set) o un probe set che include soggetti non presenti nella gallery (open set)

Validation

per validare il nostro modello dovremmo testare il nostro sistema biometrico con le diverse configurazioni possibili (TR, TS, PROBE, GALLERY), ottenere dei risultati, ripetere la procedura da capo e fare una media dei risultati ottenuti.

ALL AGAINST ALL - PROBE VS GALLERY

Utile quando si vuole testare un dataset con pochi template.

Supponiamo di avere 3 soggetti e per ogni soggetto abbiamo a disposizione 3 template, con un totale di 9 template.

Per trarre delle statistiche sul nostro sistema biometrico possiamo costruire la matrice all probe vs all gallery dove le righe verranno assunte come degli incoming probe e ogni colonna è assunta come un template.

Quindi per trarre delle statistiche in questo caso ogni probe dovrà matchare correttamente 2 volte (escludendo il match con se stessi) e le restanti 6 bisogna considerare l'incoming probe come impostore.

LEZIONE 3

TEMPLATE UPDATING

Ci sono due metodi per il template updating:

1) Label Assignment

- **Supervised**

Un supervisore assegna il label di identità ai nuovi dati acquisiti durante il processo di ricognizione

- **Semi-Supervised**

Non ci sono persone a controllare il processo. Se l'incoming probe presenta una distanza media rispettivamente a tutti i template nel gallery e questa media è al di sotto di un determinato threshold allora salvo il probe nel gallery

2) Selezione del template

- **Online**

Ogni nuovo input viene salvato

- **Offline**

La selezione avviene dopo un determinato lasso di tempo così facendo è possibile togliere i template che danno un minore contributo alla recognition.

DODDINGTON ZOO

- Colomba: raramente incorrono in errori, è il miglior utente (Genuine Acceptance);
- Pecora: matchano bene con sé stessi e poco con quelli di altri (Genuine Acceptance);
- Capra: matchano male con i loro stessi tratti (False Rejection);
- Fantasma: soffrono di basse corrispondenze (False Rejection);
- Agnello: possono essere facilmente impersonati (False Acceptance);
- Lupo: sono bravi ad impersonare altri utenti (False Acceptance);
- Camaleonte: appaiono sempre come altri utenti (False Acceptance);
- Verme: causano spesso errori di sistema, è il peggior utente. Forse è dato da un problema strutturale del nostro sistema biometrico

RELIABILITY MEASURE:

Dato che al nostro sistema biometrico possono presentarsi diversi soggetti in diverse condizioni è opportuno stabilire delle misure di affidabilità.

Se queste misure non vengono rispettate il sistema può chiedere al soggetto di ripetere l'acquisizione dell'impronta, foto ecc...

Per quanto riguarda la Face Recognition possiamo definire la qualità di un'immagine basandoci su tre valori: Yaw, Roll e Pitch (rispettivamente rotazione lungo gli assi z, x e y). Possiamo scartare i dati ininfluenti o superflui, ma una buona misura di qualità è quella che permette le migliori performance senza scartare troppi dati.

SRR - SYSTEM RESPONSE RELIABILITY

Misura l'abilità di un sistema d'identificazione nel separare i soggetti genuini dai soggetti impostori sulla base di un singolo probe. Sostanzialmente sta ad indicare quanto è affollato il probe sottomesso in quanto se è molto affollato allora la risposta non sarà affidabile altrimenti se dovesse essere poco affollato la risposta è affidabile.

L'SSR si basa sulla misura della distanza tra un probe identificato e tutti gli altri nella gallery. Se la distanza tra tale probe e i suoi vicini è bassa significa che aumenta la probabilità di errore.

Given a probe p and a system A with gallery G , the **relative distance** is defined as:

$$\varphi(p) = \frac{F(d(p, g_{i_2})) - F(d(p, g_{i_1}))}{F(d(p, g_{i_{|G|}}))}$$

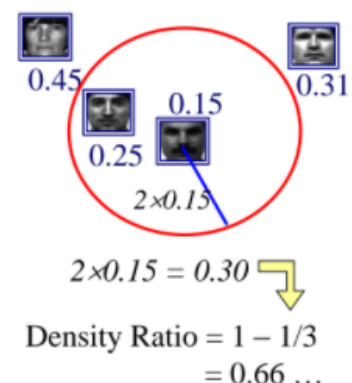
La formula è data dal rapporto tra la distanza relativa tra il probe e il secondo elemento con distanza minore dal probe meno la distanza relativa tra il probe e il primo elemento con distanza minore dal probe. Il tutto diviso dalla distanza relativa tra il probe e l'ultimo elemento presente nel similarity ranking.

Given a probe p and a system A with gallery G , the **density ratio** is defined as:

$$\varphi(p) = 1 - |N_b| / |G|$$

With

$$N_b = \{g_{i_k} \in G \mid F(d(p, g_{i_k})) < 2 \cdot F(d(p, g_1))\}$$



Nb è il numero di elementi nella gallery che hanno la distanza con il probe minore del doppio della distanza del probe e il primo elemento.

FORMULA SRR

We also define s as the width of the subinterval from $\varphi(p)$ to the proper extreme of the overall $[0,1)$ interval of possible values, depending on the comparison between the current $\varphi(p)$ and $\bar{\varphi}$:

$$s(\varphi(p), \bar{\varphi}) = \begin{cases} 1 - \bar{\varphi} & \text{if } \varphi(p) > \bar{\varphi} \\ \bar{\varphi} & \text{otherwise} \end{cases}$$

SRR index can finally be defined as:

$$SRR = (\varphi(p) - \bar{\varphi}) / S(\bar{\varphi})$$

phi segnato o critical phi rappresenta il valore in grado di minimizzare le stime errate date dalla funzione phi.

THRESHOLD PER SRR

- We would desire to have a high average (the system is generally reliable) and a low variance (the system is stable)-
- We can summarize as:

$$th_i = \left| \frac{E[\bar{S}_i]^2 - \sigma[\bar{S}_i]}{E[\bar{S}_i]} \right|$$

LEZIONE 4

I più importanti fattori che influenzano la validità di un sistema biometrico sono:

- accuratezza e accettabilità.

Ad esempio il DNA è il carattere biometrico più accurato ma per prelevarlo prevede una procedura invasiva, mentre il viso prevede un'alta accettabilità e una procedura poco invasiva.

Possibili problemi: Intra-Person Variations e Inter-Personal Similarities (rispettivamente molte differenze tra noi stessi in condizioni diverse e poche differenze tra persone diverse).

PIE

Pose, Illumination and Expression vari database cercano di raccogliere immagini facciali tenendo in considerazione il PIE.

DATABASE FACCIALI

- **FERET(1996)**
Immagini totali 14051, fatte in pose, luminosità e espressioni diverse.
- **AR-FACES(1998)**
Immagini prese con soggetti ripresi con espressioni differenti. I soggetti possono presentare delle cicatrici o occhiali
- **CMU-PIE (2000)**
risponde al PIE con 68 soggetti in 41344 immagini (608 a soggetto)
- **CASIA 3D Face V1 (2004)**
contiene 4624 render 3D di 123 persone.
- **LFW (LABEL FACES IN THE WILD, 2007)**
In the wild => perché sono foto non prese in laboratorio
- **YOUTUBE FACES (2011)**
Immagini prese da frame di video youtube

Step per la Face Recognition

Input image, Face Detection, Feature Extraction, Face Recognition e Verification (o Evaluation).

