Sistemi Biometrici basati sul Volto

Parte VI

Indice

Capitolo 1

Introduzione

I sistemi biometrici basati sul volto vengono impiegati sia per la **verifica** dell'identità, che per l'**identificazione**; si possono suddividere in:

- 2D
  - immagine fissa/video/tante immagini fisse
  - colori/toni di grigio
- 3D
  - scansione laser
  - illuminazione controllata

1.1 Vantaggi

- Ottimo compromesso tra accettabilità dell'utente e prestazioni di accuratezza
- Dispositivi di input adatti a molte condizioni operative
- Risultati direttamente verificabili da un operatore umano
- Permette l'acquisizione non consenziente (video-sorveglianza)

1.2 Svantaggi

- Soffrono di elevata variabilità intraclasse
  - posa
  - illuminazione
  - variazioni dell'aspetto dell'individuo (dimagrimento, invecchiamento, ...)

  - occlusioni (capelli, occhiali, ...)
- Possibile ingannare i sensori non dotati di sistemi anti-spoofing
- Solo di recente le accuratezze sono diventate interessanti per sistemi di larga scala

1.3 Non solo identificazione/verifica

Molti degli algoritmi servono anche per:

- riconoscimento delle espressioni facciali
- riconoscimento del movimento delle labbra
- applicazioni di computer grafica
- creazioni di protesi

1.4 Principali fasi

La catena di enroll o verifica/identificazione consiste solitamente nei seguenti passi:

1. **Face Detection** (trovare i volti nelll'immagine)
2. **Face Segmentation** (separare il volto dallo sfondo)
3. **Face Tracking** (se in un video, deve essere inseguito)
4. **Face Normalization** (crop e normalizzazione dell'immagine)
5. **Feature Extraction**
6. **Matching**

Le tecniche di deeplearning hanno permesso importanti passi in avanti, specialmente nelle fasi di *detection*, *segmentation*, *feature extraction* e *matching*.

1.5 Intraclasse e interclasse

1.5.1 Variabilità intraclasse

Moltissimi fattori possono contribuire negativamente sull'accuratezza dei sistemi basati sul volto, aumentando la variabilità intraclasse, tra cui:

- illuminazione
- espressioni del volto

- posa
- occlusioni
- sensori
- invecchiamento

Le differenze di illuminazione e posa sono quelle con i maggiori impatti sulle performance.

1.5.2 Similitudine interclasse

Similitudine tra volti, ad esempio nel caso di:

- gemelli
- fratelli, genitori (distanti nel tempo)
- sosia

Contributo della genetica

Solo una porzione tra 0,1% e 0,5% è diversa per ogni individuo, tutto il resto è uguale per tutti gli esseri umani.
Tra padre e figlio **la differenza viene dimezzata**.

1.6 Standard

1.6.1 Standard per l'acquisizione

L'ICAO (*Organizzazione Internazionale per l'Aviazione Civile*) e moltissimi governi hanno iniziato a dettare regole stringenti per l'acquisizione del tratto biometrico (per passaporti, ...).
L'obiettivo è quello di arrivare ad impiegare sistemi sempre più accurati.

1.6.2 Standard per il template

Tutti i venditori di sistemi biometrici hanno il loro algoritmo per la creazione del template, spesso **segreto** e/o **sotto brevetto**.

→ **non è possibile interoperabilità** tra i sistemi (caso opposto a quello delle impronte digitali nel caso delle minuzie) a meno che vengano condivise le foto originali e non i template.
L'ICAO impone anche degli standard sul formato delle foto per i *Machine Readable Travel Document (MRTD, ad esempio ePassport)*, tra cui:

- risoluzione dell'immagine
- dimensione approssimativa dell'immagine

1.7 Trend di ricerca

- Affrontare ambienti *unconstrained*
  - *outdoor facial images*
  - *non-frontal facial images*
- Affrontare grandi numeri
  - migliorare i tassi di falsi positivi
  - ...
- Miglioramenti teorici
  - Il deeplearning continua ad essere uno dei più promettenti

