

Ricostruzione tridimensionale a partire da immagini

Quando i matematici non entrano in aula

Luca Magri

3Dflow



3Dflow è una piccola società privata spin-off dell'Università degli studi di Udine di consulenza e di produzione software, operante nel campo della Computer Vision e Image Processing.

- Modellazione 3D automatica a partire da foto
- 3D video processing
- Realtà aumentata

Andrea Alessi - Software Engineer

Filippo Fantini - Software Engineer

Simone Fantoni - Technical Director

Andrea Fusiello - Scientific Advisor

Martino Giovanelli - External Software Engineer

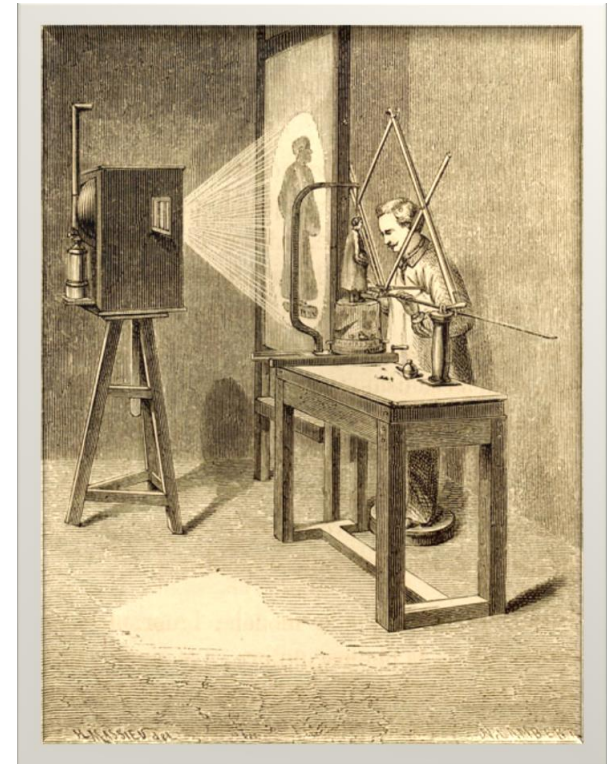
Luca Magri - External Researcher

Marco Schivi - Software Engineer

Yash Karan Singh - CEO

Roberto Toldo - Technical Director

Giacomo Vianini - Junior Sales Rep

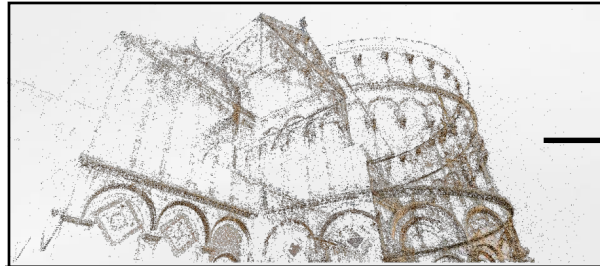


Francois Willeme, inventore della fotoscultura - 1860

3DF Zephyr

- Video pisa

3DF Zephyr



nuvola di punti sparsa



nuvola di punti densa



mesh

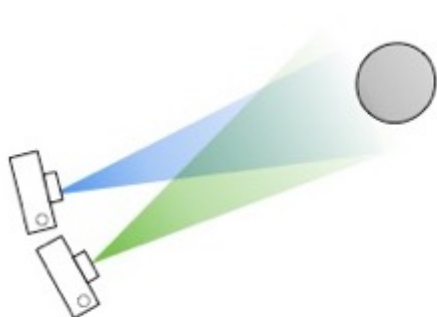
Cosa sappiamo fare con due immagini?

