

RICONOSCIMENTO E TRACCIAMENTO DELLE PERSONE

Introduzione

Il riconoscimento automatico delle persone indica il processo che si propone di determinare l'identità di una persona utilizzando informazioni relative a un gruppo di soggetti.

In generale, il riconoscimento può essere classificato in *autenticazione* (o *verifica*) e *identificazione*.

Nell'*autenticazione* una persona deve rivendicare la propria identità la quale viene poi verificata utilizzando informazioni memorizzate precedentemente su tale persona (esempi: PIN (Personal Identification Number), immagine del volto). Viene quindi effettuato un confronto 1:1.

L'*identificazione* viene eseguita in quelle applicazioni dove una persona sconosciuta deve essere identificata in un insieme di persone senza alcuna identità rivendicata (confronto 1:n).

I sistemi di autenticazione hanno semplificato la vita quotidiana in molti modi. Con una piccola carta di plastica si possono ottenere contanti in quasi ogni angolo della città o pagare direttamente beni e servizi. L'accesso a questa fonte di comfort finanziario, tuttavia, deve essere protetto dall'uso non autorizzato. Lo scambio di informazioni via Internet e via e-mail ha reso

RICONOSCIMENTO E TRACCIAMENTO DELLE PERSONE

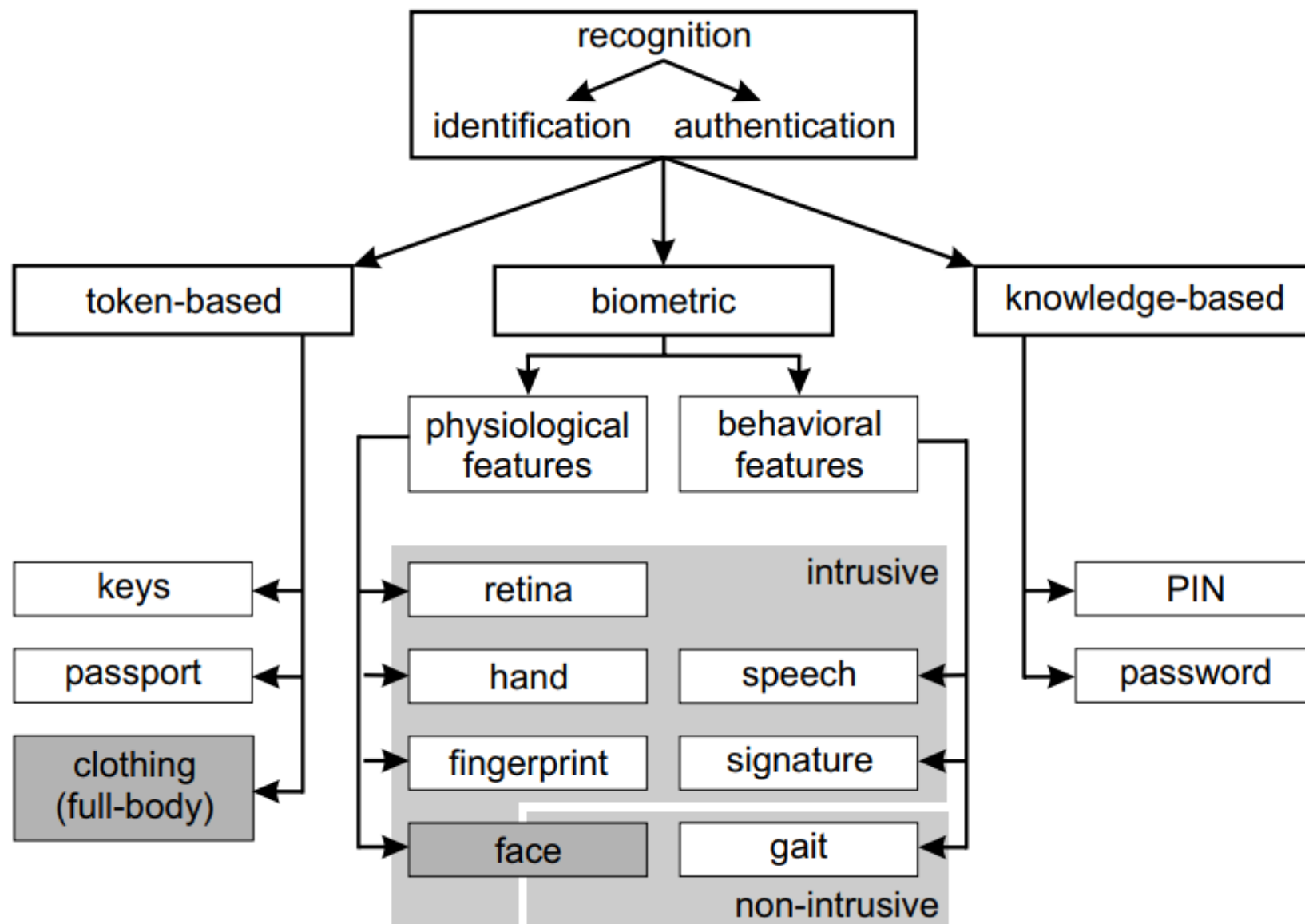
Introduzione

necessaria la personalizzazione delle informazioni, che devono essere protette in modo da non permettere accessi non autorizzati. Occorrono quindi password e codici PIN per accedere agli account e-mail o a pagine web personali. Proprio come il World Wide Web, anche le reti di computer sono protette da password.