Capitolo 2

Algoritmi di prefiltraggio ed estrazione di caratteristica

2.1 Face Detection

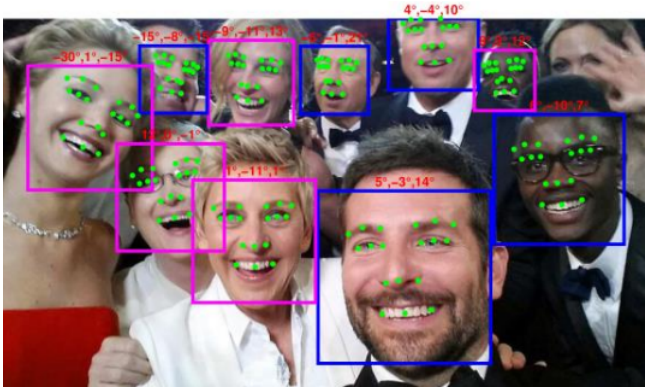
È il primo passo della catena, ha il compito di rintracciare il volto/i senza nessun prerequisito particolare; possono cambiare:

- condizioni di luce
  - intensità
  - direzione
  - banda spettrale
- i volti
  - colore
  - posizione
  - scala
  - posa
  - espressione

Molti approcci sono basati su modelli molto semplici del volto, in termini

- geometrici
- di texture

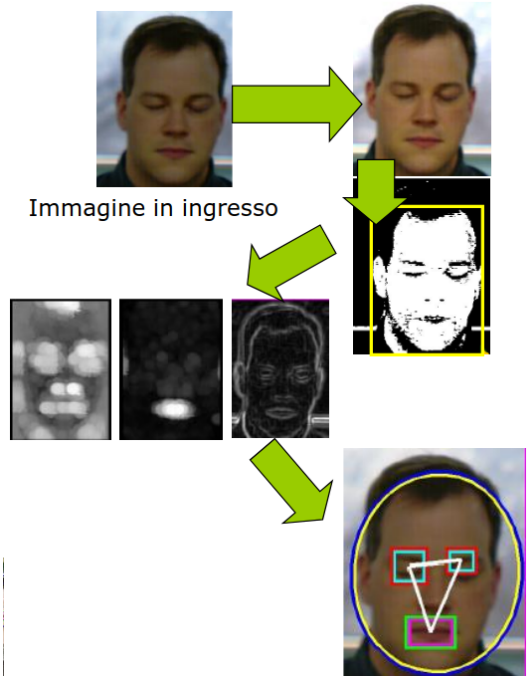
Questi modelli permettono di trovare nelle immagini le regioni dove ***fittano* meglio il modello** e quindi dove **è maggiore la probabilità che ci sia un volto**.



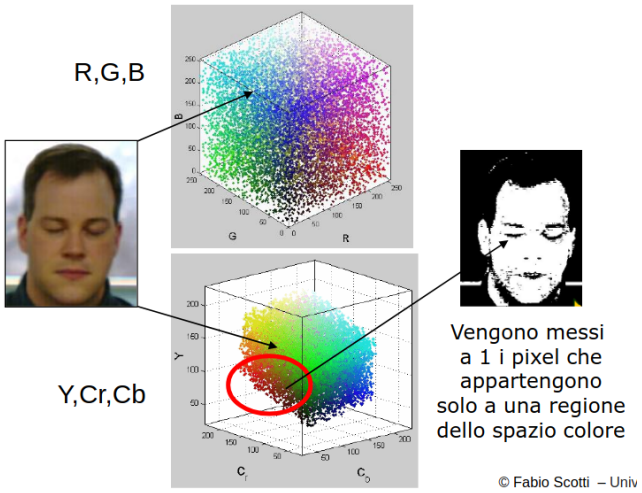
2.1.1 Color based

Nel caso delle immagini a colori uno degli schemi più diffusi è il seguente:

1. lighting compensation
2. skin tone detection
3. localization of facial features (occhi, bocca, contorno del viso)
4. aggregazione dei risultati



La skin tone detection viene effettuata controllando quali bit dell'immagine appartengono ad una determinata regione dello spazio colore che evidenzia meglio le differenze.



