

# Riconoscimento del volto

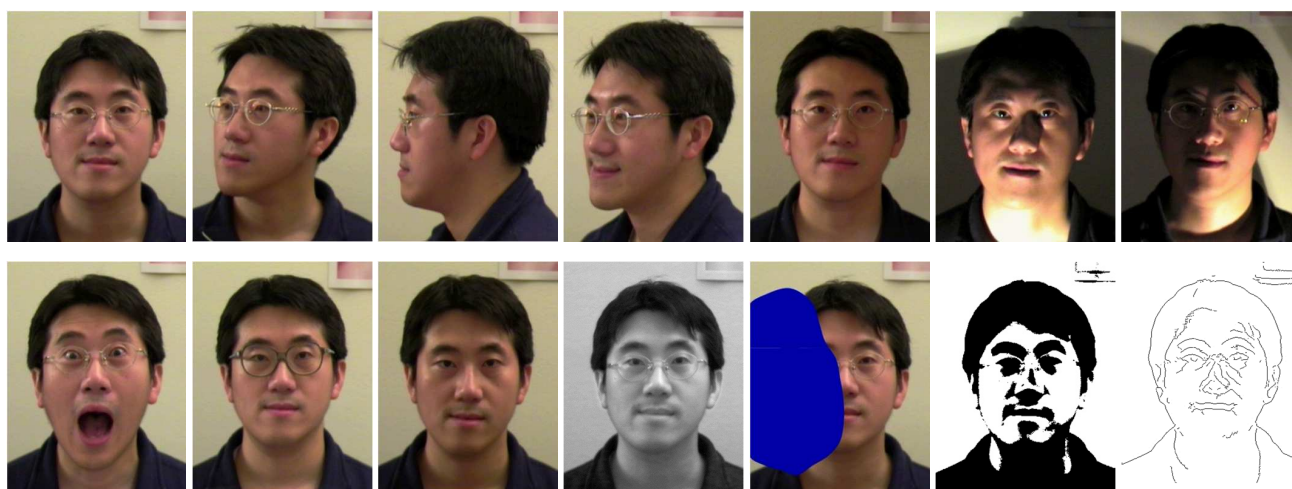
Tecniche 3D e 3D+2D

Annalisa Franco  
annalisa.franco@unibo.it

Dario Maio  
dario.maio@unibo.it

2

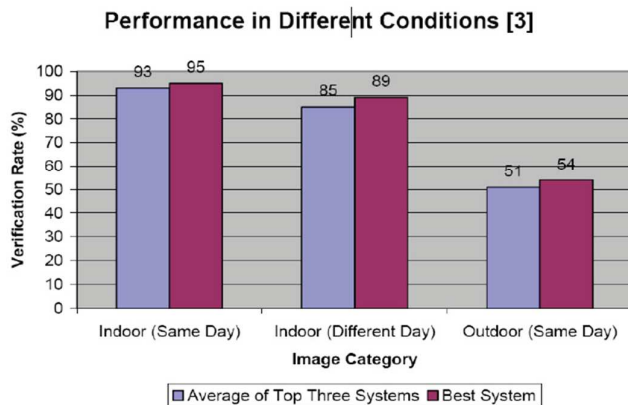
## Variabilità intra-classe



Alcuni contenuti sono in parte ripresi dalla presentazione  
<http://www.dmi.unisa.it/people/nappi/www/SB/lezione7.pdf>

## FRVT 2002

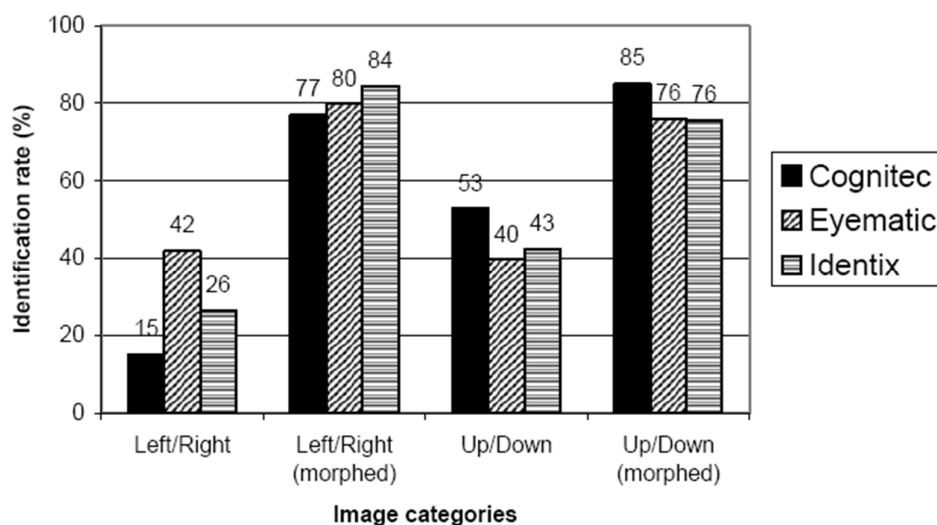
- FRVT2002 ha dimostrato che le performance degli algoritmi 2D sono drasticamente ridotte se si usano immagini dello stesso soggetto prese con illuminazioni molto diverse tra loro (immagini indoor & outdoor prese nello stesso giorno).



Alcuni esempi di immagini FRVT  
Indoor & Outdoor



## FRVT 2002



“FRVT2000 and FRVT2002 show that one of the more difficult tasks for modern facial recognition systems is recognising faces in non-frontal images. Most facial recognition systems perform well when all of the images are frontal. But, as a subject becomes more and more off angle (both horizontally and vertically), performance decreases.” **Source: Face Recognition Vendor Test 2002, Overview and Summary, March 2003.**

# FRVT 2006

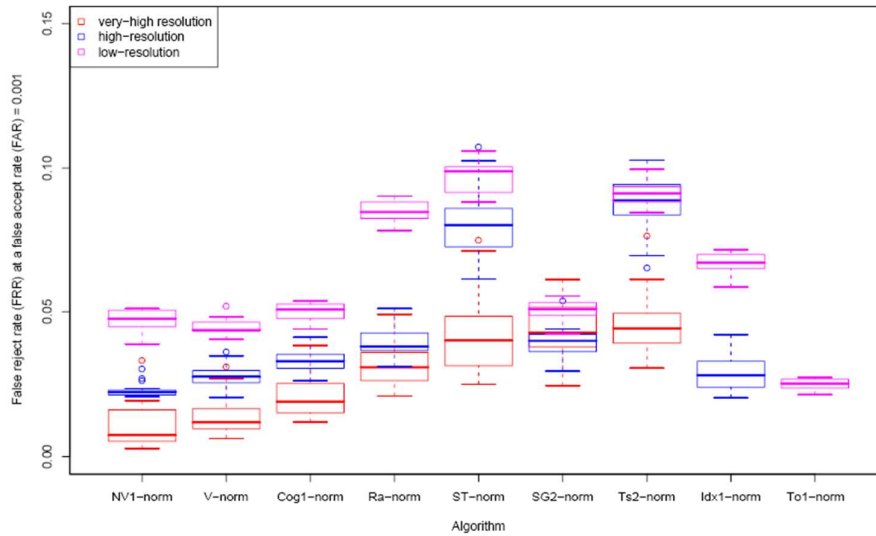


Figure 5: SUMMARY OF STILL FACE RECOGNITION PERFORMANCE ON THE VERY HIGH-RESOLUTION, HIGH-RESOLUTION, AND LOW-RESOLUTION DATA SETS. EACH COLUMN IN THE GRAPH REPORTS PERFORMANCE FOR ONE ALGORITHM WITH RESULTS PROVIDED FOR UP TO THREE DATA SETS. FOR EACH ALGORITHM, THE PERFORMANCE RESULTS ON A DATA SET ARE REPORTED BY A DIFFERENT COLOR BOXPLOT. FOR A SAGEM (SG2-NORM) ALGORITHM, THE BODY OF THE BOXPLOTS OVERLAP FOR ALL THREE DATASETS. FOR A TSINGHUA (TS2-NORM) ALGORITHM, THE BODY OF THE BOXPLOTS OVERLAP THE HIGH-RESOLUTION AND LOW-RESOLUTION DATASETS. FOR IDENTIX (IDX1-NORM) AND TOSHIBA (TO1-NORM), PERFORMANCE WAS OUTSIDE THE RANGE OF THIS GRAPH FOR AT LEAST ONE DATASET.

# FRVT 2006

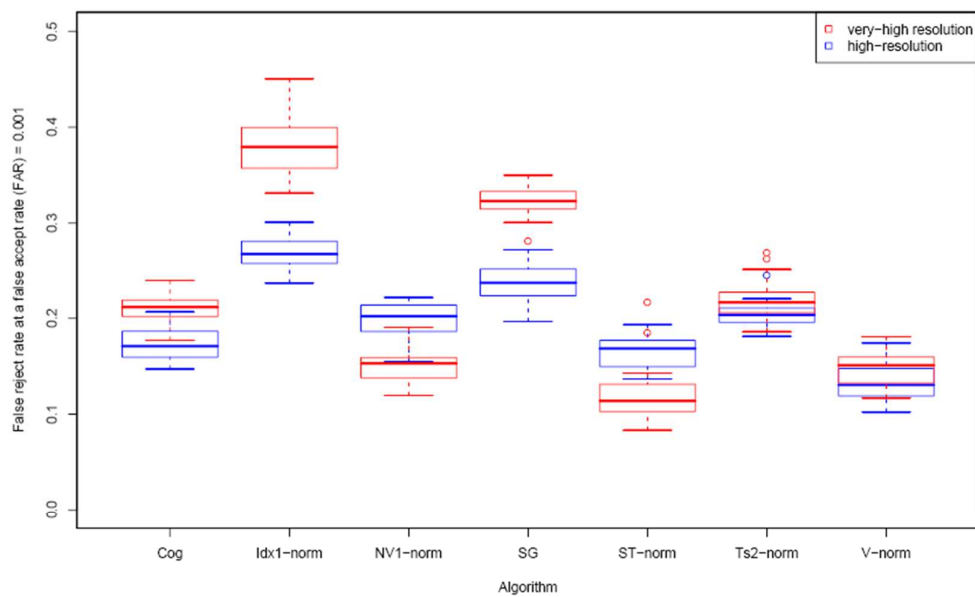


Figure 7: SUMMARY OF STILL FACE RECOGNITION PERFORMANCE ACROSS ILLUMINATION CHANGES ON THE VERY-HIGH RESOLUTION AND HIGH RESOLUTION DATASETS. FOR COGNITEC AND SAGEM, RESULTS FOR THE COG1-NORM AND SG2-NORM ALGORITHMS ARE REPORTED ON THE VERY-HIGH RESOLUTION DATASET, AND RESULTS FOR THE COG1-1TO1 AND SG1-1TO1 ALGORITHMS ARE REPORTED ON THE HIGH-RESOLUTION DATASET.