Oltre che nella sicurezza, l'autenticazione ha un ruolo importante nell'interfacciamento uomo-macchina. Per rendere le applicazioni adattabili al contesto di utilizzo, è essenziale conoscere l'utente, i suoi spostamenti e le sue attività. Il tracciamento della persona («person tracking») deve essere applicato in combinazione con l'autenticazione della persona al fine di acquisire informazioni adeguate per i sistemi interattivi e di assistenza.

L'identificazione viene solitamente eseguita in applicazioni di sorveglianza in cui le persone da riconoscere non partecipano attivamente al processo di riconoscimento.

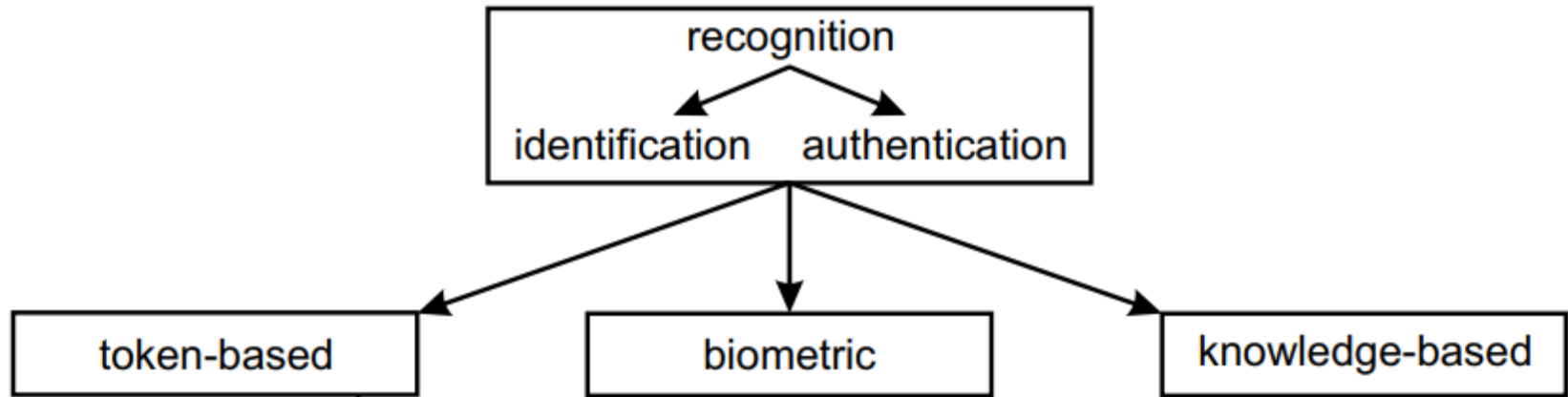
Nel seguito si analizzeranno tecniche di riconoscimento, le quali possono essere applicate sia per scopi di identificazione sia per scopi di autenticazione.



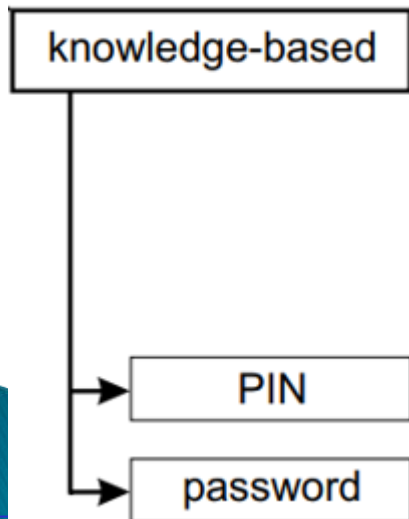
Metodi di riconoscimento delle persone. Le tecniche con rettangoli grigi verranno analizzate nel seguito.