Dato un numero sufficiente di punti corrispondenti, la costruzione di un modello 3D può essere realizzata:

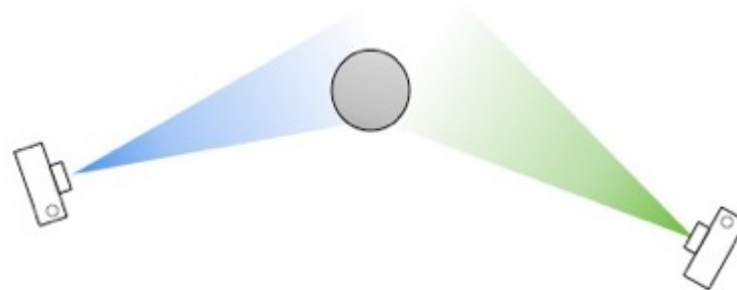
- Calcolando la matrice essenziale E
- Fattorizzando E per stimare la rotazione e la traslazione che collega le due camere
- Istanziare una coppia di camere
- Calcolare la posizione 3D tramite triangolazione.

E è definita a meno di un fattore di scala e il modello ottenuto eredita questa ambiguità (depth-speed ambiguity)

Problema:



Ricostruzione parziale

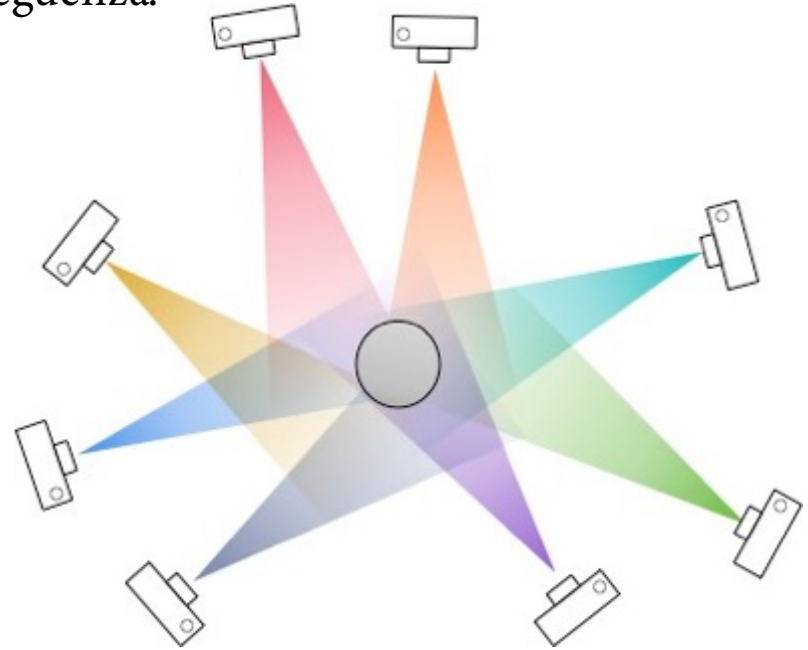


Poche corrispondenze

Non c'è due senza... n

Possiamo aumentare il numero di camere che osservano l'oggetto di interesse. La geometria epipolare si generalizza di conseguenza:

- 3 camere: tensore trifocale
- 4 camere: tensore quadrifocale
- ...
- n camere: tensore n -focale