# FRVT 2006

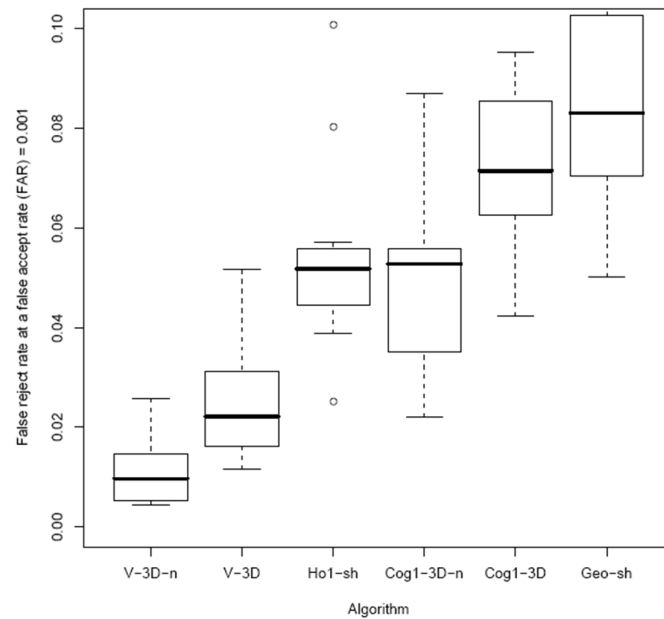


Figure 6: SUMMARY OF PERFORMANCE FOR 3D FACE RECOGNITION ALGORITHMS.

## Invarianza rispetto alle variazioni del volto

Fattori	Shape (3D)	Appearance (2D)
Posa	Sì	No
Illuminazione	Sì	No
Espressione	No	No
Invecchiamento	No	No
Makeup	Sì	No
Accessori	No	No

Né i modelli 2D né quelli 3D offrono un'invarianza rispetto a tutte le possibili variazioni del volto

# Rappresentazione di un volto

## 2D

- **Intensity Image:** un'immagine bidimensionale in cui il valore di ogni pixel rappresenta l'intensità della luce riflessa sul quel punto. Il colore di un pixel è dipende da come la luce è riflessa sulla superficie.

## 2.5D

- **Range Image:** un'immagine bidimensionale in cui il valore associato a ogni pixel rappresenta la distanza tra la sorgente e il punto

## 3D

- **Shaded Model:** una struttura di punti e poligoni connessi tra loro in uno spazio a tre dimensioni

2D



2.5D



3D



# Il riconoscimento 3D

## • Vantaggi:

- Le tecniche 3D sono più **efficaci** di quelle 2D poiché il modello 3D è molto più ricco di informazioni.
- Un sistema di riconoscimento 3D è più **robusto** rispetto a cambiamenti di posa e illuminazione.
- Il modello 3D permette anche la generazione di **pose "virtuali"** (es. con diverse espressioni del volto).

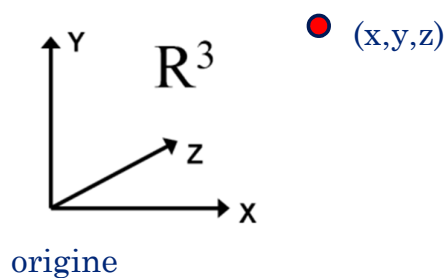
## • Svantaggi:

- I dispositivi d'acquisizione 3D sono **costosi** (il costo aumenta con la precisione di acquisizione).
- Alcuni sistemi di acquisizione 3D sono **invasivi**:
  - Tempi di acquisizione lunghi
  - Alcuni sistemi possono essere pericolosi per la retina (laser scanner)
- Sostituire in un'organizzazione i dispositivi 2D (macchine fotografiche, videocamere, ecc.) con nuove apparecchiature 3D è un processo che richiede tempo e costi elevati

## Rendering 3D (1)

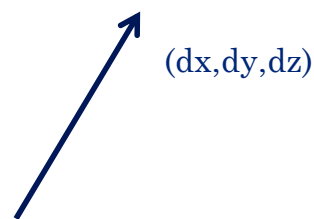
### • Vertici

- Un vertice specifica la locazione di un punto
- Non ha dimensione
- È definito da una terna di coordinate  $(x,y,z)$



### • Vettori liberi

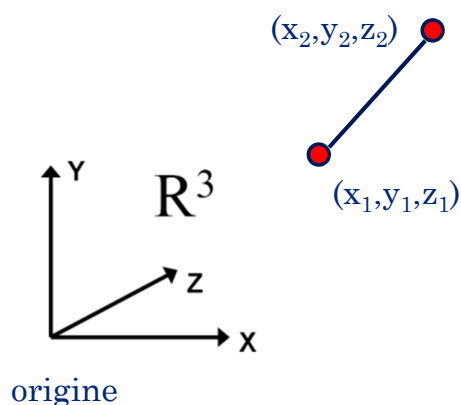
- Un vettore è caratterizzato da una direzione e una magnitudine (lunghezza)
- Non ha locazione
- Una terna di valori  $(dx, dy, dz)$  indica la direzione e il verso del vettore



## Rendering 3D (2)

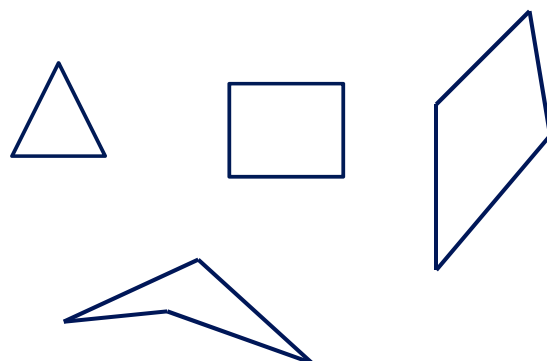
### • Segmenti

Un segmento è un percorso lineare che unisce due vertici



### • Poligoni

Un poligono è costituito da una sequenza "complanare" di punti uniti da segmenti

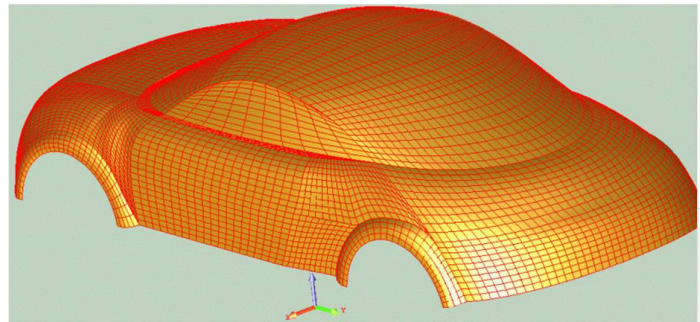
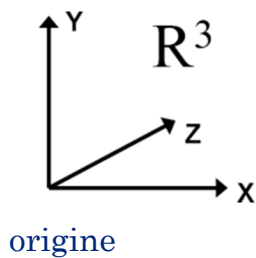
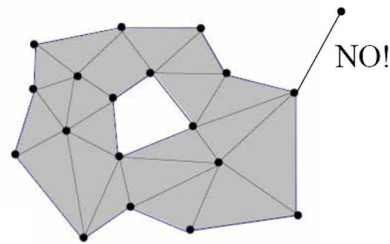


## Rendering 3D (3)

### • Mesh poligonale

Una mesh poligonale è un insieme di poligoni che godono di alcune proprietà tra cui:

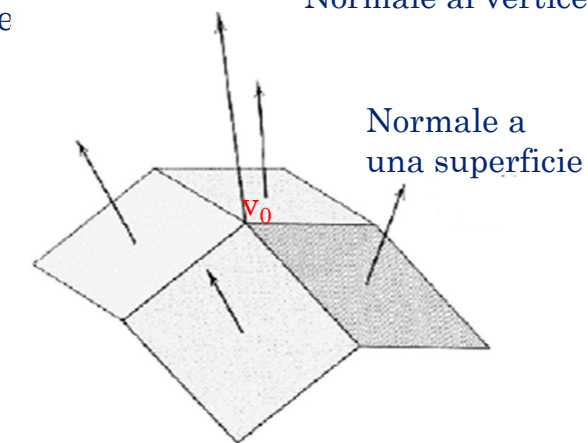
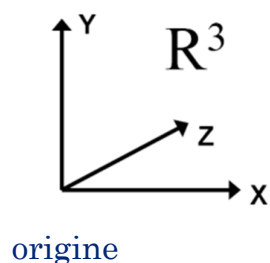
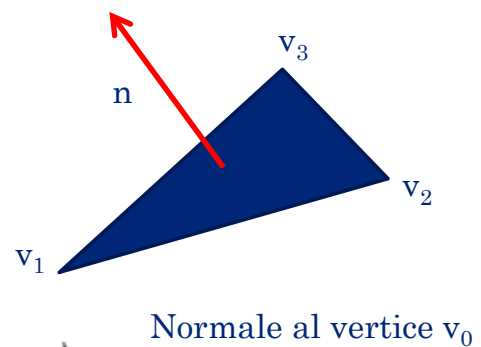
- ogni lato appartiene almeno ad un poligono
- ogni vertice ha almeno due lati



## Rendering 3D (4)

### • Normali

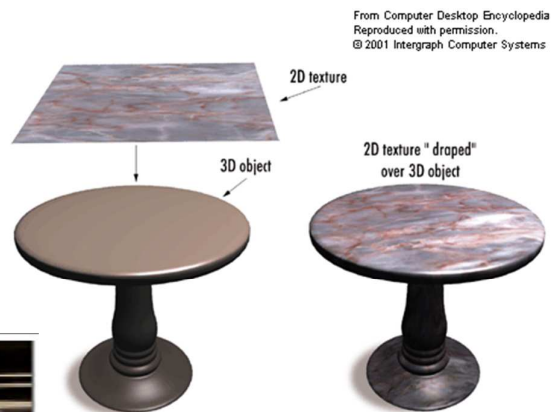
- **Una normale di un poligono** è un vettore perpendicolare a un piano o a un poligono
  - utile per comprendere la posa e l'orientamento di un poligono
- **La normale di un vertice** è la risultante della somma delle normali dei poligoni a cui il vertice appartiene
  - utile per comprendere la posa e l'orientamento di un vertice



## Rendering 3D (5)

- **Texture mapping**

Una texture aggiunge dettagli alla superficie di un oggetto 3D



Problema: come assegnare i pixel della texture?

## Rendering 3D (6)

- **È possibile assegnare un colore ad un poligono**; il poligono ha un singolo colore, uniforme, lungo la sua superficie.
- **È possibile assegnare un colore a un vertice**; il colore del poligono è la combinazione dei colori dei suoi vertici, che può essere di diversi tipi, es:
  - flat**: il colore del poligono è uno ed è la somma dei colori dei vertici
  - smooth**: il colore del poligono non è uniforme ma cambia all'avvicinarsi dei suoi vertici.