La rilevazione della posizione degli occhi e della bocca viene effettuata sottraendo immagini espresse in appositi spazi colore che ne possano mettere in evidenza le **caratteristiche cromatiche peculiari**.

2.1.2 Haar feature

Il metodo delle feature basate su funzioni di Haar consiste nel cercare nelle immagini delle regioni che presentano particolari vaori delle *trasformate di Haar*; solitamente i **i volti sono caratterizzati da particolari valori** e quindi si possono trovare tramite dei classificatori.

2.2 Face tracking

Quando si ha a che fare con un video si deve svolgere il face tracking; non è esattamente come fare del face detection in ogni frame, dato che:

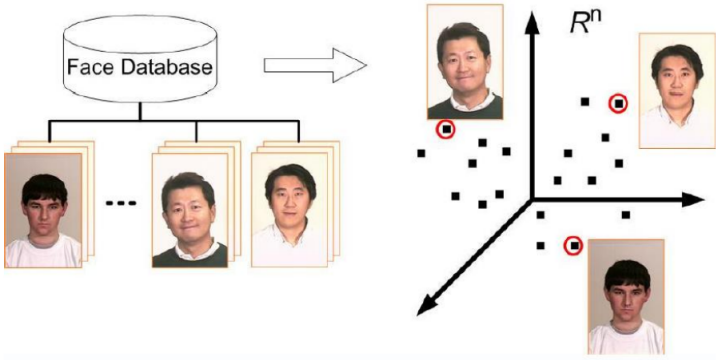
- si hanno maggiori informazioni da un filmato (i volti non possono essersi spostati di molto da un frame all'altro)
- avendo due frame è possibile effettuare una previsione sullo spostamento del volto o sulle sue deformazioni

Capitolo 3

Estrazione di caratteristica e matching

3.1 Rappresentazione dei volti

Ogni immagine $p * q$ pixel viene rappresentata come un punto nello spazio R^n, con $n = p * q$
Si crea così lo **spazio delle facce**.



→ problema: è uno spazio ad altissima dimensionalità

3.1.1 Analisi lineare di un sottospazio delle facce

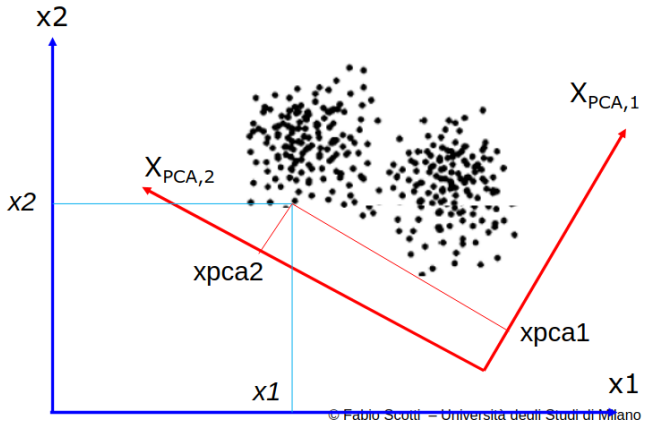
Lo spazio delle facce costruito con le immagini raw ha una dimensionalità troppo elevata; le facce possono risiedere in uno sottospazio limitato

→ il modo di stabilire il sottospazio crea il metodo di analisi

Bisogna trovare una **trasormazione lineare che mappa il vettore di immagini X in un sottospazio** di dimensione l molto più limitata ($l << n$); il match avverrà sulle feature delle immagini ottenute.

3.2 Analisi lineare di un sottospazio mediante PCA

La tecnica PCA (*Principal Component Analysis*) cerca una rotazione dello spazio (ancora la dimensionalità non cambia) che **permette di separare meglio le classi**.