Un riconoscitore facciale è suddiviso in tre parti:

- 1) Miglioramento dell'immagine.
 - aumento della nitidezza
 - elimina la sfocatura
 - aumenta il contrasto
- 2) Detection and Localization
 - Detection: ritorna vero o falso se la localizzazione ha avuto successo
 - Localization: ritorna le coordinate del ROI.
Possibile errore => False Localization: Il sistema si sbaglia a fare la localizzazione

- 3) Viene presa in considerazione solamente la regione d'interesse (ROI) per diminuire il tempo di processo

LEZIONE 5

Possibili approcci per la face localization:

A) Algoritmo A by Hsu, Mottaleb and Jain (2002)

Questo algoritmo si divide in due macro fasi:

1) Detection del viso del candidato (pre processing)

a) Compensazione dell'illuminazione

L'illuminazione viene normalizzata sulla base di un bianco di riferimento cioè il valore di bianco presente nel top 5% di pixel ordinati per valori di luma.

Luma identifica la luminosità di un pixel rgb ed è data dalla somma pesata dei componenti RGB.

Se un numero sufficiente di bianchi referenziati è stato trovato allora i componenti RGB dell'immagine vengono sistemati in modo tale di avere un'uniformità di luminosità.

b) Trasformazione dello spazio dei colori

I colori che sono vicini nello spazio RGB possono non essere percepiti come simili. In questo caso i migliori spazi dei colori rispetto a RGB sono gli spazi YUV (Componenti di luminanza e cromaticità) e HSV (prende in considerazione i valori di tonalità e saturazione)

/*

Due vettori RGB che non sono simili per la distanza euclidea possono essere percepiti come simili.

YCbCr

- Y: luminance component
- Cb: blue component
- Cr: red component

Formula per trasformare RGB in YCbCr

$$\begin{bmatrix} Y \\ C_B \\ C_R \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \frac{1}{256} \begin{bmatrix} 65.738 & 129.057 & 25.064 \\ -37.945 & -74.494 & 112.439 \\ 112.439 & -94.154 & -18.285 \end{bmatrix} \bullet \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

*/

c) Localizzazione basata sullo skin model

Due metodi possibili

- VARIANCE BASED THRESHOLDING METHOD

Dall'immagine vengono considerate due classi differenti che presentano delle scale di grigio dove C1 avrà i valori di grigio

superiori ad una determinata soglia e C2 avrà i valori di grigio inferiori sempre alla stessa soglia.
Il threshold viene quindi calcolato in modo tale da minimizzare la varianza intra-class tra le due classi.

- CONNECTED COMPONENT

I componenti connessi vengono raggruppati secondo la vicinanza spaziale e la similarità tra i colori. Il risultato è una regione che non presenta buchi.

2) Localizzazione delle caratteristiche facciali

- EYE LOCALIZATION

L'algoritmo costruisce due mappe differenti per gli occhi:

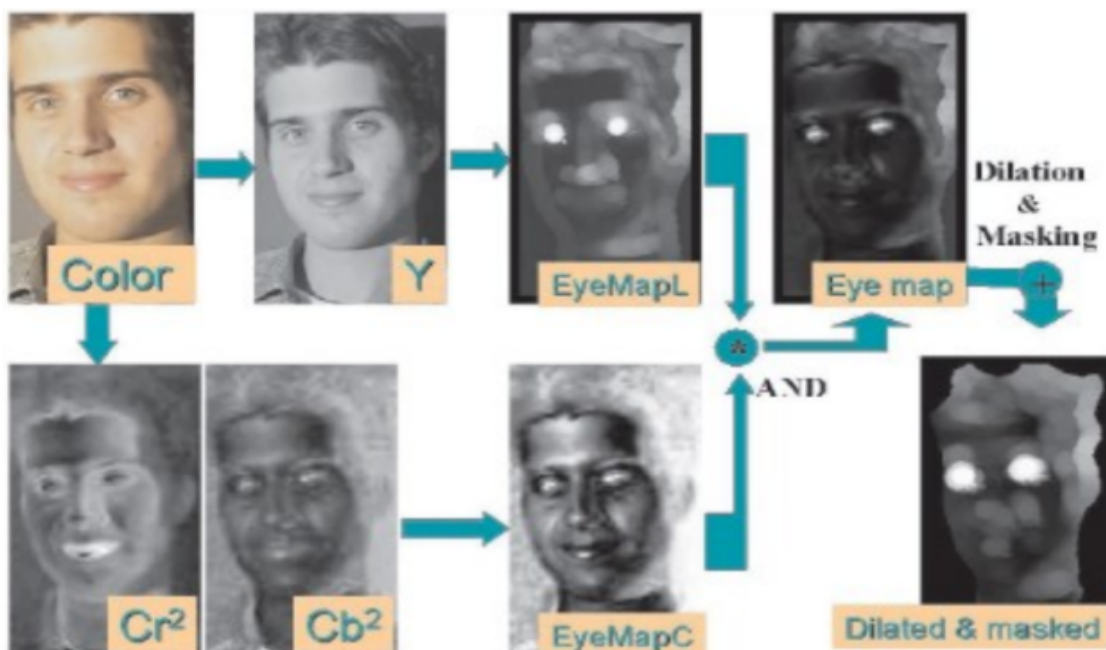
- a) CHROMINANCE MAP

Questa viene costruito prendendo in considerazione il fatto che la regione intorno agli occhi è caratterizzata da alti valori di Cb e bassi valori di Cr

- b) LUMINANCE MAP

Gli occhi generalmente contengono sia zone chiare che zone scure che possono essere evidenziate dai due operatori morfologici: Dilation and Erosion con una struttura semisferica (struttura simile all'occhio).

Successivamente queste due mappe vengono combinate insieme tramite l'operatore logico AND e al risultato ottenuto viene applicato un processo di masking e dilation in grado di esaltare la posizione degli occhi.



DILATION

- Prende in input un'immagine e un kernel. Quest'ultimo viene fatto traslare sull'intera immagine e ogni qualvolta l'intersezione tra il kernel e l'immagine non è vuota allora il centro del kernel viene colorato dando origine ad una nuova immagine.

EROSION

- In questo caso il centro del kernel viene colorato se e solo se tutti i punti del kernel fanno intersezione con l'immagine

- MOUTH LOCALIZATION

Procedimento simile a quello per gli occhi, ma si usa solo il Dilation Operator.

- FACE CONTOUR

Gli occhi e la bocca formano un triangolo. Per ogni triangolo individuato si va a studiare l'orientamento e la geometria e la presenza di un contorno del viso attorno al triangolo.

B) VIOLA-JONES

Algoritmo basato sulle immagini. Viene applicato per il face detection ma può essere comunque utilizzato per altri tipi di detection come mouth and eye detection.

Questo algoritmo prevede la creazione di un classifier e il suo relativo training.

Nella fase di training vengono sottomessi all'algoritmo due tipi di istanze della classe da identificare:

- 1) positive examples (contenenti il viso)
- 2) negative examples che possono causare errori (non contengono visi)

Il numero di Misses (oggetto non identificato) e di false alarm (oggetto identificato ma che non è presente) può essere ridotto facendo il re-training dell'algoritmo con nuovi examples.