RICONOSCIMENTO E TRACCIAMENTO DELLE PERSONE

Introduzione



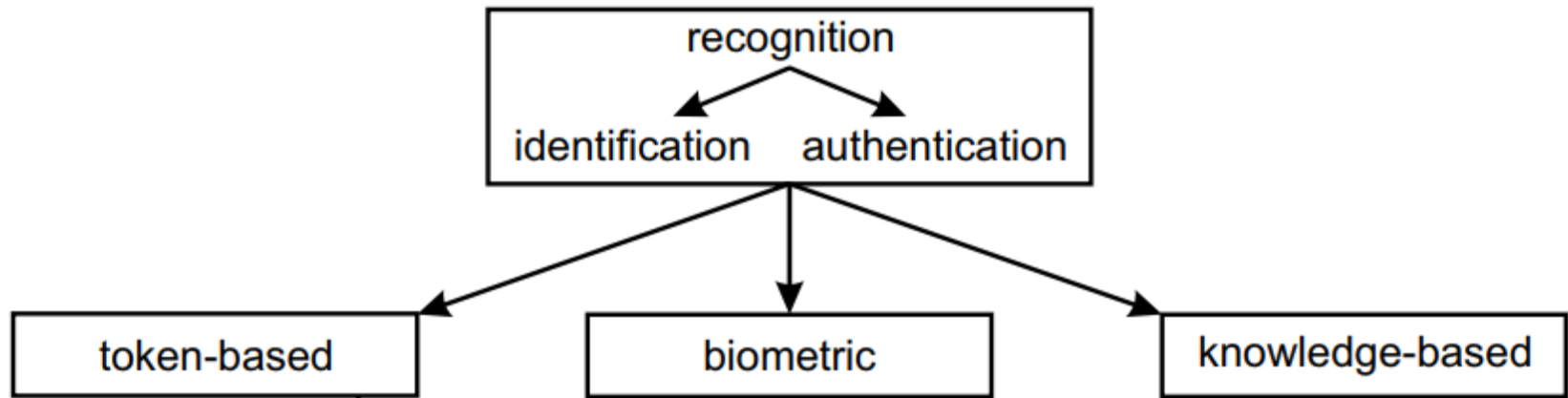
I metodi di riconoscimento possono essere classificati in tre categorie: *token-based* (chiavi, tessere, carte), *knowledge-based* (password, PIN) e riconoscimento *biometrico*.



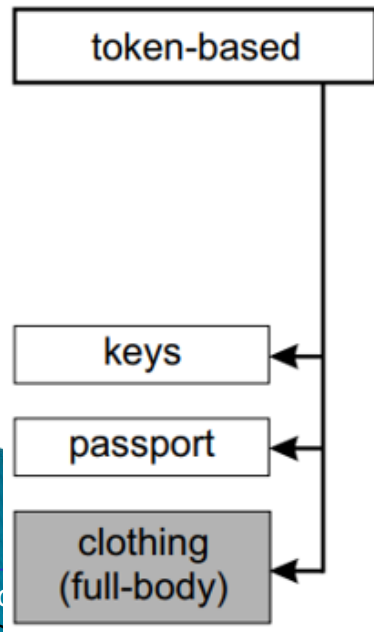
Il riconoscimento *knowledge-based* si basa su una tipologia di informazione che è unica e che identifica una persona quando viene immessa nel sistema di identificazione.

RICONOSCIMENTO E TRACCIAMENTO DELLE PERSONE

Introduzione



I metodi di riconoscimento possono essere classificati in tre categorie: *token-based* (chiavi, tessere, carte), *knowledge-based* (password, PIN) e riconoscimento *biometrico*.



Il riconoscimento *token-based* richiede il possesso di un oggetto identificativo, come ad esempio una chiave, una carta magnetica o una tessera.

RICONOSCIMENTO E TRACCIAMENTO DELLE PERSONE

Introduzione

Il vantaggio dei sistemi di autenticazione *knowledge-based* e *token-based* deriva dal fatto che la loro implementazione è semplice. Inoltre essi sono caratterizzati da facilità di utilizzo e sono molto diffusi. Tuttavia, i *token* possono essere persi o rubati, mentre le password e i PIN possono essere dimenticati. In tali situazioni, l'accesso al sistema non è più possibile o comunque richiede lavoro aggiuntivo e causa costi aggiuntivi.

Per tale motivo sono nati i sistemi biometrici che si basano su features fisiologiche (esempi: modello del volto, modello della forma della mano, modello delle vene della mano, impronte digitali, modello della retina) o dipendono da features comportamentali che sono uniche per ogni persona da riconoscere (cammino, tratto della firma, parlato).

RICONOSCIMENTO E TRACCIAMENTO DELLE PERSONE

Introduzione

I sistemi biometrici sono per lo più invasivi perché è richiesta una minima partecipazione attiva della persona da identificare. Ad esempio la persona deve appoggiare il pollice sul sensore di impronte digitali o deve guardare una camera o uno scanner per permettere l'estrazione delle features biometriche necessarie. Questa condizione è accettabile in molte applicazioni.

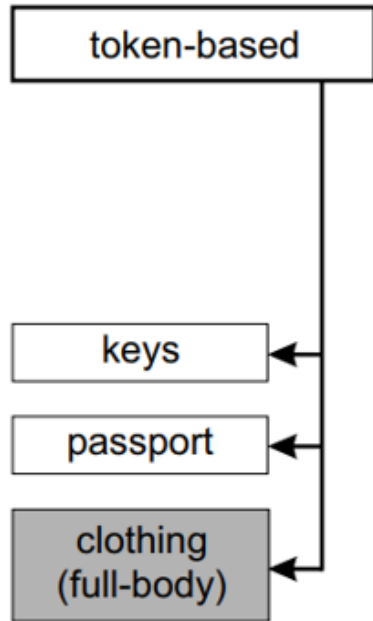
Il riconoscimento non invasivo può essere applicato in operazioni di sorveglianza in cui i soggetti sono solitamente osservati da telecamere. In questo caso, non è auspicabile che le persone partecipino al processo di riconoscimento.