Un altro modo per assegnare i colori a una mesh consiste nel mappare un'immagine 2D lungo la superficie della mesh.

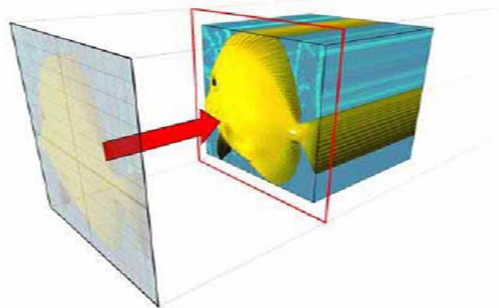
**Texture**: indica le mappe (immagini **bitmap** o **shader** procedurali) che sono applicate sulle superfici dei poligoni degli oggetti 3D.

**Texture Mapping**: consiste nella procedura utilizzata per calcolare la posizione e l'orientamento di una tessitura sulla superficie dell'oggetto 3D. Queste immagini in fase di rendering rappresentano il colore, la luminosità, la riflessione, ecc. delle superfici 3D.

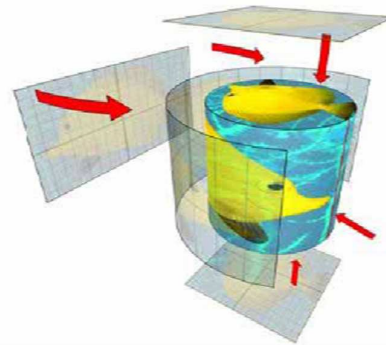


## Rendering 3D (7)

- La mappatura di tipo planare consiste in una specie di proiezione piana dell'immagine bitmap sulla superficie del modello 3D.
- È impiegata per oggetti che sono relativamente piatti o totalmente visibili attraverso un angolo qualsiasi della camera



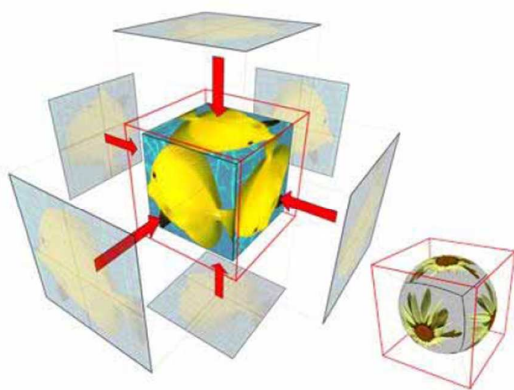
**Planar mapping**



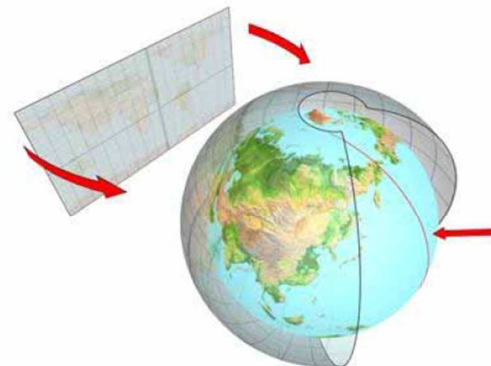
**Cylindrical mapping**

## Rendering 3D (8)

- Valgono considerazioni analoghe al cylindrical mapping ma con riferimento ad altre forme geometriche



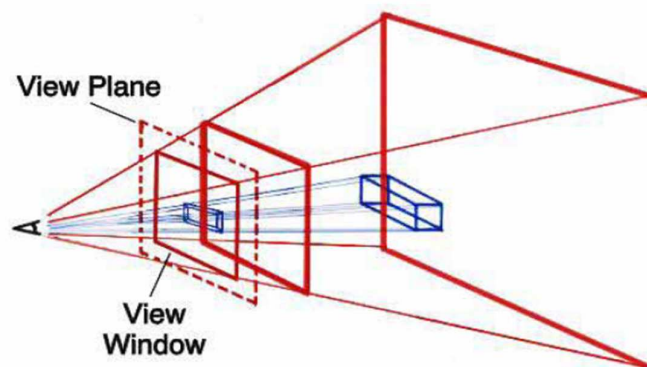
**Box mapping**






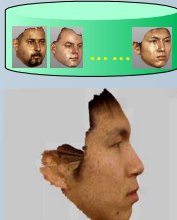

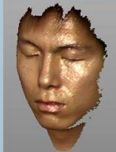
**Spherical mapping**

## Rendering 3D (9)

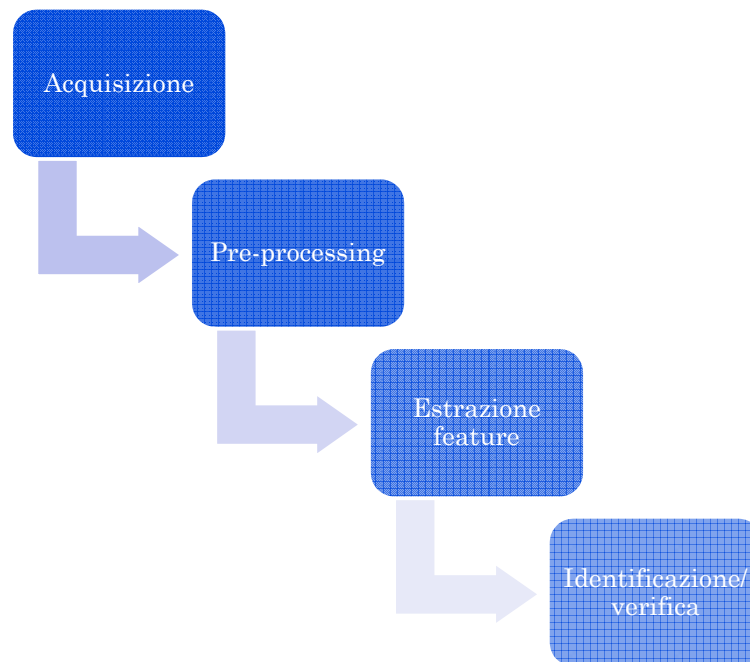
- **Rendering (presentazione fotorealistica)** : processo di generazione di un'immagine bidimensionale a partire da una descrizione (**mesh**) degli oggetti tridimensionali.
- Il rendering di un modello 3D genera un'immagine bidimensionale (**frame**) che rappresenta il modello da un preciso punto di vista.



## Modelli del volto: possibili scenari

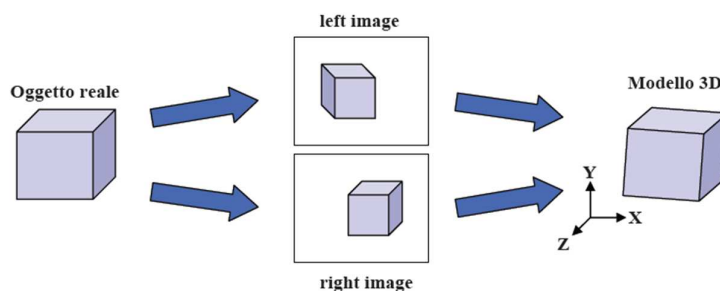
		Query/Test	
		2D	3D
Template	2D		 <p>←</p> <p>Uso di modelli 3D per reperire immagini 2D!</p>
	3D		 <p>←</p>  <p>←</p> 

# Fasi del riconoscimento 3D



## Acquisizione 3D: camere stereoscopiche

- Acquisiscono un volto con **due (o più) camere** stereoscopiche.
- La superficie 3D risultante è generata a partire dalla coppia di immagini.
  - Sono sensibili alle variazioni d'illuminazione.
  - Costo: basso.
  - Qualità d'acquisizione: media.
  - Tempo d'acquisizione: real-time.



### Algoritmo di ricostruzione

1. Trovare le feature in una delle due immagini
2. Cercare le stesse feature nell'altra immagine
3. Calcolare la corrispondenza tra la coppia di feature per trovare la coordinata z.

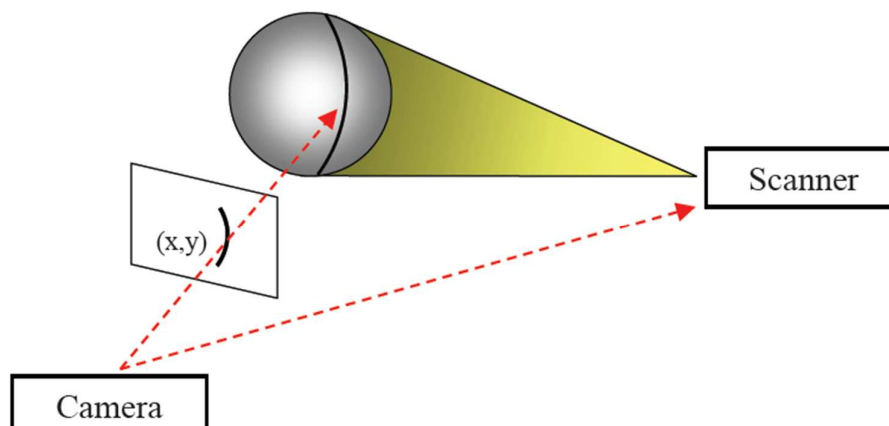
## Acquisizione 3D: Structured-Light Scanners (1)

- Gli scanner a luce strutturata tipicamente **proiettano un pattern luminoso sul volto**.
- La sorgente luminosa è solitamente una luce alogena ordinaria, quindi non causa problemi alla retina.
  - Leggermente sensibili alle variazioni d'illuminazione
  - Costo: medio-elevato
  - Qualità d'acquisizione: medio-elevata
  - Tempo d'acquisizione: 3-8 secondi



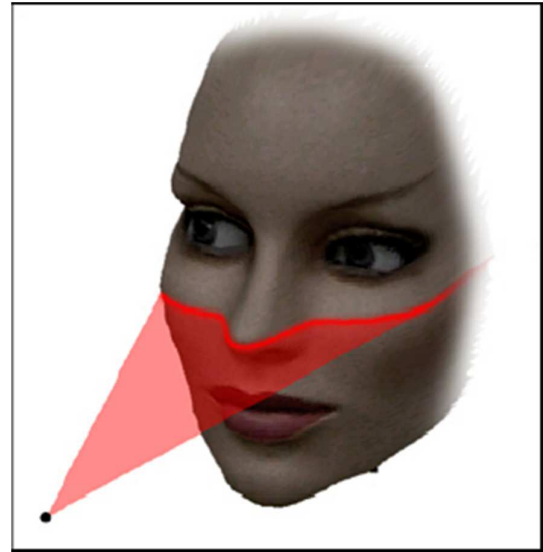
## Acquisizione 3D: Structured-Light Scanners (2)

- Il pattern luminoso è distorto a causa della superficie del volto.
- Possono essere utilizzati differenti pattern luminosi (**griglie, linee, cerchi, sinusoidi**, ecc.).
- Gli scanner a luce strutturata catturano una superficie completa da un particolare punto di vista. I dati provenienti da punti di vista multipli possono essere combinati per creare un modello 3D completo della testa.