- PROCEDURA DI LEARNING

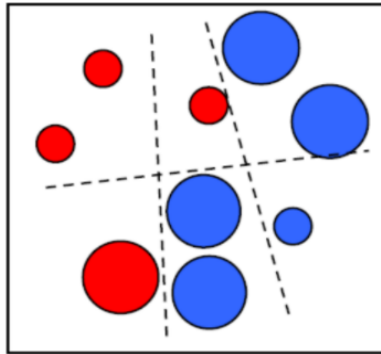
Il boosting è uno schema di classificazione che funziona mettendo insieme M weak learner (lineari) per ottenerne uno più forte (non lineare).

$$H_M(x) = \frac{\sum_{i=1}^M \alpha_i h_i(x)}{\sum_{i=1}^M \alpha_i}$$

x is a pattern to classify
 $h_i(x) \in \{-1, +1\}$ are the weak classifiers
 $\alpha_i \geq 0$ are the relative weights
 $\sum_{i=1}^M \alpha_i$ is a normalization factor

Questa procedura richiede un set di training pattern costituito da coppie x, y dove x è il pattern e y è la label associata al pattern che sta ad indicare se il pattern corrisponde o meno ad un viso (in questo caso).

Durante la procedura di learning una distribuzione di pesi viene associata al training pattern e inizialmente tutti i samples hanno peso uguale. Dopo una determinata iterazione il pattern più complicato da classificare avrà un peso maggiore. Così facendo all'iterazione successiva il sistema gli dedicherà maggiore attenzione.



HAAR FEATURE

Inoltre è possibile definire ulteriori weak learner tramite le HAAR FEATURES.

Questi sono dei rettangoli suddivisi in zone di bianco e in zone di nero. Queste features vengono applicate in una sotto regione di una sliding window (che passa sull'immagine intera) e il valore che ritornano è dato dalla sommatoria dei valori dei pixel presenti nell'area bianca meno la sommatoria dei pixel presenti nell'area nera della feature.

Per ogni round di boosting si ha:

- Evaluation dei filtri rettangolari sugli example
- Si seleziona il miglior threshold per ogni filtro
- Si seleziona la miglior combinazione threshold/filtro che ci permette di distinguere una face class da una non face class
- Si riassegnano i pesi agli examples

COME CALCOLARE HAAR FEATURE

Avviene tramite le immagini integrali. Una integral images nel punto x,y è la somma dei pixels in alto e a sinistra di questo punto. Mentre stiamo muovendo la sliding windows è possibile che alcuni pixel siano stati già calcolati. L'obiettivo di questo processo è di salvare il tempo computazionale.

WEAK CLASSIFIER

/* Il più semplice classificatore (weak) è il Albero di Decisione con un solo nodo. */

Considerando di avere a disposizione già M-1 weak classifier vogliamo costruire $h_M(x)$ classifier.

$$h_M(x) = +1 \text{ if } z_{k^*} > t_{k^*} \\ = -1 \text{ otherwise}$$

Questo classificatore compara una certa feature con un determinato threshold e assegna il valore +1 o -1 a seconda di questi due valori.

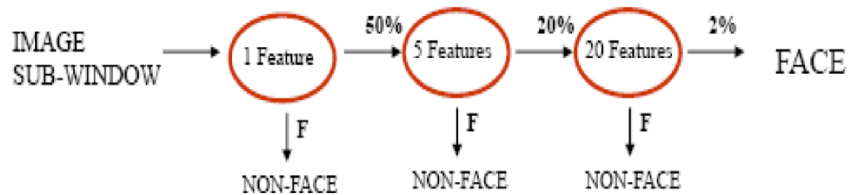
La feature z_k per il seguente classifier è scelta per ottenere l'errore minore

CLASSIFICATORI IN CASCATA

L'obiettivo è quello di minimizzare l'identificazione delle zone dove non è presente il viso. Si mettono quindi in cascata diversi classifier (dal più semplice al più complesso) e i risultati positivi del primo classifier fanno iniziare la valutazione del secondo classificatore più complesso e così via.

Un risultato negativo a qualsiasi punto determina lo scarto di questa subwindow.

Per allenare una cascata di classificatori ogni classificatore viene allenato dai false positive del classificatore precedente.



VALUTARE LA LOCALIZZAZIONE

- FALSE POSITIVE
Percentuale di finestre classificate come facce che però non contengono alcun volto
- NOT LOCALIZED FACES
Percentuale di volti che non sono stati identificati
- C-ERROR
Distanza euclidea dal reale centro del viso e da quella stimata dal sistema.

LEZIONE 6

Enrollment + Testing.

POSSIBILI RAPPRESENTAZIONI DEL VISO

- **FUNZIONE BI-DIMENSIONALE**
L'immagine può essere rappresentata come una funzione definita nel piano cartesiano che associa valori da 0 a 255 ad ogni elemento
- **MATRICE**
Un'immagine è generalmente salvata in una matrice dove ogni elemento è un pixel
- **VETTORE**
Può essere rappresentata anche sotto forma di vettore.

MALEDIZIONE DELLA DIMENSIONALITÀ

Questo problema si presenta quando si vogliono studiare e analizzare dei dati rappresentati su più dimensioni e che sono disposti in modo sparso come può succedere con le feature del volto. Più gli spazi aumentano più i dati saranno sparsi. Nel nostro caso la soluzione è quindi quella di ridurre lo spazio di rappresentazione attraverso l'Image reduction.

/*svantaggio della rappresentazione di un'immagine, si riferisce ai vari fenomeni che si presentano durante l'analisi e l'organizzazione dei dati in spazi con molte dimensioni (spesso con centinaia di migliaia di dimensioni) che invece non si verificano, ad esempio, nello spazio fisico a tre dimensioni della quotidianità.

Quando la dimensionalità aumenta, il volume dello spazio aumenta così velocemente che i dati disponibili diventano sparsi. Poiché per classificare i dati vengono individuate aree dove gli oggetti formano gruppi con proprietà simili, in dati con tante dimensioni tutti gli oggetti appaiono sparsi e di conseguenza differenti, prevenendo una organizzazione dei dati dall'essere efficiente. Quindi, poiché un'immagine è rappresentata da un feature vector composto da un altissimo numero di features, ognuna con un numero di possibili valori, è richiesto un'enorme quantità di training data. La soluzione è dunque quella di ridurre lo spazio della rappresentazione dell'immagine, attraverso l'Image Reduction. */

PRINCIPAL COMPONENT (PCA)

Metodo statistico utilizzato per ridurre la dimensionalità. Si prendono in considerazione un numero di eigenvector uguale al numero delle dimensioni. Questi devono essere perpendicolari tra di loro. Su questi vengono proiettati tutti i dati. Infine si sceglie l'eigenvector con eigenvalue maggiore (l'eigenvalue descrive la varianza) ossia la componente principale.

/* Utilizza una trasformazione ortogonale sui dati correlati in modo tale da ottenere un insieme di dati non correlati quindi non ridondanti chiamati **principal component**.

La trasformazione avviene in modo tale che il primo principal component presenti la varianza maggiore (maggior numero di informazioni) e che rispetto ai successivi componenti sia ortogonale qui non correlato.* /

Algoritmo:

- 1) Si parte da un training set TS
- 2) Dal training set si calcola il vettore media

$$\bar{x} = \frac{1}{m} \sum_{i=1}^m x_i$$

- 3) Infine si calcola matrice di covarianza

$$C = \frac{1}{m} \sum_{i=1}^m (x_i - \bar{x})(x_i - \bar{x})^T$$

- 4) Da questa matrice è possibile estrarre gli eigenvector con associati gli eigenvalues
- 5) Vengono ordinati gli eigenvalues in ordine decrescente e vengono presi i primi k eigenvector
- 6) L'eigenvector rappresenta la direzione della massima variazione dei dati, quindi dove è presente la maggiore quantità di informazione.