Il secondo vantaggio del riconoscimento non invasivo rispetto a quello invasivo è la maggiore comodità d'uso; ad esempio, se una persona può essere identificata mentre si dirige verso una porta di sicurezza, non è necessario mettere il pollice sul sensore di impronte digitali o inserire una carta magnetica nella fessura.

Il riconoscimento facciale viene molto utilizzato in sistemi che controllano l'accesso ad aree vietate, ad esempio negli aeroporti e nelle stanze ad alta sicurezza, o anche nei computer. Per quanto riguarda il riconoscimento facciale esistono differenti approcci che si basano su determinate ipotesi.

RICONOSCIMENTO E TRACCIAMENTO DELLE PERSONE

Introduzione



Un metodo *token-based* non invasivo è il riconoscimento basato su immagini del corpo intero (full-body); tale metodo può essere utilizzato ad esempio in applicazioni in cui si può ipotizzare che l'abbigliamento non cambi. La grande varietà di colori e modelli rende l'abbigliamento un indicatore prezioso che aggiunge informazioni a un sistema di identificazione, anche se gli stili della moda e i cambiamenti stagionali possono influenzare la condizione di «abbigliamento costante».

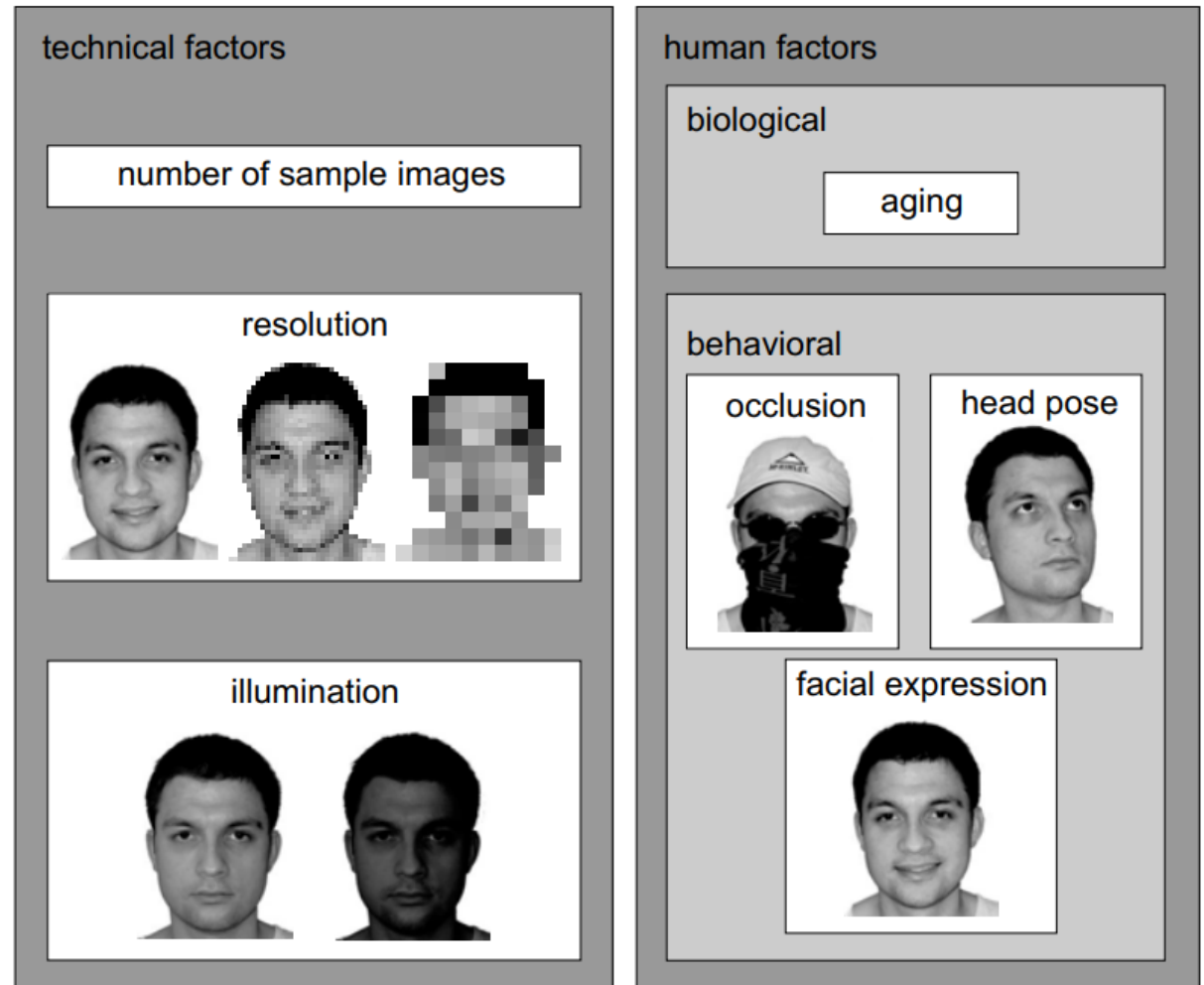
Come il riconoscimento full-body, anche il riconoscimento del cammino dipende da un adeguato rilevamento e dalla segmentazione delle persone in una scena. Tali problemi sono alcuni dei problemi principali relativi al tracciamento delle persone basato su video. Pertanto, il tracciamento delle persone può essere considerato come una fase di pre-elaborazione per questi metodi di riconoscimento. Inoltre, determinare la posizione delle persone in una scena osservata è un compito fondamentale per le applicazioni di sorveglianza in edifici, centri commerciali, aeroporti e altri luoghi pubblici.