## Acquisizione 3D: Laser scanner

- Gli scanner laser proiettano un **singolo fascio laser** sul volto.
- L'algoritmo di generazione del modello 3D è simile a quello utilizzato dagli scanner a luce strutturata.
  - La sorgente laser è invasiva e pericolosa per la retina.
  - Non sono sensibili alle variazioni d'illuminazione.
  - Costo: medio-elevato
  - La qualità d'acquisizione è elevata
  - Tempo d'acquisizione: 6-30 secondi.



## Dispositivi di acquisizione 3D



Enrollment Station – a structured light system



Door Access Control – uses “invisible” near-infrared light



**inSpeck**



Foto 2D



Fringe Image

## Dispositivi di acquisizione 3D

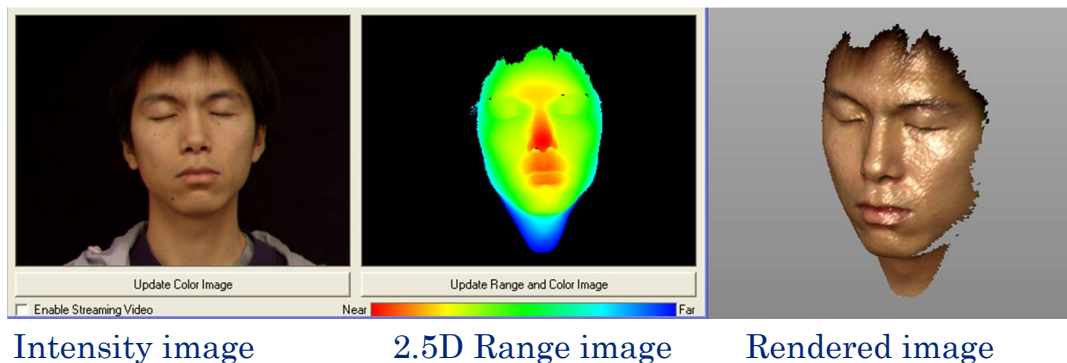


Minolta Vivid 910 scanner

Acquisition time ~ 1sec.

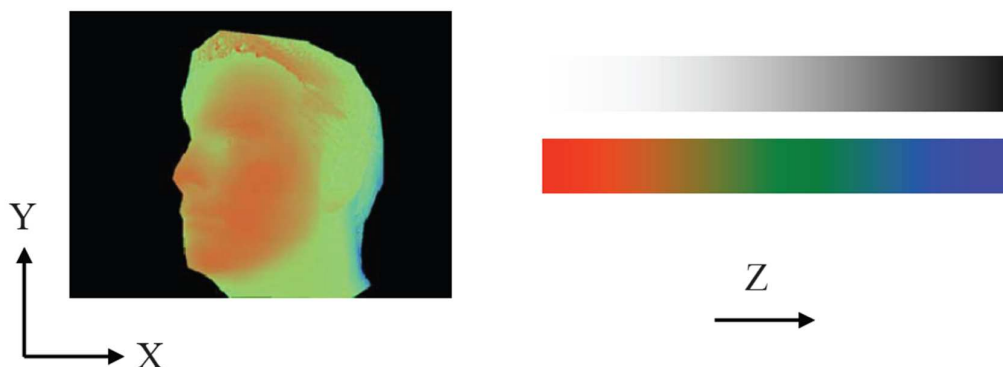
320x240 image size

Depth resolution < 0.1 mm



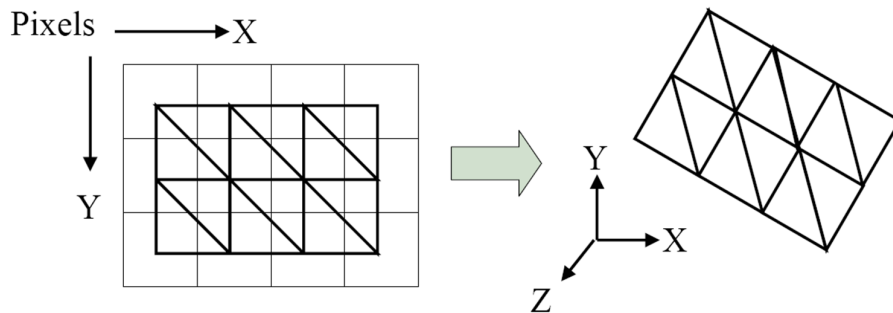
## Immagine 2.5D vs. modelli 3D (1)

- Una **Range Image** (o **immagine 2.5D**) è una immagine bidimensionale che contiene al proprio interno informazioni tridimensionali.
- Ogni pixel rappresenta un punto dello spazio il cui *colore indica la profondità* del punto dalla sorgente (scanner). Lo spazio colore può essere a toni di grigio o RGB (Red, Green, Blue).

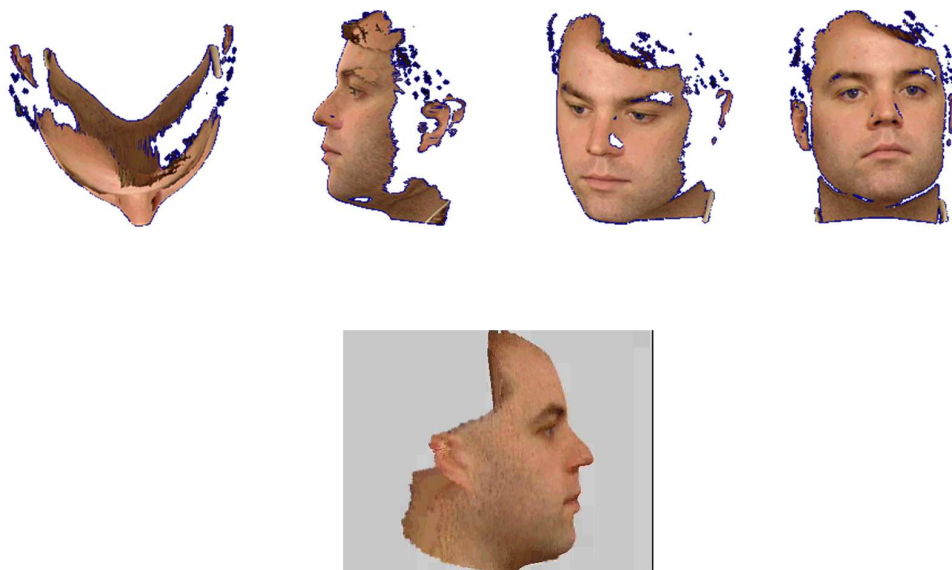


## Immagini 2.5D vs. modelli 3D (2)

- La generazione della mesh 3D avviene in due fasi:
  1. generazione della nuvola di vertici (x,y,z) dove le coordinate x e y sono equidistanti mentre il valore della coordinata z varia in dipendenza del colore del pixel (z rappresenta infatti la profondità)
  2. Triangolarizzazione della nuvola dei punti (mesh di triangoli adiacenti)



## Immagini 2.5D vs. modelli 3D (3)

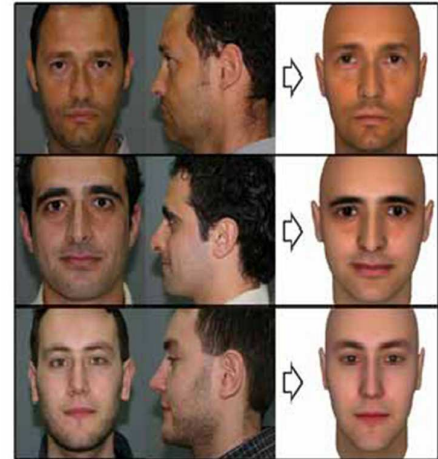


## Morphable models (1)

- ✓ Un sistema d'acquisizione atipico è quello che sfrutta una tecnica tradizionale della computer graphics, il **morphing**. A partire da un modello 3D iniziale si ottiene un altro modello 3D finale. I **morphable model** sono dunque rappresentazioni dinamiche della superficie del volto, che contengono **informazioni sulla struttura dinamica** (ad esempio muscolatura facciale) in aggiunta alle informazioni sulla geometria.
- ✓ Sono sensibili alle variazioni d'illuminazione
- ✓ costo : basso, qualità d'acquisizione: media, tempo d'acquisizione: > 1 minuto.

### Generazione del modello 3D finale

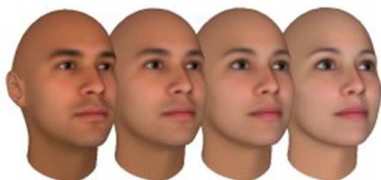
1. Si acquisiscono due o tre immagini fotografiche del volto (frontale, laterale, diagonale)
2. Si cambiano la forma e i colori di un modello generico di volto (morphable model) in accordo al contenuto delle immagini.
3. Si genera la texture combinando le due o tre immagini.



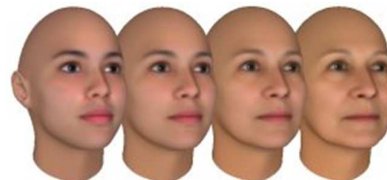
## Morphable models (2)

I **morphable model** consentono di generare differenti espressioni facciali sintetiche che possono essere sfruttate al fine di prevedere le espressioni facciali reali che un individuo può avere durante il processo di acquisizione.

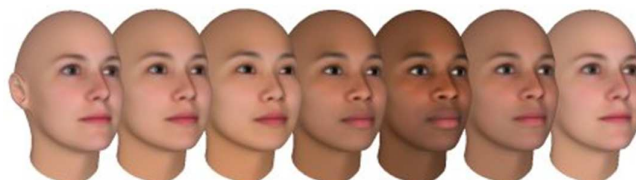
**Facegen Modeller** (<http://www.facegen.com/modeller.htm>) è uno strumento commerciale in grado di modellare diverse espressioni facciali nonché molteplici aspetti diversi tra cui:



gender morphing



age morphing



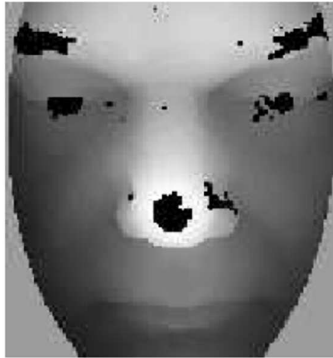
race morphing



## Pre-processing (1)

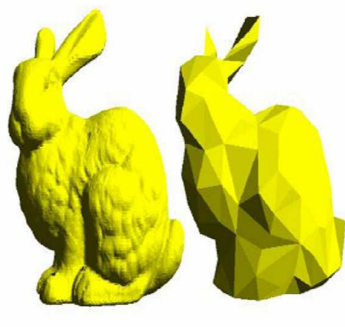
I sistemi di scansione tridimensionale sono affetti da errori di acquisizione per cui si rende necessario operare alcuni accorgimenti:

- **Noise removal:** spikes (filters), clutter(manually), noise(medianfilter)
- **Holes filling** (Gaussian smoothing, linear interpolation, symmetrical interpolation)



## Pre-processing (2) – sub-sampling

- Mesh poligonali molto accurate possono essere computazionalmente onerose da gestire, dunque è spesso necessario operare una semplificazione
- Usualmente si applica un algoritmo incrementale che rimuove un vertice alla volta e ripara il buco lasciato.
- **Obiettivo:** rimuovere il maggior numero possibile di vertici facendo in modo che la risultante maglia semplificata sia una buona approssimazione della mesh originale.