I primi ad usare la pca per la face recognition sono stati Sirovich e Kirby i quali hanno dimostrato che ogni viso può essere rappresentato nello spazio degli eigenpicture. Turk e Pentland partendo dal lavoro di S e K hanno costruito un sistema di riconoscimento facciale il quale considera la proiezione del viso dell'individuo nello spazio degli eigenfaces. In questo spazio avviene il riconoscimento.

COME AVVIENE IL RICONOSCIMENTO?

$$Proj(x) = \varphi_k^T (x - \bar{x})$$

phi di t k è la matrice di covarianza con le k colonne contenenti gli eigenvector con eigenvalue maggiori. La recognition avviene mediante la distanza euclidea tra immagine e componenti PCA.

Questo sistema soffre però le PIE variation. Per risolvere questo problema sono state introdotte le Fisherfaces utilizzate anche nel processo di LDA.

/* Una volta acquisita una nuova immagine a questa viene applicata la seguente proiezione:

$$Proj(x) = \varphi_k^T (x - \bar{x})$$

Dove φ_k^T è la matrice di proiezione dove le colonne sono i k eigenvector con varianza maggiore. La face recognition avviene mediante la distanza euclidea tra l'immagine e le principal component presenti nel training set. */

Questo sistema soffre però le PIE variation. Per risolvere questo problema sono state introdotte le Fisherfaces utilizzate anche nel processo di LDA.

LDA - LINEAR DISCRIMINANT ANALYSIS

Ha come obiettivo quello di massimizzare la separazione delle classi nello spazio ridotto e quindi di mantenere le informazioni discriminative delle classi il più possibile.

FUNZIONAMENTO

- 1) Si parte da un Training Set TS = {x=i,...m,}
- 2) Differentemente dalla PCA il TS viene partizionato in base alle classi e ogni componente della PTS={P1, ..., PS} ha cardinalità m (cardinalità del training set)
- 3) Vogliamo quindi ottenere uno scalare y dato dalla proiezione dei samples x su una linea

$$y = w^T x$$

Consideriamo due classi P1 e P2 (possono essere considerate anche più classi) che hanno rispettivamente i vettori m_1 e m_2 .

- 1) Per ogni classe calcoliamo il vettore media:

$$\mu_i = \frac{1}{m_i} \sum_{j=1}^{m_i} x_j$$

and

$$\tilde{\mu}_i = \frac{1}{m_i} \sum_{j=1}^{m_i} y_j = \frac{1}{m_i} \sum_{j=1}^{m_i} w^T x_j = w^T \mu_i$$

(μ barretta è la proiezione del punto medio delle classi sull'asse)

- 2) Adesso vogliamo massimizzare la distanza tra i due mean vector delle due classi

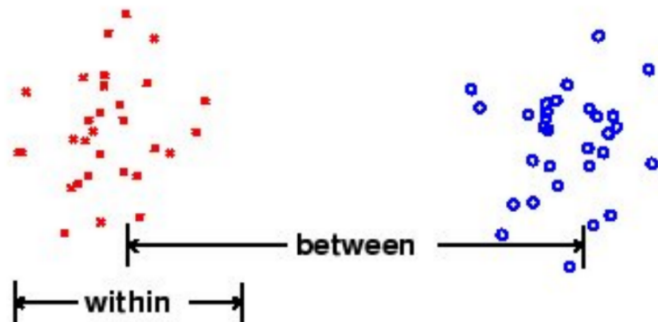
$$J(w) = |\tilde{\mu}_1 - \tilde{\mu}_2| = |w^T(\mu_1 - \mu_2)|$$

- 3) μ barretta 1 meno μ barretta 2 non sono però delle buone misure perché non tengono in considerazione la standard deviation (non ci informa su quanto i dati varino) tra le classi per questo motivo si usa la **SCATTERING MATRICES**

SCATTERING MATRICES

Obiettivo:

Minimizzare la varianza all'interno della stessa classe e massimizzare quella tra le classi



Funzionamento

- 1) Si parte da un training set TS
- 2) Differentemente dalla PCA il TS viene partizionato in base alle classi e ogni componente della PTS ha cardinalità m (cardinalità del training set)
- 3) Per ogni classe calcoliamo il vettore medio e poi la media delle medie

$$\mu_i = \frac{1}{m_i} \sum_{j=1}^{m_i} x_j \quad \mu_{TS} = \frac{1}{m} \sum_{i=1}^S m_i \mu_i$$

- 4) Per ogni classe calcoliamo la matrice di covarianza

$$C_i = \frac{1}{m_i} \sum_{j=1}^{m_i} (x_j - \mu_i)(x_j - \mu_i)^T$$

- 5) Si calcola la within class scatter matrix S_W

$$S_W = \sum_{i=1}^S m_i C_i$$

m è il numero di elementi collezionati per ogni classe. S_W è dato dalla somma pesata della matrice di covarianza.

- 6) Si calcola la between class scatter matrix

$$S_B = \sum_{i=1}^S m_i (\mu_i - \mu_{TS})(\mu_i - \mu_{TS})^T$$

(for two classes it is also defined as $S_B = (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T$)

- 7) Questa definizione vale anche per S_W cappelletto e S_B cappelletto con y considerato come una proiezione su un asse. Quindi lo scatter di una classe proiettata è definita come:

$$\tilde{s}_i^2 = \sum_{y \in P_i} (y - \tilde{\mu}_i)^2$$

FISHER SOLUTION

Obiettivo:

Massimizzare la distanza tra le medie normalizzate attraverso una misura dello scatter intra-class.

Esempio con due classi:

$$J(w) = \frac{|\tilde{\mu}_1 - \tilde{\mu}_2|^2}{\tilde{s}_1^2 + \tilde{s}_2^2}$$

(Numeratore: massimizzazione; Denominatore: minimizzazione)

Vogliamo quindi ottenere una proiezione dove gli examples sono vicini tra loro e contemporaneamente le medie delle proiezioni lontane tra loro.

Sia il numeratore che il denominatore di questa funzione possono essere riscritti anche in funzione delle scatter matrix nello spazio delle features di partenza.

$$J(w) = \frac{w^T S_B w}{w^T S_W w}$$

BUBBLE EXPERIMENT

Confronta l'osservatore umano e l'osservatore ideale con due tipi di task:

- 1) Capire il sesso del soggetto
- 2) Capire l'espressione del soggetto

Il secondo andrà a prendere in considerazione delle regioni dell'immagine che presentano la varianza maggiore e che l'osservatore umano invece non prende in considerazione

WAVELET TRANSFORM

La trasformata di Fourier ci dà informazioni sulla frequenza di un certo segnale quindi ci informa sulla quantità di frequenza di un segnale.

La trasformata wavelet invece ci permette di conoscere il tempo di ogni frequenza e il procedimento per applicare questa trasformata è il seguente:

Viene fatta passare una Onda Madre nel dominio del tempo con diverse dimensioni.

Maggiore sarà la dimensione dell'onda madre maggiori saranno le feature catturate e discorso analogo vale per la dimensione minore della Onda Madre.

Esempio: Da un'immagine 2D vogliamo conoscere i livelli di grigio e inoltre dove questi livelli appaiono nell'immagine.

GABOR FILTER

Filtro lineare per rilevare i bordi di un'immagine.

Questo filtro funziona come un filtro passa banda. Sono utili perché fanno passare le informazioni umane.