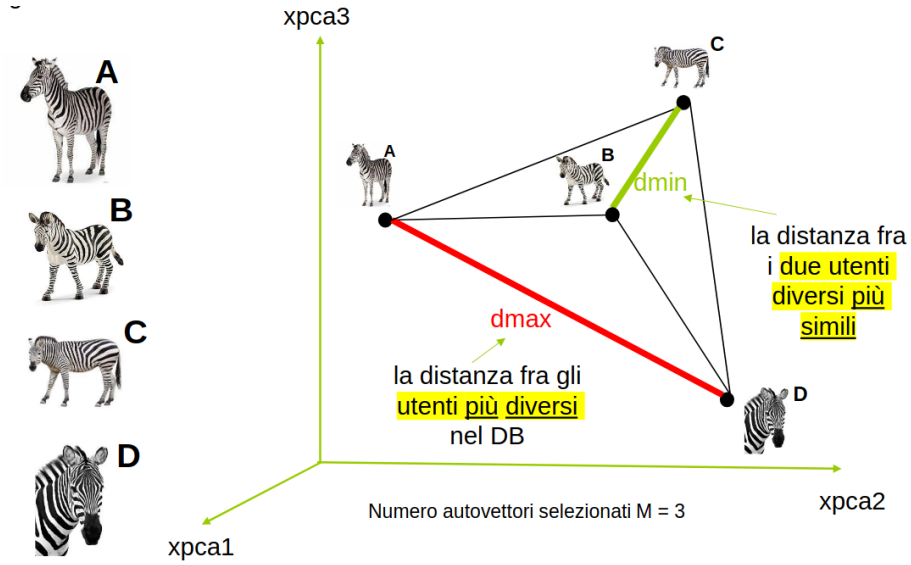
È possibile passare da un sistema di coordinate ad un altro tramite una moltiplicazione tra matrici.

→ **PCA cerca una rototraslazione** che permette di mettere la variabilità dei dati tutta nei primi parametri.

3.2.1 Vantaggi e Svantaggi

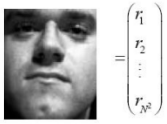
- **Vantaggi**
  - compatta la maggior parte della varianza nel minor numero di coefficienti di trasformazione
  - riduce lo spazio dei dati in ingresso
  - minimizza errore medio tra i dati ricostruiti e quelli originali
  - minimizza entropia totale dei dati costruiti
- **Svantaggi**
  - non ci sono algoritmi veloci per la sua implementazione
  - è costoso in termini di risorse computazionali

3.2.2 Creazione DB e distanze tra utenti diversi



3.2.3 Arrivo di una faccia nuova - Spazio PCA

- Arriva un faccia nuova da identificare, ecco i passi:



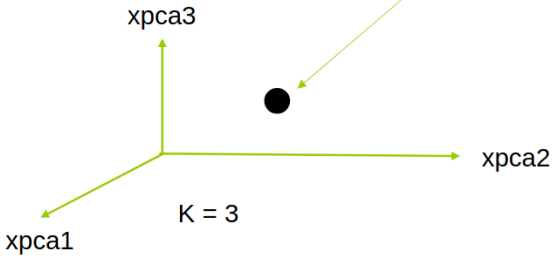
$$= \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_{N^2} \end{pmatrix}$$

$$\vec{r}_m = \begin{pmatrix} r_1 - m_1 \\ r_2 - m_2 \\ \vdots \quad \vdots \\ r_{N^2} - m_{N^2} \end{pmatrix}$$