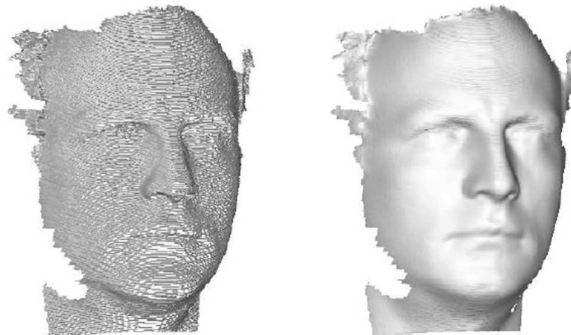


## Pre-processing (3) – smoothing

□ Operazione che consente di levigare gli angoli tra i poligoni al fine di creare una superficie più liscia.

**Tipico algoritmo:**

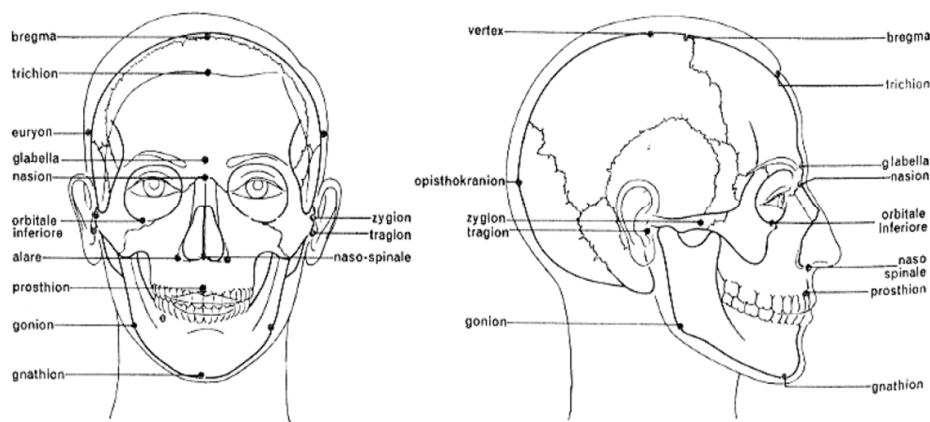
- Calcolare la curvatura locale  $\alpha$  tra un insieme di vertici vicini.
- Se  $\alpha$  contiene un punto  $p$  di massimo o di minimo più grande di un valore di soglia  $t$  “smussare” la superficie sostituendo  $p$  con il valore medio dei punti adiacenti.



**prima e dopo lo smoothing**

## Pre-processing (4) – allineamento

□ Principali **punti di repere** nell'anatomia di un volto

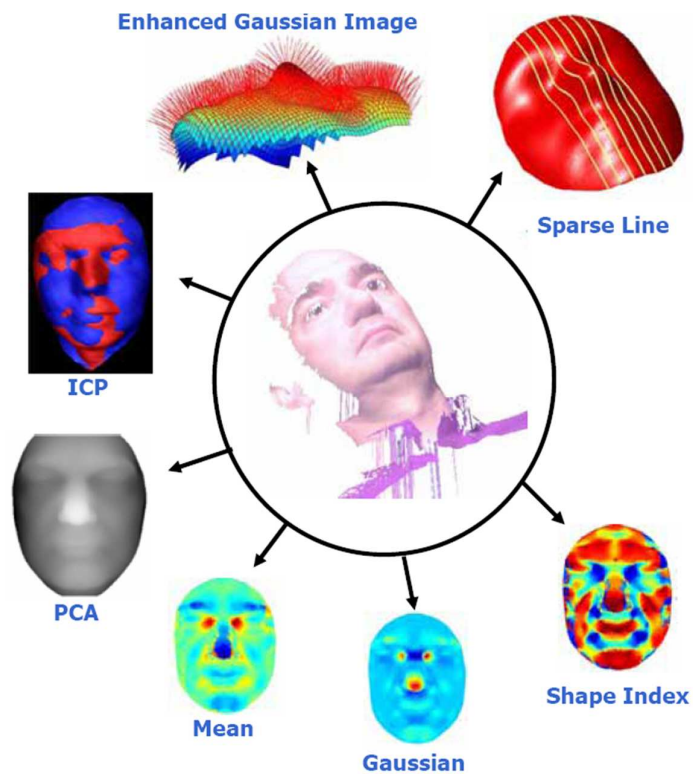


*I punti possono essere localizzati sulle immagini 2D oppure direttamente sul modello 3D acquisito.*

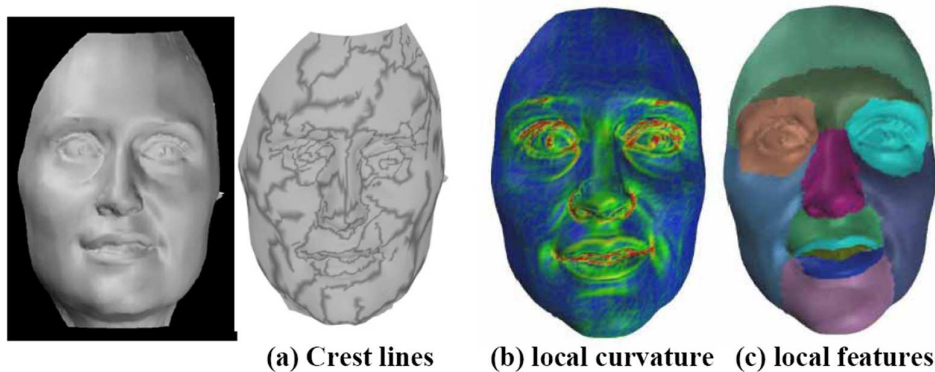
1. localizzare un numero finito di punti caratteristici di un volto (es. angolo degli occhi, punta del naso, centro della bocca);
2. allineare (rotazione, traslazione, scala) i volti riducendo la distanza tra punti corrispondenti al minimo

## Estrazione feature (1)

- ❑ Alcuni metodi d'estrazione e analisi delle feature 3D sono indipendenti dall'orientamento; non è richiesto in tal caso alcun allineamento.
- ❑ Gli algoritmi di riconoscimento facciale 3D operano spesso sull'analisi della curvatura locale e globale del modello del volto.



## Estrazione feature (2) – analisi curvatura



**Crest Lines:** si individuano le zone a maggior curvatura

**Local Curvature:** si rappresenta la curvatura locale con un colore

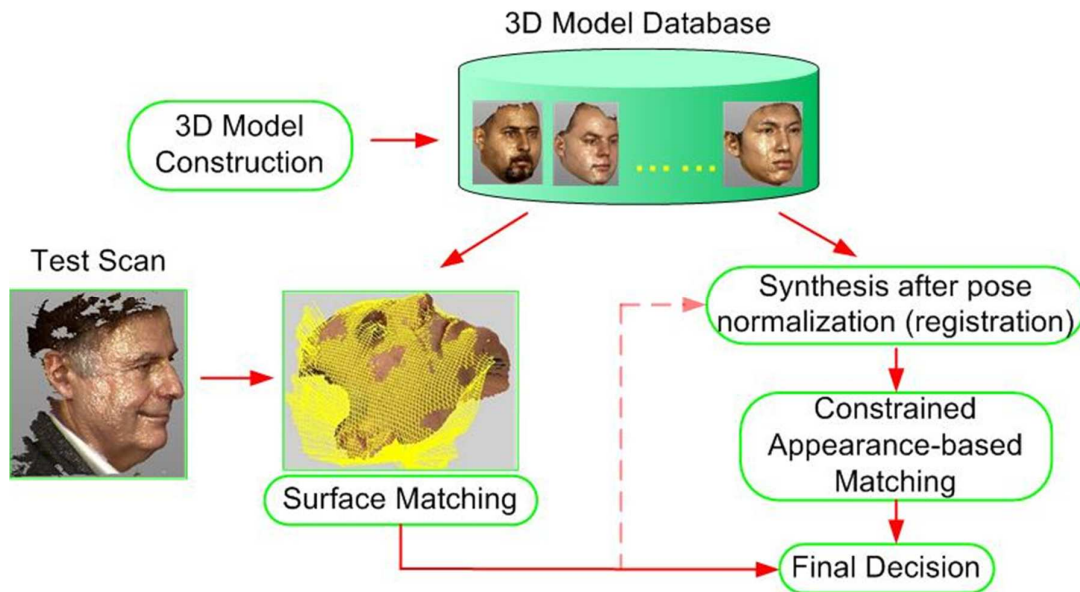
**Local Features:** il volto è segmentato in alcune zone d'interesse

# Classificazione e verifica

- ❑ I metodi di classificazione hanno l'obiettivo di archiviare, organizzare e indicizzare i modelli 3D in un database.
- ❑ I metodi sono analoghi a quelli applicati per il 2D ma adattati al caso 3D (PCA, non-linear PCA, Fisher mapping, ...).
- ❑ La verifica d'identità si ottiene misurando la distanza (euclidea, di Manhattan, di Chebychev, ..) tra coppie di feature facciali e fissando un valore di soglia.

## Una tecnica di riconoscimento 3D

# Il sistema di riconoscimento 3D



1. X. Lu, D. Colbry, A.K. Jain, "Three-Dimensional Model Based Face Recognition, *In Proc. International Conference on Pattern Recognition*, 2004.
2. X. Lu, A.K. Jain, D. Colbry, "Matching 2.5D Face Scans to 3D Models", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 1, pp. 31-43, 2006.