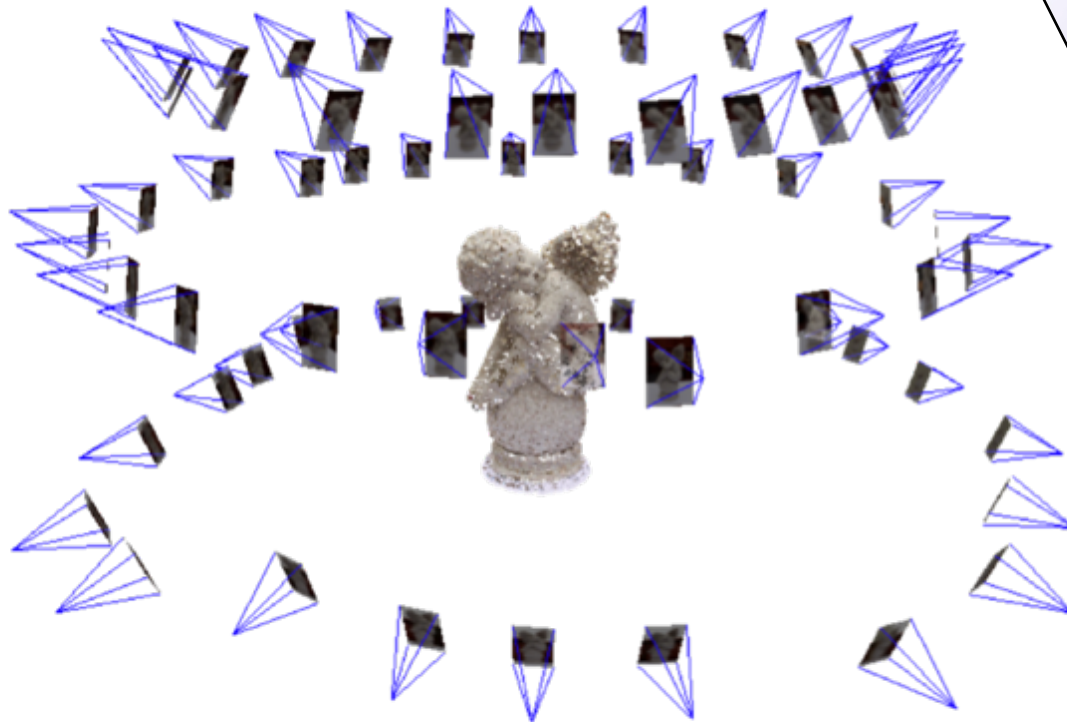
Si può dimostrare che i vincoli che riguardano n camere possono essere fattorizzati in termini di vincoli che coinvolgono coppie, triple o quadruple di immagini.

Quindi sfutteremo i vincoli bilineari tra *coppie* di immagini (meno onerosi da calcolare).

Structure And Motion

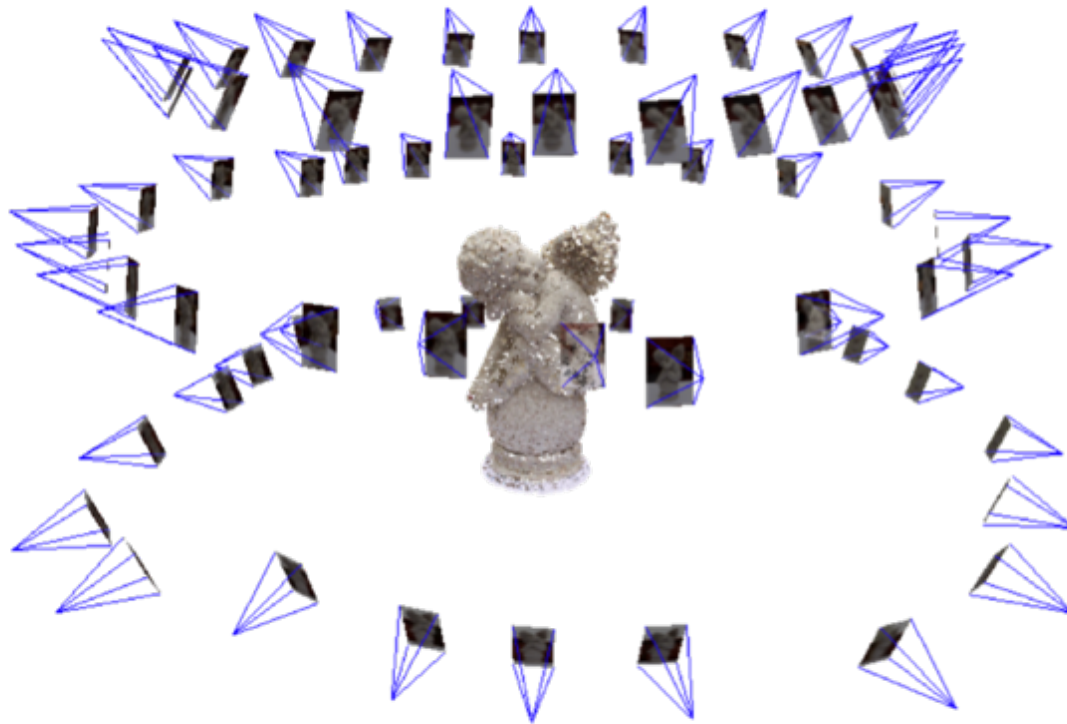
Analisi di molte immagini con sovrapposizione per il recupero di