1. sottraiamo la media

2. proiettiamo la faccia nello spazio della PCA

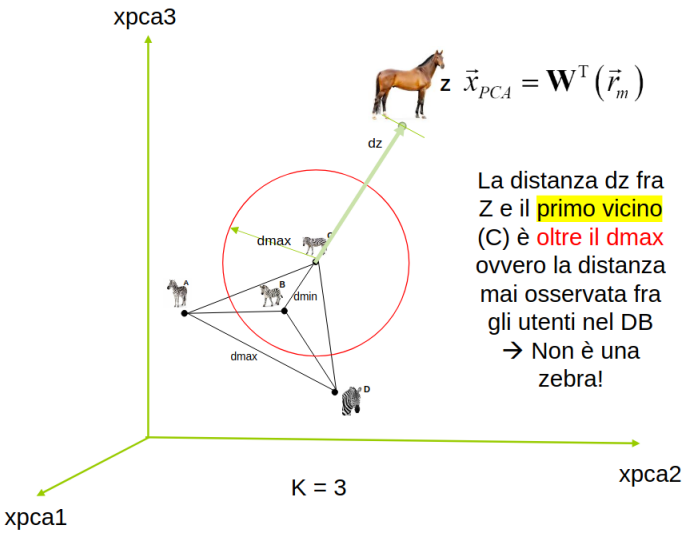
$$\vec{x}_{PCA} = \mathbf{W}^{\mathrm{T}} \left(\vec{r}_m \right)$$



Immagini in Enrollment

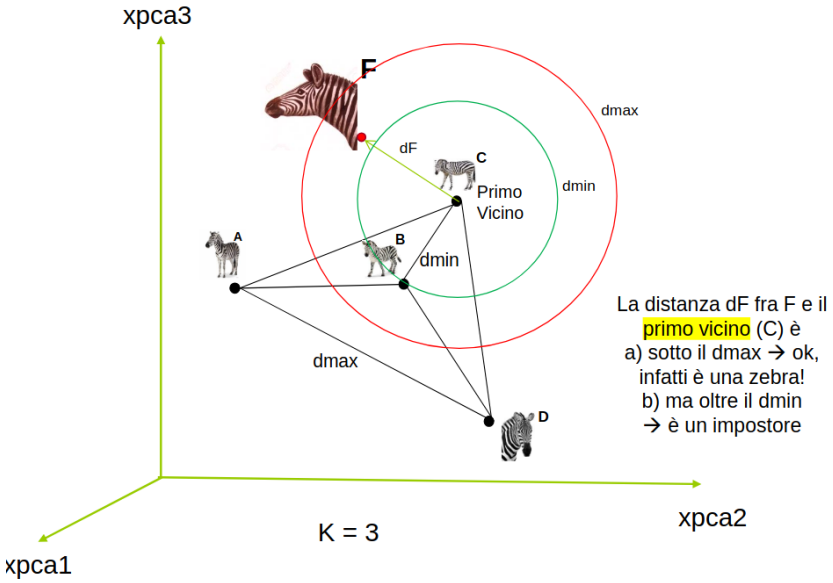
Immagine nuova

PCA: Immagine nuova

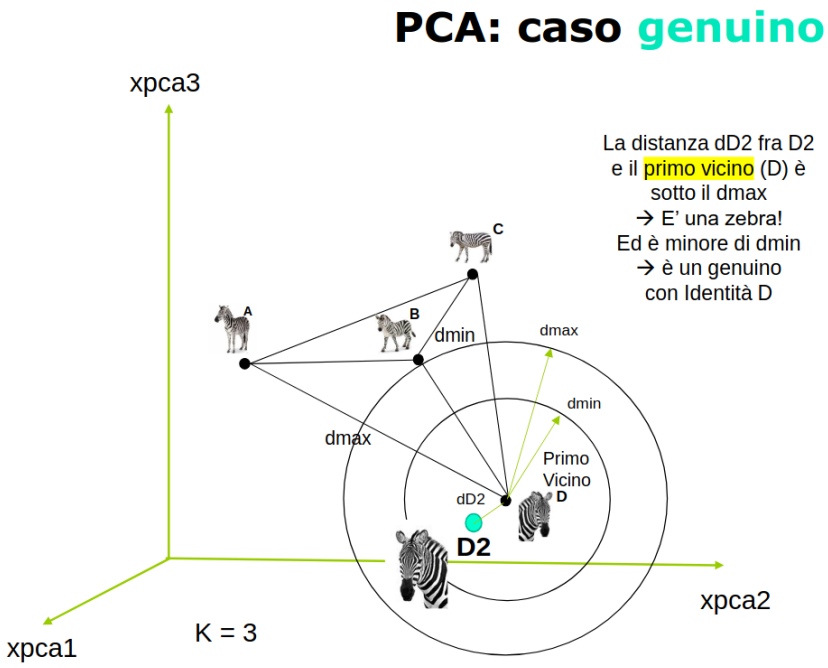


La distanza dz fra Z e il primo vicino (C) è oltre il dmax ovvero la distanza mai osservata fra gli utenti nel DB → Non è una zebra!