# Costruzione del modello 3D

2.5D Scans

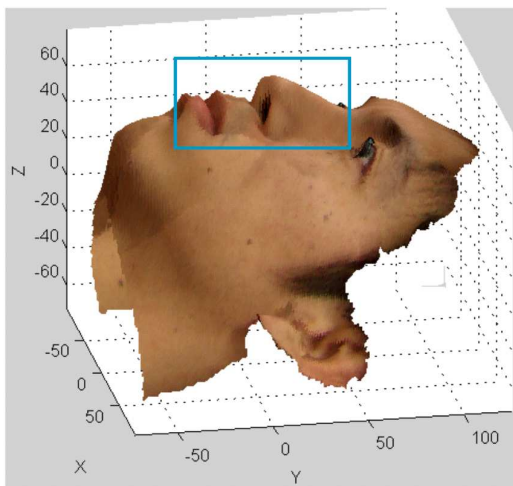


3D Model



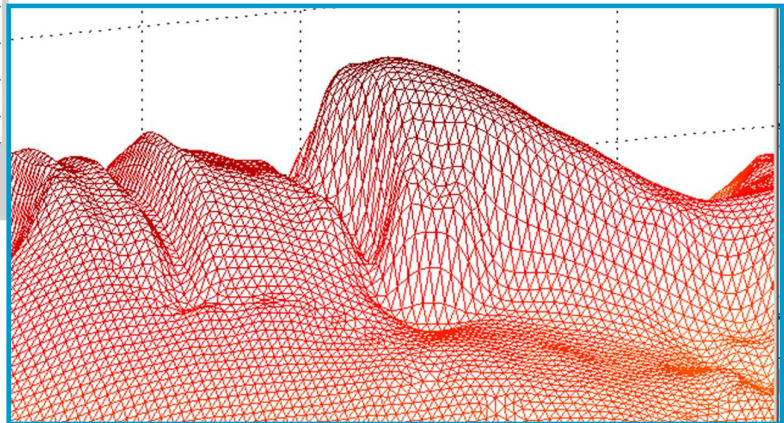
- Creazione del modello 3D a partire da 5 immagini 2.5D utilizzando software commerciali (Geomagic Studio e Rapid Form);
- Rimozione del rumore e riempimento dei buchi
- Semplificazione del modello (riduzione a 50.000 poligoni).

# Rappresentazione del modello 3D

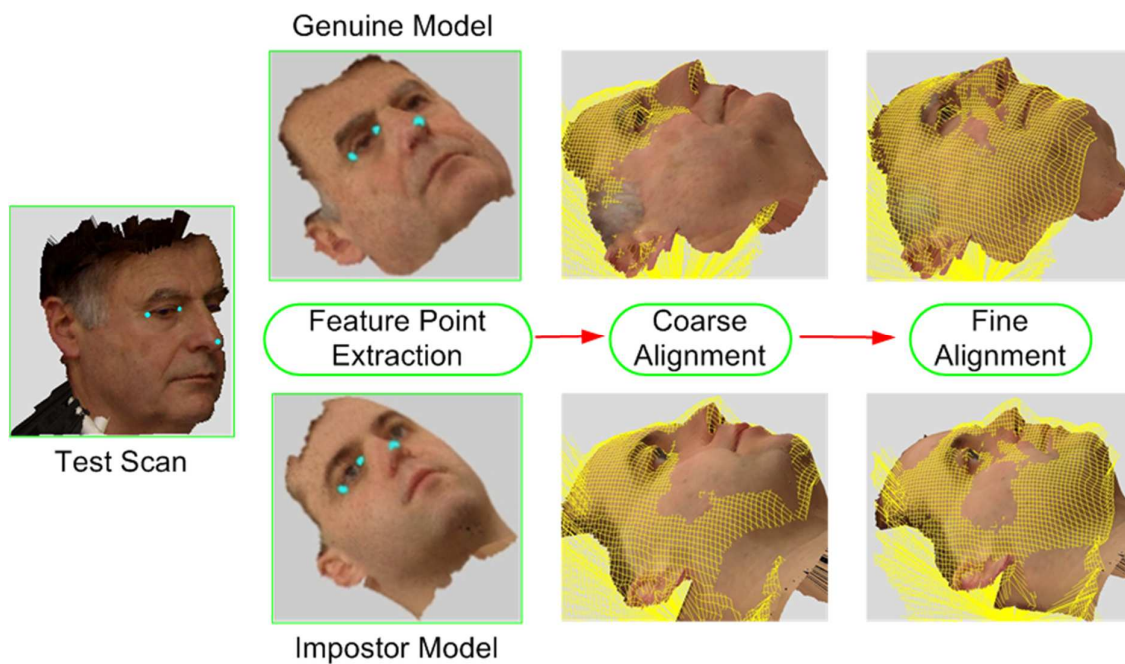


Modello 3D con texture

Mesh associata



# Matching di superfici



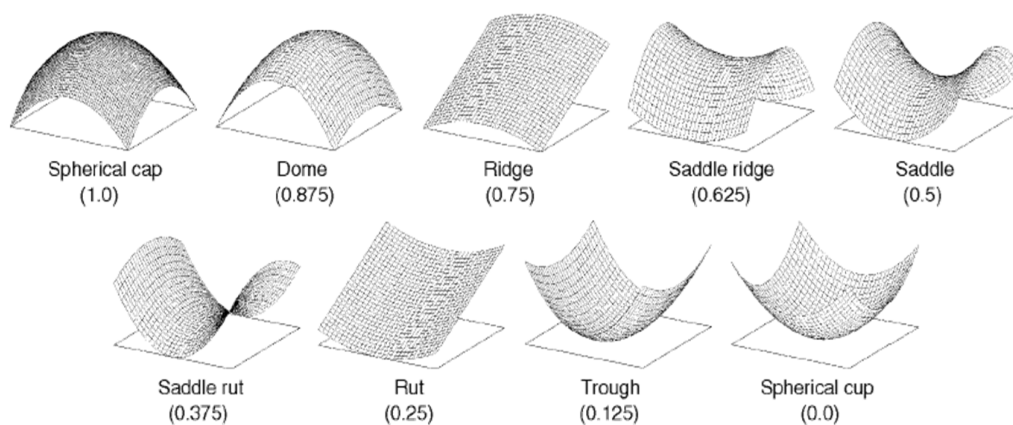
## Estrazione dei feature point

- A partire da un'immagine 2.5D sono individuati 3 feature point che permettono:
  - di calcolare la trasformazione rigida per l'allineamento dell'immagine 2.5D e il modello 3D;
  - di allineare una griglia di punti di controllo usata poi per l'allineamento finale.
- La localizzazione dei feature point avviene sulla base di un indice di forma (*shape index*), valutato in ogni punto dell'immagine 2.5D.
- Lo shape index nel punto  $p$  è definito usando il valore massimo ( $k_1$ ) e quello minimo ( $k_2$ ) della curvatura locale.

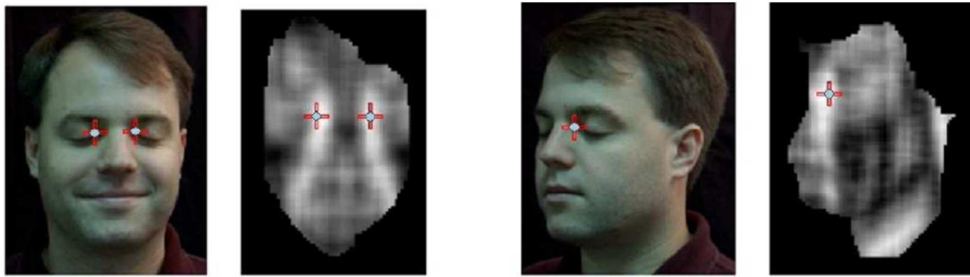
$$S(p) = \frac{1}{2} - \frac{1}{\pi} \tan^{-1} \frac{\kappa_1(p) + \kappa_2(p)}{\kappa_1(p) - \kappa_2(p)}$$

- Lo shape index assume valori nell'intervallo  $[0,1]$ .
- Il valore calcolato è indipendente dal sistema di coordinate e può quindi essere usato per la ricerca di punti simili in immagini 2.5D con pose diverse.

## Shape index



## Localizzazione dei feature point



(a)

(b)

Gli angoli interni degli occhi hanno un valore di shape index molto vicino a 0



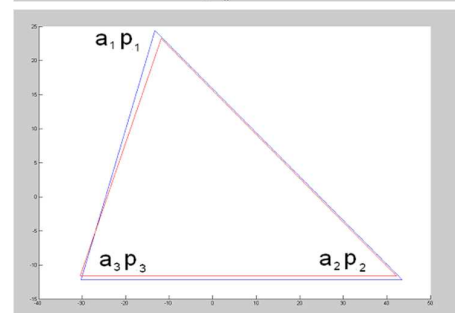
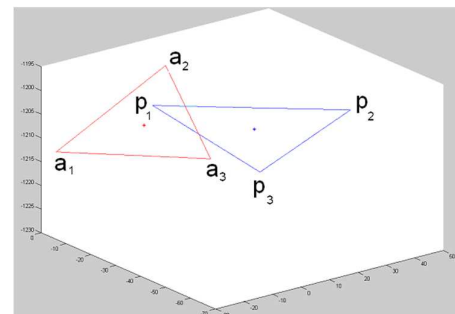
I feature point sono usati per un allineamento approssimativo

## Allineamento approssimativo

- La trasformazione rigida  $T$  è approssimata sulla base della corrispondenza di tre feature point:

dove 
$$T = T_{c_p} \cdot R_{p^t} \cdot \Theta \cdot R_A \cdot T_{C-a}$$

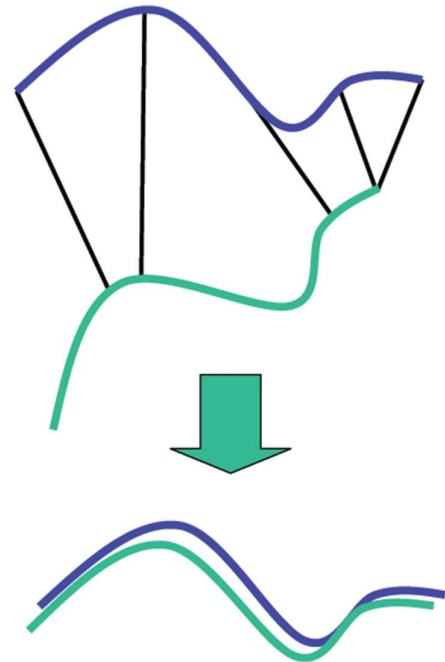
- $T$  è la trasformazione totale dal set di punti  $a$  al set  $p$ ;
- $T_{C-a}$ : traslazione che porta il centroide del triangolo  $a$  nell'origine degli assi;
- $R_a$ : rotazione nel piano  $xy$ ;
- $\Theta$ : rotazione ottimale ai fini dell'allineamento dei vertici nel piano  $xy$ ;
- $R_{p^t}$  rotazione al di fuori del piano  $xy$  nel sistema di coordinate di  $p$ ;
- $T_{c_p}$  traslazione dei vertici in modo che i centroidi dei due set coincidano.