RICONOSCIMENTO E TRACCIAMENTO DELLE PERSONE

Riconoscimento Facciale

Per gli esseri umani, il volto è la parte del corpo più discriminante per determinare l'identità di una persona. Molti fattori influenzano l'aspetto del volto. In molte applicazioni non è necessario processare tutti questi fattori.

Classificazione di alcuni dei fattori che influenzano le prestazioni dei sistemi di riconoscimento del volto. Inoltre, occorre menzionare l'aspetto (comportamentale) e i volti gemelli (biologici).



RICONOSCIMENTO E TRACCIAMENTO DELLE PERSONE

Riconoscimento Facciale

La sfida principale del riconoscimento automatico del volto (non invasivo) consiste nel riuscire a diminuire l'influenza dei fattori che riducono le prestazioni dei sistemi di riconoscimento associati. In base all'applicazione e all'ambiente, alcuni fattori possono avere una priorità minore (ad esempio i cambiamenti di illuminazione quando è disponibile un'illuminazione abbastanza costante) e altri possono essere molto importanti (ad esempio il fatto di avere a disposizione una sola immagine per persona nella fase di training). Questi fattori possono essere classificati in fattori umani e tecnici (si veda l'immagine riportata nella slide precedente).

Fattori tecnici (technical factors)

I fattori tecnici derivano dalla tecnologia impiegata e dall'ambiente in cui il sistema di riconoscimento viene utilizzato. Qualsiasi processo di riconoscimento di oggetti basato su immagini deve tenere conto di questi fattori. La loro influenza può spesso essere facilmente ridotta scegliendo hardware adeguato, come la luce (per l'illuminazione) e le camere (per l'illuminazione e la risoluzione).

RICONOSCIMENTO E TRACCIAMENTO DELLE PERSONE

Riconoscimento Facciale

Fattori tecnici (technical factors)

- *Risoluzione dell'immagine*: la scelta della risoluzione dell'immagine necessaria è un compromesso tra le prestazioni di elaborazione desiderate (\Rightarrow bassa risoluzione) e la separabilità delle features utilizzate per il riconoscimento (\Rightarrow alta risoluzione). È consigliabile utilizzare la risoluzione più alta possibile che consenta di rispettare i limiti di tempo definiti.
- *One-sample problem*: molti metodi sono in grado di estrarre buone descrizioni del viso di una persona da un insieme di immagini scattate in condizioni diverse, ad esempio illuminazione, espressioni facciali e posizioni. Tuttavia, spesso non sono disponibili diverse immagini della stessa persona. A volte, per addestrare un sistema di riconoscimento si possono usare poche immagini, o addirittura una sola. In questi casi, si deve utilizzare la conoscenza generica a-priori (ad esempio *head models* e database di immagini di volti di altre persone) in modo da modellizzare le variazioni.

RICONOSCIMENTO E TRACCIAMENTO DELLE PERSONE

Riconoscimento Facciale

Fattori tecnici (technical factors)

- *Illuminazione*: le sorgenti di luce (caratterizzate da differenti direzioni e da differenti intensità) e gli oggetti tra la sorgente di luce e il volto osservato fanno apparire lo stesso volto in modi diversi a causa delle ombre proiettate e delle regioni (sovra)illuminate. Non è possibile rilevare né le features importanti del volto né l'intero volto, oppure le features non descrivono correttamente il volto; in tali condizioni si verifica una classificazione errata o un rifiuto dell'immagine. Esempi di approcci che possono essere utilizzati per risolvere tali problemi sono:
 - utilizzare features e classificatori invarianti rispetto all'illuminazione;
 - modellizzare i cambiamenti di illuminazione e pre-elaborare l'immagine considerando i modelli ottenuti. Tali modelli devono integrare molta conoscenza a-priori acquisita da immagini di volti caratterizzate da identità e illuminazione variabili (eventualmente anche in posizioni diverse e con diverse espressioni facciali). Tali immagini spesso non sono disponibili (*one-sample problem*).

RICONOSCIMENTO E TRACCIAMENTO DELLE PERSONE

Riconoscimento Facciale

Fattori umani (human factors)

I fattori umani influenzano il processo di riconoscimento a causa delle speciali caratteristiche e proprietà del volto come oggetto da riconoscere. I fattori umani possono essere classificati in fattori comportamentali (*behavioral*) e biologici (*biological*).