PCA: caso impostore



La distanza dF fra F e il primo vicino (C) è
a) sotto il dmax → ok, infatti è una zebra!
b) ma oltre il dmin → è un impostore



3.2.4 Spazio delle facce

All'arrivo di una nuova faccia, la possiamo classificare usando anche lo spazio delle facce:

1. Ricostruiamo la faccia usando le autofacce
2. Calcoliamo la distanza della faccia ricostruita da quella iniziale

L'errore indica quanto **l'immagine iniziale era vicina alle immagini con cui abbiamo calcolato le autofacce**; quindi quanto probabilmente potrebbe essere genuino o un impostore

- se ricostruisco bene l'immagine nuova (l'errore è basso) allora probabilmente era un'immagine simile a quelle memorizzate
- se il volto è troppo diverso rispetto a tutti quelli del DB del autofacce (l'errore è alto), allora è un impostore

3.2.5 Significato delle autofacce

- **Enroll:** dai volti si calcolano le autofacce con la tecnica PCA
- **Verification:** Ogni volto del DB può essere ricostruito esattamente con una combinazione lineare di autofacce; se un nuovo volto viene ricostruito male, allora:
  - non è un volto
  - è un impostore

Considerazioni sulle autofacce

Sono una tecnica base, che è stata nel tempo raffinata per far fronte agli evidenti svantaggi:

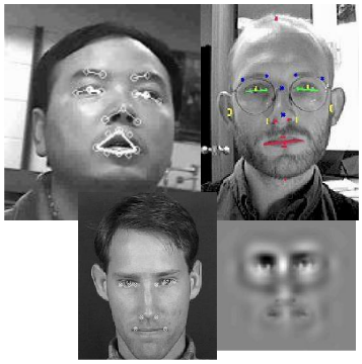
- non usano l'informazione di classe (apprendimento non superivisionato), ma lavorano solo sulla distribuzione dei punti nello spazio delle facce
- soffrono delle variazioni di
  - illuminazione
  - posa
  - allineamento
  - espressioni facciali

3.3 Altri metodi di estrazione delle feature e matching

3.3.1 Metodi di analisi locale

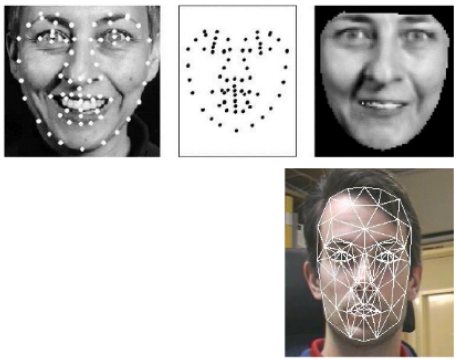
I metodi di analisi locale lavorano sul riconoscimento, misurazione e confronto dei dettagli del volto, come ad esempio:

- misure del naso, bocca, distanza tra gli occhi, ...
- valori ritornati da particolari funzioni di trasformazione



3.3.2 Metodi *model-based*

I metodi *model based* estraggono non una feature singola, ma adattano un modello sulla faccia trovandone i coefficienti per poi confrontarli con quelli delle altre facce.



Graph Matching

Nei metodi di *graph matching* il volto è rappresentato come una rete di nodi sui punti peculiari del volto; servono **almeno due immagini per trovare il modello del volto** (il grafo).
La comparazione di due facce viene eseguita misurando quanto *sforzo* è necessario fare per adattare un grafo all'altro *tirando* i nodi, che sono legati ad un modello elastico.