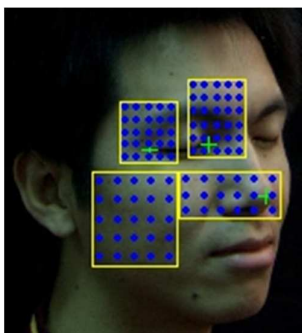
## Allineamento preciso

- L'allineamento è effettuato usando l'algoritmo Iterative Closest Point (ICP).
- L'algoritmo *prevede l'uso combinato di caratteristiche geometriche di tipo differente* (punti, superfici, linee, curve) che, estratte dai volumi da correlare, sono poi descritte come un insieme di punti.
- Per ogni punto di riferimento è identificato il punto del volume da registrare geometricamente più vicino ed è poi calcolata la trasformazione che minimizza ai minimi quadrati la distanza tra le coppie di punti.
- Si applica la trasformazione identificata e si ripete la procedura di allineamento fino a convergenza.

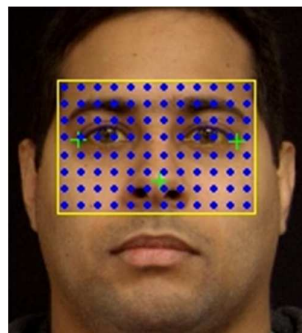


1. P. Besl, N. McKay, "A Method for Registration of 3D Shapes", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239-256, 1992.
2. Y. Chen, G. Medoni, "Object Modeling by Registration of Multiple Range Images", *Image and Vision Computing*, vol. 10, no. 3, pp. 145-155, 1992.

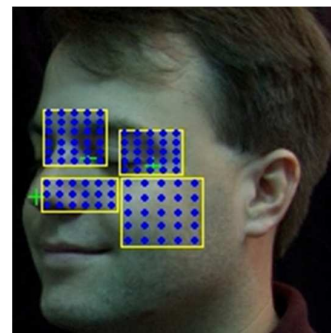
## Selezione dei punti di controllo



Right profile



Frontal scan



Left profile

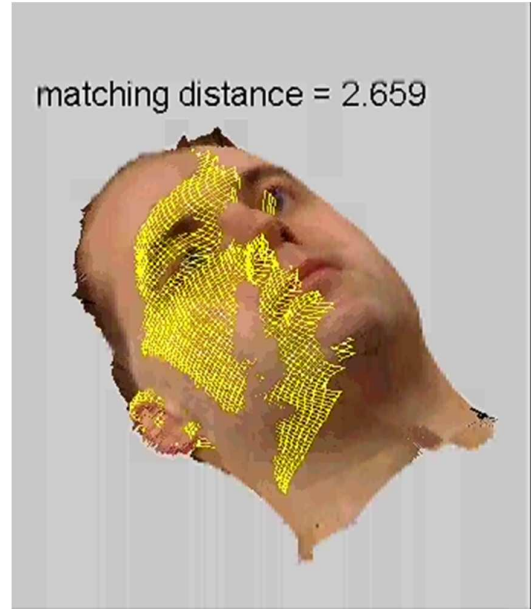
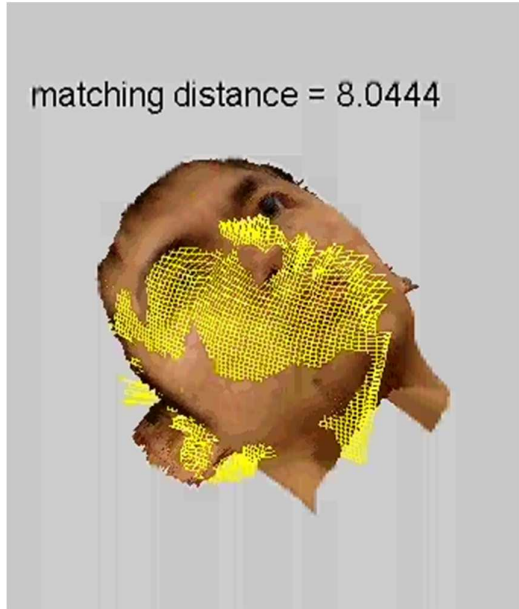
- Caratteristiche dei punti di controllo:
  - regioni meno deformabili a causa di espressioni del volto;
  - nel complesso offrono la maggior copertura possibile del volto.

## Esempi di allineamento (1)

Test scan



**Matching con modello vero**   **Matching con impostore**

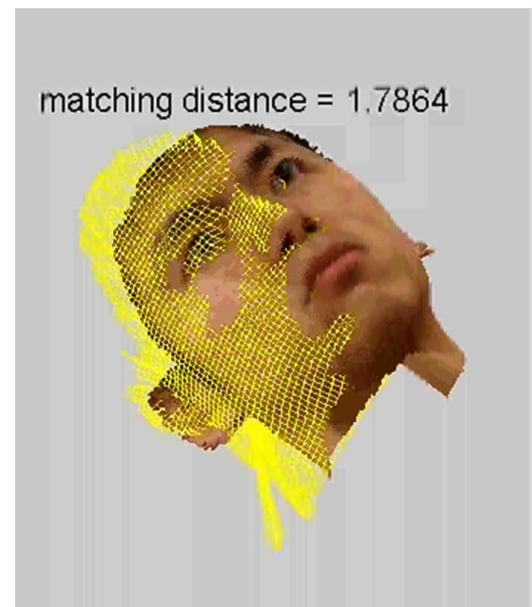
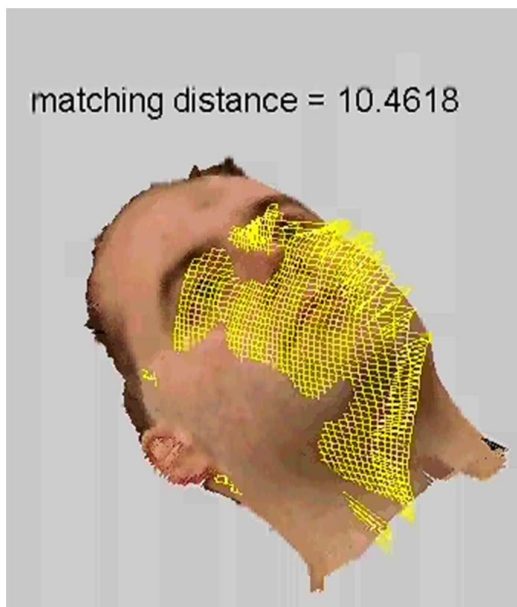


## Esempi di allineamento (2)

Test scan



**Matching con modello vero**   **Matching con impostore**



## Matching

- Per il matching si utilizza la distanza minimizzata dall'algoritmo ICP in fase di allineamento:

$$MD_{ICP} = \sqrt{\frac{1}{N_c} \sum_{i=1}^{N_c} d^2(\Psi(p_i), S_i)}$$

- $d(.,.)$  è la metrica di distanza di un punto da un piano;
  - $\Psi(.)$  è la trasformazione rigida applicata a ciascun punto di controllo  $p_i$  nell'immagine 2.5D di test;
  - $S_i$  è il piano tangente corrispondente a  $p_i$  nel modello 3D
  - $N_c$  è il numero di punti di controllo.
- Inoltre a ogni punto di controllo è associato un valore di shape index.
- Lo score finale si ottiene correlando i valore di shape index tra i due set di punti di controllo e la distanza ottenuta dall'algoritmo ICP, normalizzando opportunamente i valori e usando la regola della somma per la fusione.

## Risultati sperimentali: database



Training scans for 3D model construction

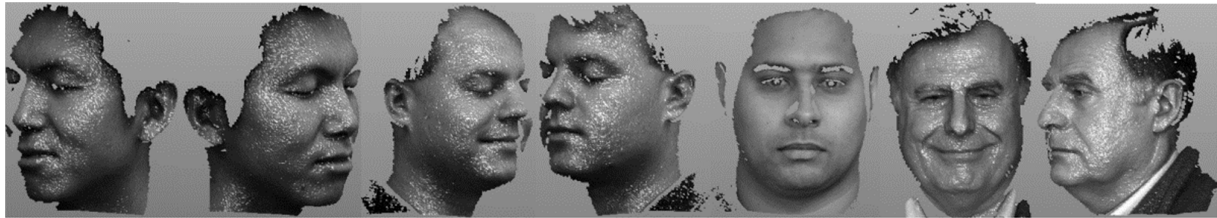


Test scans

- 100 modelli 3D
- Scansioni del volto multi-view
- Espressioni del volto variabili (es. sorriso)
- 598 immagini 2.5D di test

# Risultati sperimentali

Immagini di test  
(con espressione variabile e dati mancanti)



<i>Categoria immagine di test</i>	<i>Accuratezza</i>
Frontale, espressione naturale (99)	98% (2)
Profilo, espressione naturale (213)	97% (7)
Frontale, sorridente (98)	71% (28)
Profilo, sorridente (188)	78% (41)

Accuratezza  
complessiva  
87%

## Correct Matches



Test scans (top row), and the corresponding 3D models correctly matched. The 3D model is shown in a pose similar to the corresponding test scan

# Matching Errors



Test scans (top row), and the corresponding 3D models. The 3D model is shown in a pose similar to the corresponding test scan