Fattori comportamentali: poiché la testa umana non è un oggetto rigido e statico, essa può subire alcuni cambiamenti dovuti a movimenti intenzionali e non intenzionali e inoltre può subire alcune variazioni (esempio: indossare gli occhiali). Alcuni fattori comportamentali sono le *espressioni facciali*, la *posizione della testa*, l'*occlusione* e i *cambiamenti dell'aspetto*.

Oltre al parlato, le diverse *espressioni facciali* sono la fonte di informazione più importante per la comunicazione attraverso il volto. Tuttavia, la capacità di modificare l'aspetto del volto rappresenta un grande problema per i metodi di elaborazione e riconoscimento delle immagini. Decine di muscoli facciali permettono un movimento molto sottile della pelle, in modo da far apparire rughe e rigonfiamenti e da spostare o distorcere regioni associate a texture importanti.

RICONOSCIMENTO E TRACCIAMENTO DELLE PERSONE

Riconoscimento Facciale

Per quanto riguarda la *posizione della testa*, la rotazione 3D della testa provoca ampie variazioni nell'aspetto. Ciò non porta solo a cambiamenti associati alla texture, ma anche a cambiamenti geometrici, poiché la testa umana non è invariante alla rotazione. Quando il volto viene ruotato, regioni importanti del volto (parte frontale) non sono più visibili. Più la testa viene ruotata, più una metà del volto diventa non visibile e si perdono informazioni preziose per il processo di riconoscimento.

Per quanto riguarda l'*occlusione*, una riduzione delle informazioni utilizzabili è spesso conseguenza dell'occlusione da parte di oggetti, come sciarpe, occhiali, berretti. La regione della bocca o quella più importante degli occhi vengono coperte in parte e non consentono l'estrazione di features. Tuttavia, l'occlusione può essere gestita da approcci locali per il riconoscimento del volto. La variazione dell'occlusione è potenzialmente infinita perché l'aspetto degli oggetti occludenti (esempi: occhiali, giornali, berretti) di solito non è noto.

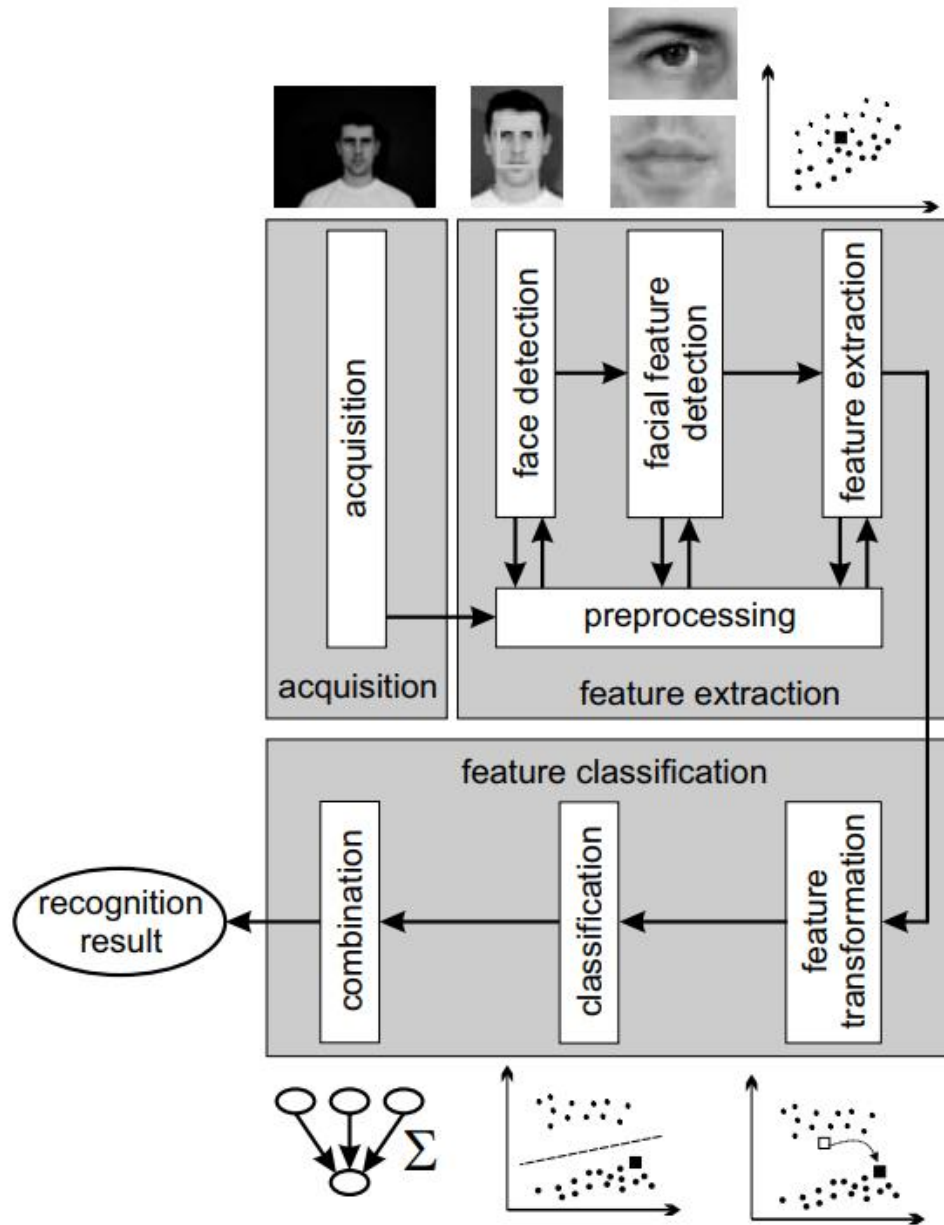
Per quanto riguarda i *cambiamenti dell'aspetto*, cambiamenti intenzionali (esempi: baffi, barba, trucco) possono complicare il rilevamento e l'estrazione delle features se tali cambiamenti associati alla texture non vengono considerati dall'algoritmo di rilevamento.