Il filtro si basa su una funzione Gaussiana.

$$\psi_{f,\theta}(x,y) = \exp\left[-\frac{1}{2}\left\{\frac{x_{\theta_n}^2}{\sigma_x^2} + \frac{y_{\theta_n}^2}{\sigma_y^2}\right\}\right] \exp(i(2\pi f x_{\theta_n}))$$

La matrice di trasformazione viene moltiplicata con il vettore originale in modo tale da calcolare la risposta.

$$\text{where, } \begin{bmatrix} x_{\theta_n} \\ y_{\theta_n} \end{bmatrix} = \begin{bmatrix} \sin \theta_n & \cos \theta_n \\ -\cos \theta_n & \sin \theta_n \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

M. TRASFORMAZIONE

I feature vector vengono ottenuti facendo la convoluzione dell'immagine con una serie di gabor filter.

Se si facesse la convoluzione pixel per pixel la dimensionalità sarebbe troppo elevata per ovviare a questo problema si possono applicare due soluzioni:

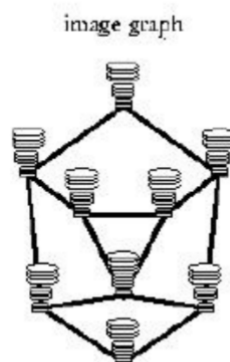
- 1) Applicare una griglia fissata di punti inserita a mano
- 2) Considerare solo le regioni con il maggior quantitativo di informazione

EBGM

Ogni posa è rappresentata come un bunch(mazzo/gruppo) graph.

I nodi che vengono presi in considerazione sono riposti sui bordi del viso, sugli occhi e sulla bocca avendo così tutte le distanze tra questi punti.

Su ogni nodo sono presenti dei jets calcolati tramite i wavelet che ci danno quindi informazioni riguardanti la frequenza in quel punto.



Per far funzionare correttamente il sistema bisogna dargli un insieme di samples che presentano diverse pose e caratteristiche. In questo modo l'algoritmo sarà in grado di creare dei bunch appositi per ogni feature. Ad esempio un bunch (è un set di jets che fanno

riferimento ad uno stesso punto) per gli occhi può contenere informazioni riguardanti gli occhi: aperti, chiusi, maschili e femminili.

LBP - LOCAL BINARY PATTERN

E' un descrittore di texture.

Si basa su un operatore di texture che cerca di trovare dei pattern in modo tale da distinguere le regioni.

Lavora pixel per pixel e nella sua versione basica considera 8 vicini, quindi la sliding window è grande 3x3.

Il centro della sliding window funziona da threshold e viene comparata con tutti i suoi vicini.

Se il valore del vicino è maggiore o uguale al threshold allora sarà assegnato valore 1 a quel vicino altrimenti zero.

Infine si fa la somma pesata per calcolare il valore di LBP.

example		
6	5	2
7	6	1
9	8	7

thresholded		
1	0	0
1		0
1	1	1

weights		
1	2	4
128		8
64	32	16

Pattern = **11110001**

LBP = $1 + 16 + 32 + 64 + 128 =$ **241**

C = $(6+7+8+9+7)/5 - (5+2+1)/3 =$ **4.7**

Si può inoltre calcolare il contrasto facendo la media dei valori sopra il threshold meno la media dei valori sotto il threshold.

Sono presenti inoltre delle variazioni di questo algoritmo dove vengono usate window più grandi e con raggi più grandi. All'aumentare del raggio cambierà anche il numero di vicini considerati.

Un pattern è uniforme se contiene al più due transizioni 0-1 o 1-0. Un pattern uniforme serve per salvare la memoria.

Dopo la recognition si può dividere l'immagine di un viso tramite una matrice e su ogni elemento possiamo calcolare l'istogramma LBP. L'istogramma LBP è calcolato partendo da un'immagine LBP considerando i livelli di grigio.

A seconda del punto di partenza il pattern cambia e per questo bisogna scegliere sempre lo stesso. Il punto di partenza per tutti è quella che presenta il maggior numero di zeri.

GLOBAL APPEARANCE METHOD

Si basano sull'intera immagine quindi non perdono alcuna informazione a discapito però del costo computazionale.

Molti algoritmi sono flessibili e quindi possono essere modificati in modo tale da gestire la PIE.

Richiedono però una grande correlazione tra il training set e il test set.

Metodi global: PCA e LDA.

LOCAL FEATURE BASED METHOD

- EBGM, LBP

PRO: Il matching è rapido e leggero computazionalmente.

CONTRO: Approccio handmade per decidere quali feature estrarre .

LEZIONE 7

FACE RECOGNITION 3D

Molto più complesso rispetto al face recognition 2D ma possono essere analizzate maggiori informazioni ed è inoltre possibile sintetizzare immagini 2D partendo da un modello 3D.

Rappresentazione:

- 2D il valore di ogni pixel è dato dall'illuminazione
- 2.5D è dato dalla distanza tra quel punto e la sorgente di luce
- 3D la struttura è rappresentata come un insieme di punti e poligoni connessi in uno spazio 3D.

ACQUISIZIONE 3D

- Camera stereoscopica

ha due o più lenti con un frame d'immagine per ogni lente, permette di simulare la visione binoculare umana.

- Structured Light Scanner

viene proiettato sul soggetto un determinato pattern. L'informazione riguardante il modello 3D viene quindi catturato a seconda di come questo pattern viene deformato sul viso del soggetto.

- Laser Scanner

Un laser è proiettato sul volto ma è pericoloso per gli occhi ma è molto preciso.

PREPROCESSING DI UNO SCANNER 3D

Rimozione delle distorsioni come buchi o spikes.

Integrating several 2.5D face scans

integrazione di numerose scansioni 2.5D da punti di vista differenti per comporne una singola 3D. Per ogni immagine 2.5D vengono creati dei punti 3D. Alla fine del processo questi punti vengono triangolarizzati in modo tale da ottenere il modello finale.

MODELLO 3D

Un poligono è una sequenza coplanare di punti connessi da dei segmenti.

Coplanare = appartenenti allo stesso piano.

L'orientamento delle superfici è determinato dalle normali, ossia da vettori perpendicolari al piano.

L'immagine 3D composta da una maglia (mesh) di poligoni, indicanti l'approssimazione della figura.

FROM 2D TO 3D - SHAPE FROM SHADING

Calcola una forma 3D dall'illuminazione del viso, assunta come Lambertiana (che riflette la luce in tutte le direzioni). Si può calcolare unendo forme 3D e applicando il PCA o usando soltanto l'immagine 2D in input usandola come guida.

MORPHABLE MODEL

Iniziando da un generico modello 3D ottenuto da un insieme di visi 3D, si manipola per arrivare a un modello 3D di un soggetto specifico.

3D FEATURE

Dobbiamo estrarre le Features, rappresentate da linee locali o globali che possono essere Crest Lines, Local Curvatures e Local Features.

- Crest Lines: aree che presentano la maggiore curvatura
- Local Curvature: curvatura locale di un determinato colore
- Local Feature: Segmenta la faccia in ROI e estrae le feature da ogni singola regione.

3D ALIGNMENT

- COARSE

L'allineamento può avvenire tramite dei punti di riferimento.