Capitolo 4

Spoofing e Anti-Spoofing

4.1 Limiti oggettivi e tipi di attacco

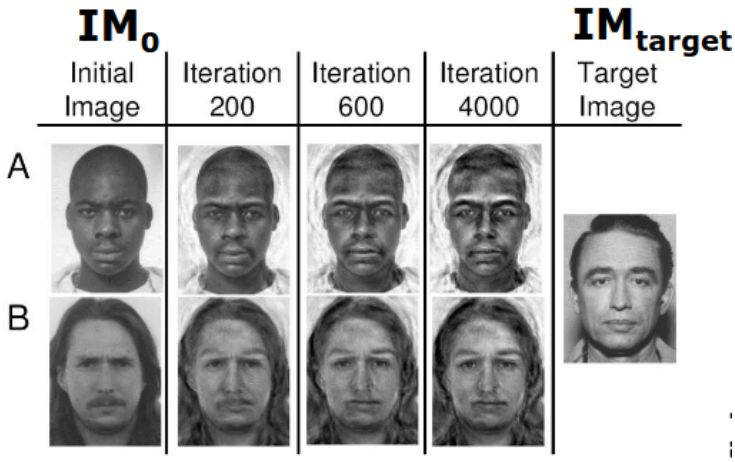
- Il volto può essere coperto da peli o indumenti durante le fasi di enrollment o verification
  - non accettare acquisizioni se la percentuale coperta supera quella necessaria per garantire l'accuratezza certificata dal sistema
- Maschera di pelle artificiale
  - tecniche termografiche e spettrometriche a diverse lunghezze d'onda per rilevare la pelle
- Immagine di un volto stampata o proiettata con uno schermo
  - controllo tridimensionalità
  - controllo sincronia tratti biomerici indipendenti (esempio variazioni di volto e voce mentre si parla)

4.2 Attacco Hill Climbing

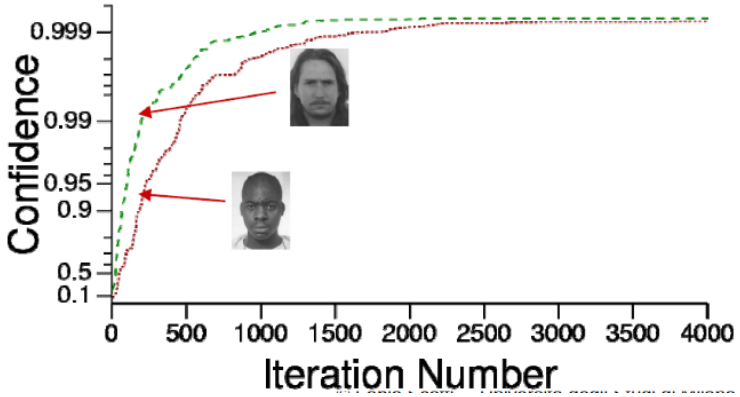
È possibile ricreare un ***sample*** *da dei* ***template*** *memorizzati in un sistema?*
→ **Sì!** Basta avere accesso al match score
La tecnica ***Hill Climbing*** prevede di:

1. iniziare con un sample casuale
2. portare delle piccole modifiche che incrementino il match score
3. proseguire finché si riesce ad aumentare il match score

Per velocizzare il processo si parte da un database locale di immagini (accessi trafugati); l'obiettivo è quello di manipolare l'immagine di partenza per farla *matchare* con un'immagine target ingannando il sistema.



Il problema è che qualunque sia il valore di soglia settato . . . basta proseguire con le iterazioni e *prima o poi si entra!*



4.3 Metodi di controllo

4.3.1 Controllo della sincronia parlato - movimenti delle labbra

Il sistema chiede all'utente di pronunciare di fronte al sensore una frase (casuale) richiesta dal sistema.
Il sistema misura la variazione dei movimenti delle labbra nel tempo controllando se si sta cercando di ingannare il sistema, ad esempio con una maschera o con una immagine proiettata.