RICONOSCIMENTO E TRACCIAMENTO DELLE PERSONE

Riconoscimento Facciale

Fattori biologici: esempi di fattori biologici sono l'*invecchiamento* e il *twin-problem*. Per quanto riguarda l'*invecchiamento*, cambiamenti sostanziali della superficie cutanea e lievi cambiamenti geometrici (considerando intervalli di tempo più lunghi) causano problemi nel processo di riconoscimento a causa delle forti differenze nelle features estratte. Per quanto riguarda il *twin-problem*, la «copia» naturale di un volto da identificare è quella del volto gemello. Gli esseri umani hanno difficoltà a distinguere tra due gemelli non conosciuti, così come i sistemi automatici. Tuttavia, gli umani sono in grado di riconoscere i gemelli se essi sono conosciuti in modo adeguato.

A differenza di un sistema di riconoscimento facciale non invasivo, che deve gestire la maggior parte dei fattori sopra citati, un sistema invasivo può trascurare molti di questi fattori. Se si chiede a un cosiddetto «utente disponibile» («friendly user») di entrare in un'area con illuminazione controllata o di guardare direttamente nella camera, non è necessario gestire forti variazioni (esempi: illuminazione, posizione). Le condizioni operative possono spesso ridurre la necessità di algoritmi più sofisticati grazie alla scelta di ambienti con hardware adeguato. Tuttavia, ad esempio è difficile che ci siano «utenti disponibili» nelle attività di sorveglianza.



Struttura generale di un sistema di riconoscimento facciale.

RICONOSCIMENTO E TRACCIAMENTO DELLE PERSONE

Riconoscimento Facciale

Acquisizione

Durante l'acquisizione, la camera acquisisce immagini istantanee o flussi video relativi ad una scena contenente uno o più volti. In questa fase si deve considerare la posizione di installazione della camera. Se possibile, si dovrebbe utilizzare più di una camera in modo da ottenere differenti scene della stessa persona per combinare i risultati del riconoscimento o per costruire un modello di volto che viene utilizzato per gestire diversi fattori come la posizione o l'illuminazione.

Oltre al posizionamento della camera, si deve considerare anche la posizione per un'illuminazione adeguata. Se è possibile garantire una condizione di illuminazione costante sia nella fase di training che nella fase di test, è necessario ridurre gli sforzi per modellizzare i cambiamenti di illuminazione e l'accuratezza del sistema aumenta. Se possibile, i parametri della camera, come la velocità dell'otturatore e il bilanciamento del bianco, dovrebbero essere controllati in modo adattativo da un apposito algoritmo che valuta la situazione di illuminazione corrente e adatta i parametri di conseguenza.

RICONOSCIMENTO E TRACCIAMENTO DELLE PERSONE

Riconoscimento Facciale

Preprocessing

Ci possono essere diverse fasi di preelaborazione come ad esempio la normalizzazione geometrica dell'immagine e il miglioramento di luminosità e contrasto. Le fasi di preelaborazione da eseguire dipendono dallo step successivo.

Ad esempio, per la fase di *face detection* (si veda l'immagine riportata nelle slide precedenti) potrebbe essere utile determinare le regioni dell'immagine che potrebbero contenere un volto. Ciò può essere ottenuto ad esempio utilizzando gli istogrammi del colore (pelle).

Face Detection

Esistono differenti algoritmi in tale contesto. Una difficoltà comune per tutti gli algoritmi è rappresentata dalle condizioni del mondo reale. Il risultato della maggior parte degli algoritmi di rilevamento dei volti è rappresentato da rettangoli contenenti l'area del volto. Di solito una buona stima della dimensione del volto può essere ricavata da tali rettangoli. Tale dimensione è necessaria ad esempio per la fase di *facial feature detection*.

RICONOSCIMENTO E TRACCIAMENTO DELLE PERSONE

Riconoscimento Facciale

Facial Feature Detection

Tale fase viene eseguita solo nei metodi di riconoscimento che utilizzano informazioni locali. La normalizzazione delle dimensioni delle immagini viene spesso effettuata utilizzando la distanza tra gli occhi, che deve essere rilevata per tale scopo. Un approccio olistico si basa su modelli *AAM* (*Active Appearance Model*).