- 1) vengono trovati un numero finito di punti caratteristici
- 2) viso viene allineato minimizzando la distanza tra i punti corrispondenti

- ICP

Allineamento tra modelli:

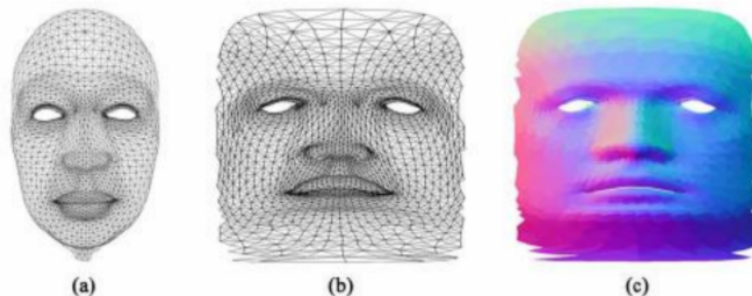
- 1) Si trova un match iniziale tra le due superfici
- 2) Si calcola la distanza tra le due superfici
- 3) Si calcola la trasformazione che minimizza tale distanza
- 4) Applica la trasformazione e reitera la procedura fin quando la distanza non sia sotto un certo threshold

3D FACE RECOGNITION

Bisogna estrarre le feature e fare l'allineamento.

Diversi tipi di recognition:

- NORMAL MAP
 - 1) acquisizione modello 3D
 - 2) proiezione da 3D a 2D
 - 3) Generazione della normal Map (immagine RGB)



Utile per salvare informazioni del volto tridimensionale in uno spazio bi-dimensionale. La difference map è la differenza tra due normal map. Ogni pixel della difference map rappresenta la distanza angolare tra le due normal map in quel punto.

- FACEGEN MODELLER

Tool che permette tramite i morphable model di creare dei modelli 3D partendo da un'immagine. E' possibile quindi emulare le espressioni facciali

- ISO-GEODESIC STRIPES

La distanza Geodetica è il percorso più breve tra due punti in uno spazio curvo, tali distanze sono influenzate dai cambiamenti delle espressioni. Le informazioni sono codificate in una rappresentazione sotto forma di grafo e il riconoscimento facciale è ridotto alla corrispondenza dei grafi. L'immagine del viso 3D è divisa in strisce facciali iso-geodetiche di uguale larghezza e distanza crescente dalla punta del naso.

Ogni stripe può essere rappresentata come un nodo di un grafo. Due nodi sono connessi tra di loro tramite 3D Weighted Walkthroughs (3DWW).

La similarità tra i due modelli avviene confrontando la distanza tra i nodi appartenenti alle stesse stripes dei due grafi.

LEZIONE 8

PROBLEMI RICONOSCIMENTO FACCIALE

a) ILLUMINAZIONE

Soggetti diversi illuminati allo stesso modo possono apparire simili.

Tre algoritmi cercano di risolvere questo problema:

1) SHAPE FROM SHADING

2) REPRESENTATION BASED METHOD

Vengono utilizzati dei classificatori che sono costruiti in modo tale da essere robusti alle variazioni di illuminazione

3) GENERATIVE METHOD

Partendo dall'immagine 3D viene generato un insieme grande di immagini con il maggior numero possibile di variazioni di luminosità.

b) POSA

La posa del soggetto durante i test potrebbe essere differente da quella catturata durante l'enrollment.

Per risolvere questo problema:

1) MultiView System

Durante l'enrollment viene salvato il soggetto in diverse posizioni

2) Pose Correction System:

Partendo da un'immagine di test il sistema prova a dedurre la posizione cercando poi di correggerla.

c) OCCLUSIONE

L'approccio più popolare per riuscire a comparare un'immagine che presenta un'occlusione è il seguente:

1) si trova l'occlusione

2) si applica una patch

Così facendo si vengono a creare delle zone che non sono d'interesse.

FERET PROTOCOL

E' stato il primo db condiviso nella comunità dei ricercatori così facendo è stato possibile valutare tutti gli algoritmi di face detection allo stesso modo.

L'obiettivo di questo protocollo era di sponsorizzare la ricerca.

- Agosto 94

Composto da 3 test

1) riconoscimento facciale

2) false alarm

3) pose variation

- Mar 95
- 1) testing effettuati su gallery di maggiori dimensioni e che contenevano doppioni

- Set 96
- 1) Match di 3000+ immagini

- FRVT

Challenge aperta anche ai privati, non solo research

I dati utilizzati erano dei dati che i researcher o i developer non avevano mai visto fino a quel momento. Così facendo hanno reso i risultati più generali.

- FRGC (Face Recognition Grand Challenge)

E' una sfida tra diversi gruppi di ricercatori. La sfida è progettata per ottenere un aumento delle prestazioni degli algoritmi di riconoscimento 2D e 3D. FRGC mirava a raggiungere in modo affidabile un tasso di verifica del 98%.

LEZIONE 8BIS

BIOMETRIC SPOOFING

E' un attacco verso il sistema biometrico. Si cerca di imitare il tratto biometrico in modo tale da fingersi qualcun altro.

CAMOUFLAGE

Il soggetto che attacca il sistema si camuffa cercando di non essere riconosciuto

ATTACCHI FACE SPOOFING

- 2D face spoofing: si presenta una foto o un video di uno sconosciuto
- 3D face spoofing: tramite delle maschere
- Operazioni chirurgiche

LIVENESS DETECTION

Obiettivo: cercare di capire se il soggetto che si sta presentando al nostro sistema è effettivamente un utente e non un'immagine.

Metodi per il detect delle foto:

- Analisi del battito degli occhi
- Le foto contengono un materiale (metallo) che riflette in maniera diversa la luce rispetto ad un viso reale quindi LBP può servire nel rilevare queste differenze. A trarne le conclusioni è un SVM classifier.
- Tramite una foto del soggetto sottomesso si è in grado di determinare se è stata sottomessa una foto perché la foto di una foto presenta una risoluzione inferiore.
- Controllare il background di un soggetto al variare della posizione della testa. Se la testa si muove il background dovrebbe rimanere fermo.

Video spoofing:

- Detect di un moirè pattern, le bande che si presentano quando viene fatta una foto a un display

Gaze Stability

- Vengono mostrati all'utente una serie di impulsi sullo schermo in modo random così facendo chi attacca non è in grado di preparare degli attacchi specifici. Ad ogni impulso vengono ripresi i landmark più importanti è tramite questi il sistema determina se è una persona o un'immagine o un video.

3D FACE DETECT

Per il detect si studia come la luce viene riflessa dal materiale della maschera oppure si possono studiare i battiti cardiaci presi dalla variazione del colore sotto la pelle.

ANTI SPOOFING FOR MOVING FACES

Un modo finale per l'anti-spoofing di video/foto è verificare la tridimensionalità del viso attraverso invarianti geometrici.

Gli invarianti geometrici sono descrittori di forma, che non sono influenzati dalla posa e dalla scala dell'oggetto, dalla proiezione prospettica e dai parametri intrinseci della telecamera. Sono espressi come rapporti di distanze / misure o come una combinazione di coordinate 3D/2D dei punti dell'oggetto.

Vengono quindi applicati i landmark a dei punti del viso che sappiamo non essere coplanari, viene quindi richiesto di muovere il viso. Se questi punti dopo la rotazione risultano essere coplanari allora si ha la certezza che stiamo analizzando una foto/video e non un volto reale.

ANTI SPOOFING EVALUATION

- SFAR: percentuale di attacchi spoof che hanno avuto successo
- FRR: Soggetti genuini non accettati
- FAR: Attacchi zero effort che hanno avuto successo
- FLR(FALSE LIVING RATE): percentuale di spoof attack che vengono rilevati come legittimi viventi (e non come foto/video)
- FFR(FALSE FAKE RATE): percentuale di accessi reale che è stata categorizzata come falsa.