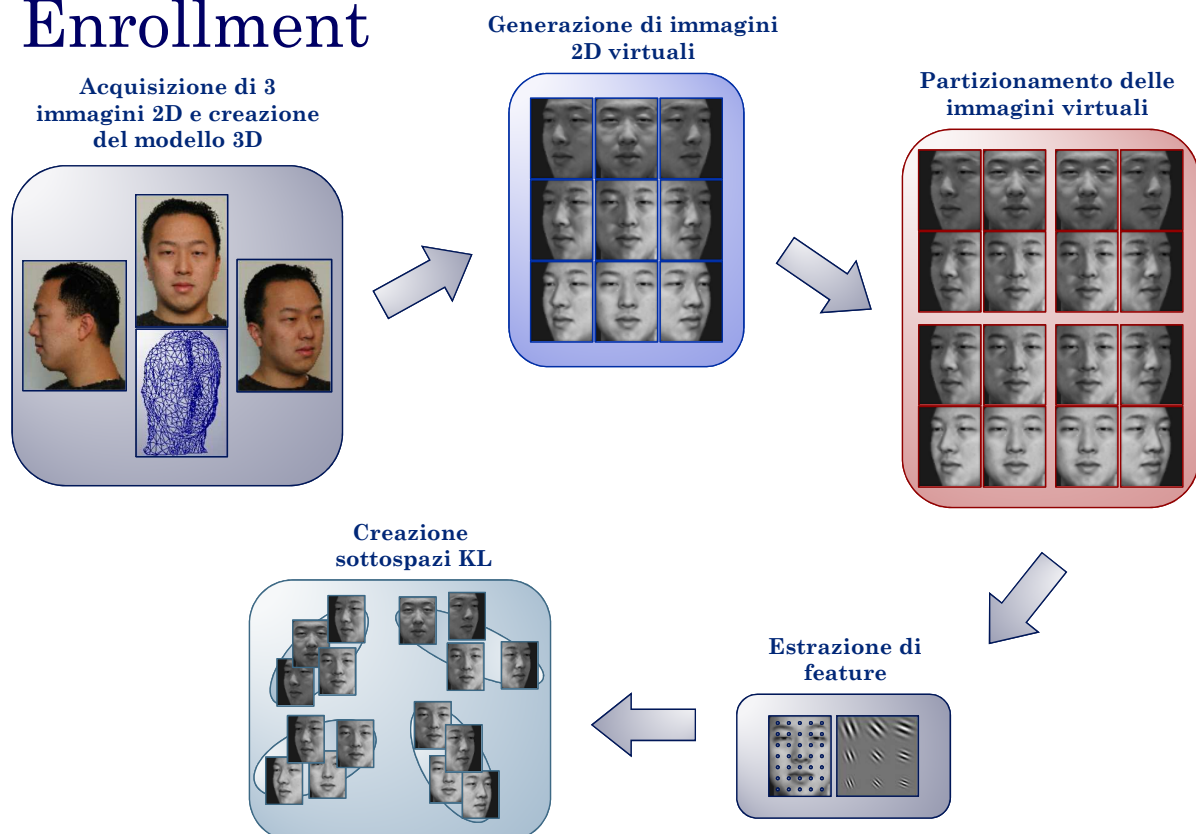
Una tecnica di  
riconoscimento 3D + 2D

# Riconoscimento 3D-2D

- Idea di base:
  - Usare il **modello 3D** solo in fase di **training** per la generazione di immagini del volto con pose e illuminazioni variabili.
  - Il **riconoscimento** avviene solo sulla base di **immagini 2D**.
- Vantaggi:
  - L'uso del modello 3D in fase di training permette la creazione di un template molto **robusto**.
  - Il riconoscimento è basato su semplici immagini 2D che non richiedono particolari apparecchiature di acquisizione.
  - La creazione del template degli utenti ha una complessità elevata (confrontabile con quella di altri metodi 3D), ma il riconoscimento è molto **efficiente** poiché basato su immagini 2D.

A. Franco, D. Maio, D. Maltoni, "2D Face Recognition based on Supervised Subspace Learning from 3D Models", *Pattern Recognition*, vol.41, no.12, pp.3822-3833, 2008.

## Enrollment

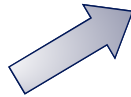
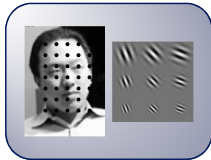


# Riconoscimento

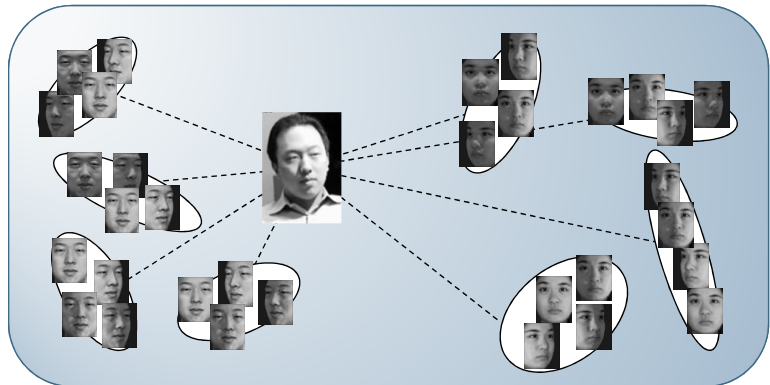
Acquisizione di una  
singola immagine 2D



Estrazione di  
feature



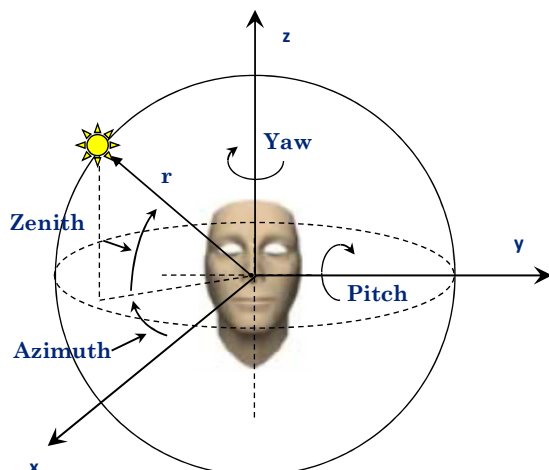
Riconoscimento in base alla minima  
distanza dallo spazio



$$I = \{I_{\alpha_i, \beta_j, \gamma_k, \theta_l} \mid i = 1, \dots, nPo, j = 1, \dots, nPv, k = 1, \dots, nLo, l = 1, \dots, nLv\}$$

## Generazione delle immagini virtuali (1)

- La generazione delle immagini virtuali è controllata da 4 parametri:
  - Posa:
    - Yaw e Pitch
  - Illuminazione:
    - Azimuth e Zenith
- Per ogni individuo viene generato un insieme di immagini  $I$  variando con passo discreto i 4 parametri.



## Generazione delle immagini virtuali (2)

- L'insieme di immagini di training generate è:

$$I = \{I_{\alpha_i, \beta_j, \gamma_k, \theta_l} \mid i = 1, \dots, nPo, j = 1, \dots, nPv, k = 1, \dots, nLo, l = 1, \dots, nLv\}$$

- I parametri  $\alpha$  e  $\beta$  controllano le variazioni di posa e rappresentano rispettivamente i valori di Yaw e Pitch.  $nPo$  e  $nPv$  indicano rispettivamente il numero di angoli considerati per Yaw e Pitch.
- i parametri  $\gamma$  e  $\theta$  controllano le variazioni di illuminazione e rappresentano rispettivamente i valori di Azimuth e Zenith.  $nLo$  e  $nLv$  indicano rispettivamente il numero di angoli considerati per Azimuth e Zenith.
- Un generico angolo  $\alpha_i, i=2, \dots, nPo$ , è dato da  $\alpha_i = \alpha_1 + (i-1) \times \Delta\alpha$  dove  $\Delta\alpha$  è il passo d'incremento.

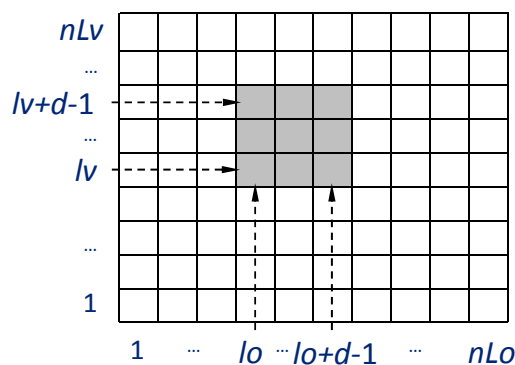
## Definizione dei sottoinsiemi di immagini

- Le immagini di training vengono organizzate in  $s$  sottoinsiemi (non necessariamente disgiunti):

$$P_{po, pv, lo, lv} = \left\{ X_{\alpha_i, \beta_j, \gamma_k, \theta_l} \mid \begin{array}{l} i = po, \dots, po + d - 1, j = pv, \dots, pv + d - 1, \\ k = lo, \dots, lo + d - 1, l = lv, \dots, lv + d - 1 \end{array} \right\}$$

$$po \in [1, \dots, nPo - d + 1] \quad pv \in [1, \dots, nPv - d + 1] \quad lo \in [1, \dots, nLo - d + 1] \quad lv \in [1, \dots, nLv - d + 1]$$

- dove  $d$  rappresenta il numero di pose/illuminazioni consecutive considerate.





## I sottoinsiemi

- Il numero totale di sottoinsiemi per ogni soggetto è determinato dal parametro  $d$  e dalla sovrapposizione ( $ov$ ) tra sottoinsiemi adiacenti:

$$s = \left\lceil \frac{nPo - ov}{d - ov} \right\rceil \times \left\lceil \frac{nPv - ov}{d - ov} \right\rceil \times \left\lceil \frac{nLo - ov}{d - ov} \right\rceil \times \left\lceil \frac{nLv - ov}{d - ov} \right\rceil$$

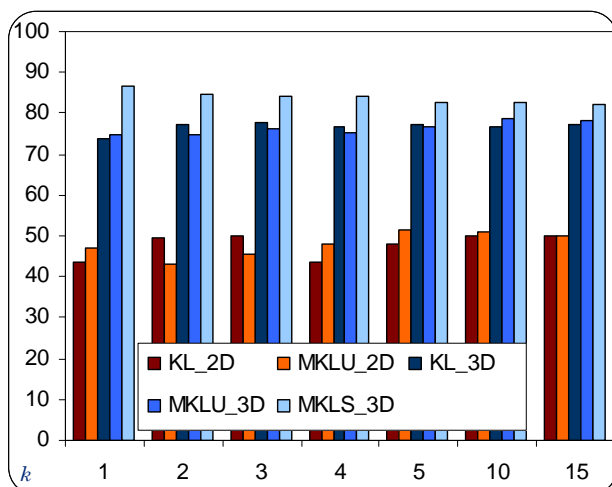


## Prove sperimentali: database MIT-CBCL

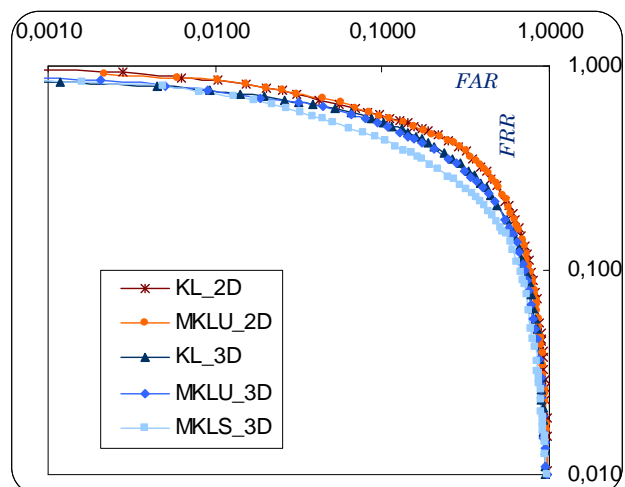
- 59 immagini 2D a elevata risoluzione, alcune delle quali usate per la creazione dei modelli 3D.
- 3240 immagini 2D virtuali (324 per ogni individuo) generate dai modelli 3D.
- 2000 immagini 2D di test, con notevoli cambiamenti di posa, scala e illuminazione.
- Per la localizzazione del volto è stato usato il localizzatore di Viola/Jones.



## Risultati sperimentali



Accuratezza



Curva DET

## Prove sperimentali



ID	RR	EER
0000	100.00	12.97
0001	100.00	7.61
0002	100.00	9.67
0003	99.50	13.61
0004	99.00	15.03
0005	97.50	17.17
0006	100.00	9.47
0007	100.00	20.19
0008	95.50	22.97
0009	85.00	39.58

Immagini di test  
acquisite 6 mesi dopo il  
training...