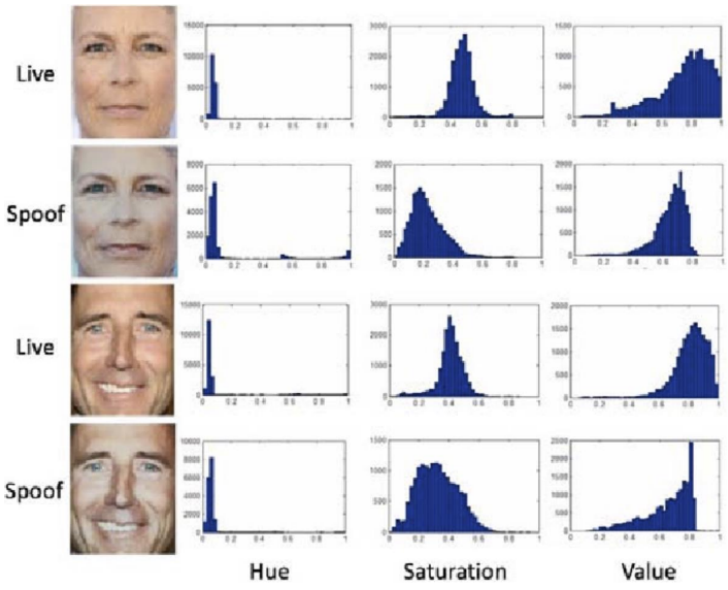
4.3.2 Facial Fingerprinting

La mappa di vene e capillari al di sotto del volto viene rilevata da telecamere ad infrarosso ad alta risoluzione, diventando un nuovo tipo di tratto biometrico:

- unico
- difficilmente falsificabile
- non intrusivo
- accurato

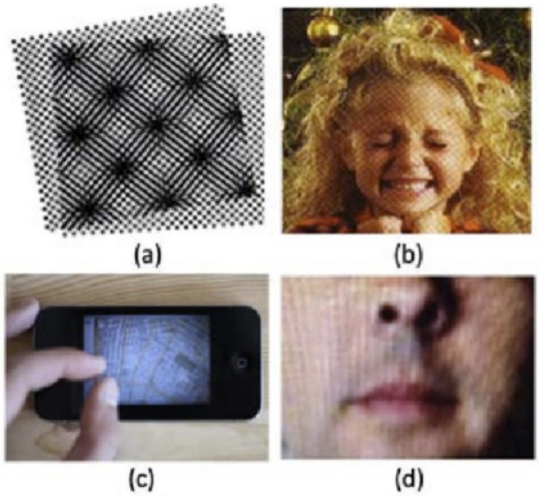
4.3.3 Controllo istogrammi del colore

In caso di attacchi con immagini stampate o proiettate, il controllo avviene controllando la distribuzione dei colori.



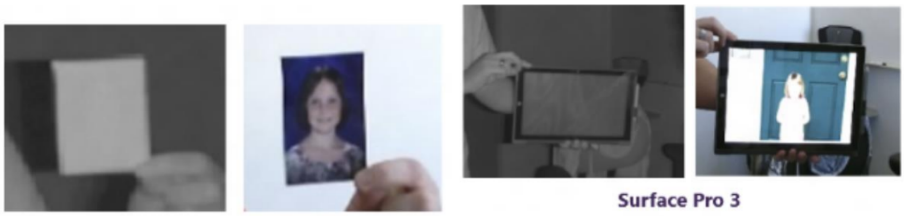
4.3.4 Controllo di pattern

Quando con una camera si inquadrano delle stampe o display avvengono dei particolari pattern; si possono vedere direttamente nell'immagine inquadrata o nella *trasformata di Fourier*.



4.3.5 Infrared Anti-Spoofing

Le telecamere ad infrarossi offrono un punto di vista diverso e più robusto alle condizioni ambientali.
Inoltre, i display o le stampe diventano non usabili con camera IR.



Analisi 3D

I dispositivi in grado di scansionare il volume degli oggetti possono migliorare le capacità anti-spoofing.