CAPTCHA

Modo di rilevare se l'utente è o meno una persona. Utilizzato per non far abusare certi bot di determinati servizi

- reCaptcha: Google usa distorsioni
- imaginationCaptcha: foto con distorsioni dove viene richiesto se un determinato soggetto è presente
- Image orientation: Viene chiesto di far ruotare un'immagine fino a farla mettere nella posizione precisa
- artificial: Chiede di indicare dove sono certi soggetti
- Fatcha: Viene chiesto di fare determinate pose con il viso

LEZIONE 9

EAR

L'orecchio presenta dei dati unici, universali e permanenti e per questo è un tratto biometrico valido per l'identificazione.

LOCALIZATION

Attraverso rete neurale e localizzazione tramite AdaBoost

RECOGNITION

Simili al face recognition, prevedono l'utilizzo di PCA e Gabor Filter

LEZIONE 10

IRIDE

membrana situata dietro la cornea, davanti la lente e bucata dalla pupilla. È pigmentata e costituita da uno strato fibroso muscolare circolare. Grazie alle texture irregolari, al colore e alla pigmentazione ha un buon livello di diversificazione.

In quanto presenta pattern di texture regolari e irregolari ha un alto grado di discriminazione.

Vantaggi:

- Non cambia nel tempo

Svantaggi:

- Dimensione e acquisizione.

ACQUISIZIONE

1) LUCE VISIBILE

La membrana dell'occhio riflette la luce creando rumore

2) LUCE NON VISIBILE

La melanina riflette solamente la luce rossa permettendo una chiara identificazione della sua texture

SISTEMA DOUGMAN

Estrazione Feature.

1) LOCALIZZAZIONE E UNWRAPPING

Posizione dell'iride

questo approccio utilizza una specie di rilevatore di bordi circolari per localizzare sia la pupilla che l'iride.

Posizione delle ciglia

come prima ma rileviamo anche le linee delle palpebre.

Segmentazione dell'iride

otteniamo una maschera in modo che solo i pixel dell'iride vengano ulteriormente elaborati.

Unwrapping

Estrae la texture dell'iride normalizzandola per poter poi essere matchata con un db. Un altro motivo per cui viene normalizzata è perché la pupilla e l'iride non sono perfettamente concentrici e le dimensioni della pupilla possono cambiare.

2) FEATURE EXTRACTION AND CODING

Viene effettuata utilizzando i filtri Gabor.

NICE (Noisy Iris Challenge Evaluation)

Esperimento utilizzato per comporre il dataset UBIRIS. Il dispositivo per l'acquisizione di immagini è stato installato in una sala sia in condizioni naturali che artificiali. Sono state eseguite due sessioni di acquisizione distinte della durata di due settimane ciascuna, separate da un intervallo di una settimana. Dalla prima alla seconda sessione, sono stati cambiati sia la posizione che l'orientamento del dispositivo di acquisizione e delle sorgenti luminose artificiali.

NICE I

Contest per la valutazione delle tecniche di segmentazione dell'iride e di rilevazione del rumore.

Definiamo Alg l'algoritmo per la segmentazione dell'iride della regione rumorosa libera dell'iride, che vogliamo valutare. Due misure di valutazione dell'algoritmo: Classification error rate e Type-I and Type-II error rate. L'algoritmo vincente è quello proposto da CASIA con quattro fasi:

- 1) pre-processing: filtri per il miglioramento dell'immagine utili a correggere eventuali errori
- 2) pupil localization (eseguito da circular detection),
- 3) linearization
- 4) limbus location

Per il pre-processing ci sono alcuni fattori che influenzano negativamente il rilevamento dell'iride (ad esempio pori della pelle, ciglia). A tale scopo vengono utilizzati alcuni filtri. Il seguente schema rappresenta il sistema di riconoscimento N- IRIS.

NICE II

Obiettivo

valutazione delle strategie di codifica e abbinamento per le firme biometriche.

Ad ogni algoritmo viene sottoposto un set di iridi e vengono calcolate le rispettive mappe binarie.

Per ogni coppia di iridi nel set viene fatta calcolare la distanza tra di esse.

Attraverso le distanze calcolate si ricava una statistica che prende in esame le distanze calcolate tra le immagini inter class e intra class.

Vincitore

Casia

LEZIONE 11

IMPRONTE DIGITALI

Serie di linee scure che rappresentano porzioni rigide di pelle e che sono in rilievo rispetto agli spazi intermedi tra queste.

Formazione:

Si formano nei sette mesi del periodo fetale e sono influenzate dal liquido amniotico e dalla posizione fetale.

Unicità:

Ogni impronta digitale è micro diversa dalle altre. Anche tra gemelli le impronte sono molto simili ma presentano tratti micro diversi.

Acquisizione:

- OFFLINE

L'impronta viene passata su una tavola di inchiostro e poi viene trasferita su un foglio di carta attraverso una pressione del dito sul foglio.

Infine viene scannerizzata l'immagine.

- LIVE SCAN

Tramite un contatto diretto tra il dito e un sensore speciale

Problemi Acquisizione

- troppo movimento causa distorsione
- pressione variabile sul sensore
- errore nell'estrazione delle feature

Parametri Digitalizzazione

- Risoluzione
- Area di acquisizione
- Profondità
- Contrasto
- Distorsione geometrica.

Tipi di scanner

Optical Scanner

Poco costoso, robusto alla variazione di clima e con buona risoluzione, ma è grosso e va pulito molto bene dopo ogni utilizzo;



Capacitive Scanner

migliore risoluzione dell'impronta e dimensioni ridotte ma dura molto l'acquisizione;



Thermal Scanner

Non può essere ingannato da impronte artificiali perché riconosce pulsazione, temperatura, pori e cambiamento di colore della pelle tramite pressione, anche se l'immagine scompare rapidamente, è infatti possibile riprodurre Fake Fingerprints tramite gelatina, silicone e lattice.

BASI PER IL MATCHING

Bisogna trovare un pattern globale tra le due impronte.

Successivamente si deve trovare un certo numero di dettagli minimi tra le due impronte che devono essere uguali.

METODI PER IL MATCHING

- CORRELAZIONE
Sovrapposizione delle due impronte
- FEATURE DELLE CRESTE
Quando l'immagine presenta una bassa qualità allora conviene studiare l'orientamento e la forma delle creste
- MINUZIE
Le minuzie delle due impronte vengono estratte e posizionate in uno spazio a due dimensioni.
Si cerca quindi di sovrapporre i due spazi per determinare se il risultato finale del matching.

FEATURE EXTRACTION

Segmentation

- separare il foreground anisotropico (dipendenza direzionale) dallo sfondo isotropico (indipendenza direzionale).

Directional Map

- E' una matrice discreta dove ogni elemento denota l'orientamento delle tangenti alle creste del fingerprint.

Mappa della frequenza

- Si studia per ogni punto presente in una impronta digitale il numero di creste per unità di lunghezza.

Singularità

- Si basano sulla matrice delle direzioni e sull'indice di Poincaré
- Vengono considerate le curve intorno ad ogni elemento della matrice. Si sceglie quindi un orientamento e si calcolano le differenze tra i diversi orientamenti.
- Per ogni punto questo indice può assumere questi valori:

0 => no singularità

360 => spira

180 => loop

-180=>delta

ESTRAZIONE MINUZIE

Prevede i seguenti step:

- Binarization: si convertono i livelli di grigio dell'immagine in un livello binario;
- Thinning: riduce lo spessore delle linee a 1 pixel;
- Location: localizza i pixel corrispondenti alle minuzie;

Ridge Count: numero di linee tra due punti significativi.

RICONOSCIMENTO TRAMITE METODO IBRIDO

Questo procedimento mette insieme la rappresentazione dei fingerprint basata sulle minuzie con una rappresentazione basata sui filtri Gabor.