Feature Extraction

Oltre a una corretta preelaborazione, l'estrazione delle features è la fase più importante del riconoscimento facciale, poiché in questa fase si costruisce una descrizione di un volto che deve essere sufficiente per separare le immagini del volto di una persona da quelle di tutti gli altri soggetti.

Feature Transformation

La *feature extraction* e la *feature transformation* molto spesso sono combinate in un'unica fase. Il termine *feature transformation* descrive un insieme di metodi matematici che cercano di diminuire la *within-class distance* e di aumentare la *between-class difference* delle features estratte dall'immagine. Un esempio di metodo di *feature transformation* è la Principal Component Analysis (PCA). La PCA trasforma lo spazio delle features in una rappresentazione più adatta per l'elaborazione da parte di un classificatore.

RICONOSCIMENTO E TRACCIAMENTO DELLE PERSONE

Riconoscimento Facciale

Classification

In tale fase del processo di riconoscimento, un metodo di classificazione confronta un dato *feature vector* con i dati di training precedentemente memorizzati. Un esempio di algoritmo è il classificatore *k*-nearest-neighbor. Altri algoritmi si basano sull'utilizzo di reti neurali e di *SVMs* (*Support Vector Machines*).

Combination

L'input di un algoritmo di combinazione è rappresentato dai risultati delle classificazioni di differenti parti del volto o dai risultati associati a differenti algoritmi di riconoscimento e classificatori (esempio: metodi di riconoscimento locali e ibridi).

Le strategie di combinazione possono essere classificate in tre gruppi:

- *Voting*: il risultato finale è dato dall'ID che è stato giudicato come il migliore dalla maggioranza di tutti i singoli classificatori.
- *Ranking*: se i classificatori possono costruire una classifica per ogni immagine di test allora vengono sommate le posizioni in classifica di ogni classe. Viene scelta la classe con la somma minore.

RICONOSCIMENTO E TRACCIAMENTO DELLE PERSONE

Riconoscimento Facciale

- *Scoring*: se le classifiche sono caratterizzate da un punteggio che integra le informazioni su quanto un'immagine di test è simile all'immagine di training più vicina di ciascuna classe, è possibile applicare metodi di combinazione più sofisticati che combinano questi punteggi per ottenere un punteggio complessivo. La classe vincitrice è quella con il punteggio migliore.

RICONOSCIMENTO E TRACCIAMENTO DELLE PERSONE

Riconoscimento Facciale

Gli algoritmi di riconoscimento facciale possono essere classificati in 4 gruppi:

- Metodi *geometry-based*: tali metodi dipendono solo dalle relazioni geometriche tra le differenti regioni facciali. Essi sono stati sviluppati nei primi tempi della ricerca sul riconoscimento facciale.
- Metodi globali (*global*) o olistici (*holistic*): tali metodi utilizzano l'informazione associata all'intera immagine. In questo caso le informazioni geometriche non vengono utilizzate affatto o solo indirettamente attraverso la distribuzione dei valori di intensità (ad esempio, gli occhi in immagini diverse). Un esempio di approccio appartenente a tale categoria è l'*Eigenface Approach*.

RICONOSCIMENTO E TRACCIAMENTO DELLE PERSONE

Riconoscimento Facciale

Gli algoritmi di riconoscimento facciale possono essere classificati in 4 gruppi:

- Metodi basati su *local feature (analysis)*: tali metodi considerano solo l'informazione della texture delle regioni del volto senza integrare direttamente le informazioni geometriche. Le regioni locali contengono solo alcuni pixels ma sono caratteristiche di una specifica immagine del volto. Tali metodi includono un algoritmo per rilevare le regioni caratteristiche che viene applicato durante la fase di training (per ottenere una descrizione appropriata del volto) e durante la fase di test (per trovare le regioni dell'immagine di test che sono simili alle regioni precedentemente descritte nella fase di training).

RICONOSCIMENTO E TRACCIAMENTO DELLE PERSONE

Riconoscimento Facciale

Gli algoritmi di riconoscimento facciale possono essere classificati in 4 gruppi:

- Algoritmi ibridi (*hybrid*): essi combinano informazioni geometriche e informazioni associate all'aspetto per riconoscere un volto. Tali metodi di solito includono alcuni tipi di informazioni locali, cioè essi utilizzano informazioni associate all'aspetto relative a regioni facciali con una grande entropia riguardo al riconoscimento facciale (esempio: la regione degli occhi), ignorando regioni con scarso contenuto informativo. Tali regioni con scarso contenuto informativo sono quelle regioni caratterizzate da textures omogenee (esempi: fronte, guance). Inoltre, le informazioni geometriche riguardanti la relazione tra i punti di estrazione delle features vengono incluse nel processo di classificazione. L'utilizzo di informazioni locali riduce l'influenza dei cambiamenti dell'illuminazione e delle espressioni facciali e permette di gestire l'occlusione se essa viene rilevata.

Riferimenti Bibliografici

- [1] Kraiss, K. -F. (2006). Advanced Man-Machine Interaction: Fundamentals and Implementation. Springer-Verlag Berlin Heidelberg. ISBN-10: 3-540-30618-8