STEP:

- ALLINEAMENTO IMMAGINI

Inizia con l'estrazione delle minuzie sia dal template in input che dal template da matchare.

I due insiemi di minuzie vengono comparati tramite un algoritmo che prende in considerazione una coppia di minuzie (una dal template nel gallery e una dal template in input).

Questi due punti vengono sovrapposti determinando un insieme di minuzie sovrapposte.

La coppia migliore è quella che presenterà il maggior numero di sovrapposizioni delle minuzie.

Viene escluso lo sfondo dalle immagini. Immagine e template sono divise in un set di finestre in cui l'immagine e la luminosità viene normalizzata

- FEATURE EXTRACTION

Vengono estratte tramite 8 Gabor filter che presentano la stessa frequenza ma con orientamento variabile.

- MATCHING

Vengono comparati i vettori caratteristici del template e dell'immagine in input.

Vengono inoltre confrontate le minuzie ed infine tramite un threshold si determinerà il risultato del matching.

Fake Fingerprint

E' possibile emulare una impronta digitale tramite il silicone, il lattice e la gelatina. Per determinare se al nostro sistema si sta sottomettendo una fake fingerprint basterà effettuare una liveness Detection basandoci sulla pressione, pulsazione, colore della pelle dovuto alla pressione e sudore.

LEZIONE 11bis

1. **APPLE TOUCH ID:** Apple's iPhone 5s and later, and the iPad Air 2 and later, and the MacBook Pro 2016, all boast a Touch ID fingerprint sensor, which allows you to ditch your passcode in favour of your fingerprint. The first port of call in fixing an unresponsive fingerprint scanner is to clean the Touch ID sensor. This takes a high-resolution picture of the sub-epidermal layers of your skin to read your fingerprint and compare it to the fingerprint it has on file, so if there's dirt or grime on the Home button, the likelihood is your iPhone or iPad will have difficulty confirming that you really are who you say you are. You can clean the Home button with a lint-free cloth (like the kind you use to clean glasses or a tablet screen). Others problems can be: position of the fingertip on the home button, covers having cutouts for the home button quite tight, fingers not clean².
2. **APPLE FACE ID:** even if wired spends a lot of dollars for testing the robustness of FaceID, various events have occurred: it seems that FaceID is a little bit sensitive to kinship relation. Two brothers (not twins) have been able to unlock the same phone and the same things is happened with a child who unlocked his mother phone. Some videos have also shown that FaceID is subject to mask attack. Such a mask is printed in 3D, made with a mixture of crushed stone and polymers and with two-dimensional images of the eyes (crushed stone seems the secret!).
3. **FINGERPRINTS IN FORENSICS:** there are vary cases where the fingerprints are not used in a right way causing the sentence for innocents.

LEZIONE 12

MULTIBIOMETRIC SYSTEM

Vengono utilizzati molteplici tratti biometrici invece di uno singolo.

Differenti feature possono essere estratte tramite metodi diversi e possono essere fuse insieme in modo tale da creare una nuovo feature vector in grado di rappresentare l'individuo.

FEATURE LEVEL FUSION

Questa fusione può essere:

- SERIALE
Dopo un processo di normalizzazione i feature vector vengono messe in serie
- PARALLELA
Dopo un processo di normalizzazione i feature vector vengono combinati parallelamente
- CCA (Canonical Correlation Analysis)
Sono una coppia di trasformazioni lineari in grado di massimizzare la correlazione tra i coefficienti delle caratteristiche

PROBLEMI FEATURE:

- La selezione delle feature con eventuale riduzione. Si cerca quindi di escludere tutte quelle che sono le feature che presentano un basso tasso di discriminazione andando ad analizzare solo le feature rilevanti.
- Problemi dimensionalità
- Il matcher deve essere più complesso
- I vettori combinati possono presentare dati ridondanti per questo si vanno a togliere tutte quelle che sono le feature che una volta estratte si vanno a sovrapporre.

SCORE LEVEL FUSION

Per ottenere uno score finale possiamo utilizzare due approcci:

- 1) TRANSFORMATION BASED
Gli score ottenuti dal match di ogni feature vengono normalizzati e messi insieme secondo delle regole
- 2) CLASSIFIER BASED

Gli score ottenuti dai diversi classifier sono considerati come feature che vengono salvate all'interno di un feature vector.

REGOLE SCORE LEVEL FUSION

- 1) MAJORITY VOTE (basata sui risultati dei classificatori su singola label)
- 2) BORDA COUNT
Ogni classificatore ritorna il proprio ranking. Successivamente per ogni classe vengono sommati i risultati ottenuti e viene scelta la classe con il voto più alto.
- 3) MISURE
Somma, somma pesata, media...

PROBLEMI SCORE LEVEL FUSION

A supporto di queste tecniche è possibile utilizzare dei processi di normalizzazione dei risultati ottenuti.

Durante questo processo è opportuno tenere in considerazione i seguenti possibili problemi:

- 1) Robustezza: La normalizzazione non deve essere influenzata dai valori anomali
- 2) Efficacia: I parametri ottenuti dalla normalizzazione devono essere il più possibile approssimati ai valori reali

RELIABILITY

Data la qualità differente che possiamo avere dai diversi sottosistemi è opportuno stabilire delle misure di affidabilità prima di trarre la nostra conclusione.

Una possibile soluzione è quella di applicare dei margini di confidenza basati sulle stime del FAR e FRR

$$M(\Delta) = |FAR(\Delta) - FRR(\Delta)|$$

DECISION LEVEL FUSION

Ogni classificatore mostra la propria scelta (accept o reject).

La decisione finale può essere presa in diversi modi. I più semplici sono tramite gli operatori logici AND e OR. Si può anche utilizzare un sistema di majority voting.

LEZIONE 13

PIFS: Partitioned Iterated Function System. Un potente approccio basato su frattali alla compressione e all'indicizzazione delle immagini. Sfrutta e codifica le somiglianze con l'immagine. L'immagine è suddivisa in intervalli quadrati non sovrapposti chiamati regioni.

FARO: Face Recognition Against Occlusions

FACE - Face Analysis for Commercial Entities

1. Rilevazione del viso con Viola Jones 2. Vengono rilevati occhi naso e bocca 3. La rotazione del viso viene corretta utilizzando la posizione degli occhi 4. La faccia è divisa simmetricamente in sinistra e destra

5. La regione del viso R con una migliore illuminazione viene preservata mentre l'altra viene scartata

6. R viene allungato in modo tale che i punti chiave predefiniti rientrino in posizioni predefinite e infine l'immagine ottenuta venga specchiata sull'altro lato per creare una faccia intera.

FOVEA: video Frame Organizer Via Identity Extraction and Analysis

Traccia i volti in un video. Abbiamo due tipi di identità: temporaneo (appartiene a un soggetto rilevato in fotogrammi consecutivi) e permanente (identità che sono state trovate nell'intero video, sono distinte e a ciascuna può corrispondere un insieme di modelli che rappresentano il volto con espressioni diverse). Per tracciare le facce in un video, il sistema mantiene una serie di facce attualmente esaminate e lo aggiorna fotogramma per fotogramma.

HERO: Human Ear Recognition against Occlusions. L'immagine è divisa in quattro sottoregioni (quadrati in alto a sinistra, in basso a sinistra, in alto a destra e in basso a destra) che chiamiamo riquadri. Al fine di rendere il sistema robusto rispetto alle modifiche locali, i quattro riquadri sono indicizzati in modo indipendente mediante i sistemi PIFS come regioni di interesse separate.