- *moto*: i parametri esterni di un insieme di fotocamere
- *struttura*: la posizione dei punti 3D della scena



Samantha al lavoro

- *Estrazione dei punti salienti*
- *Accoppiamento*
- *Clustering*
- *Hierarchical Structure and motion*
- *Bundle adjustment*

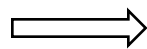


Punti salienti

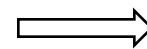
Non tutti i pixel sono facilmente distinguibili all'interno delle immagini.
I punti di un'immagine che si distinguono chiaramente dal loro vicinato sono detti *punti salienti* (ad esempio: spigoli, macchie, forti variazioni di colore...)



Estrazione



Descrizione



Accoppiamento

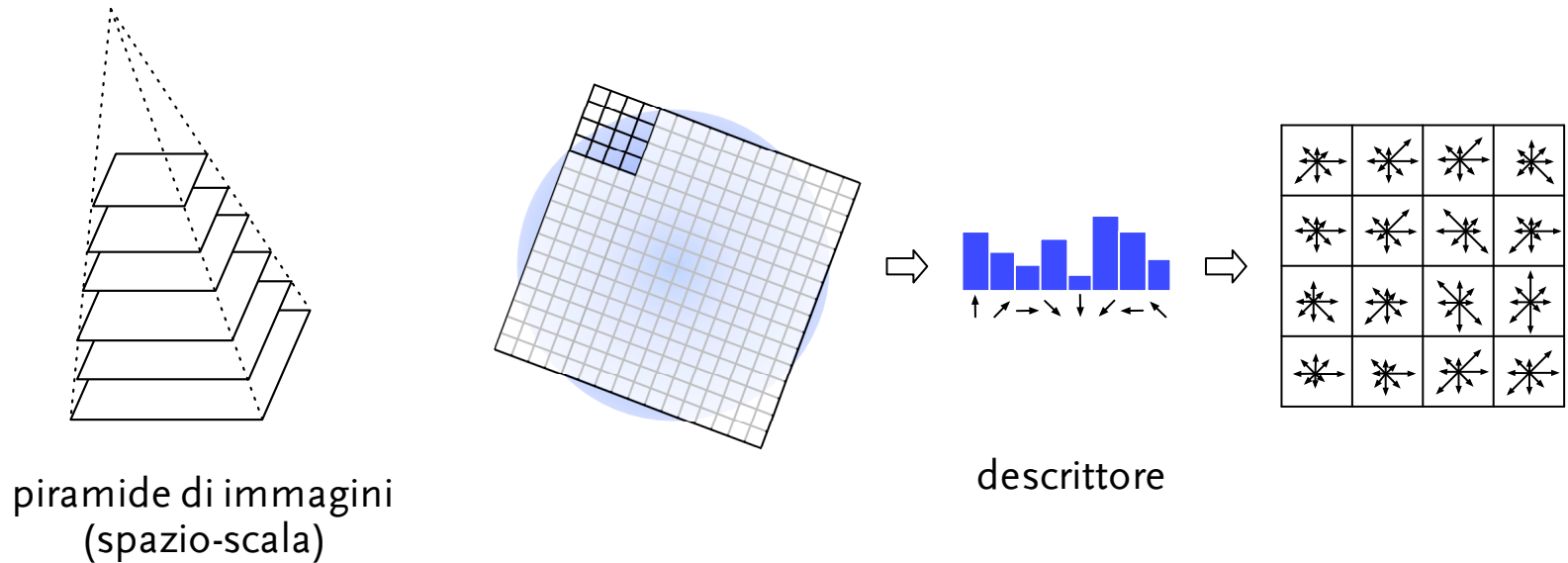
in ogni immagine si
cercano i punti salienti

ogni regione intorno ai
punti salienti viene descritta
in forma compatta e stabile
(invariante)

si confrontano i
descrittori per trovare
delle corrispondenze

Estazione e descrizione dei punti salienti

Ogni immagine viene riprodotta a diverse risoluzioni e lavorata con opportuni filtri, per mettere in evidenza i dettagli a diverse scale.

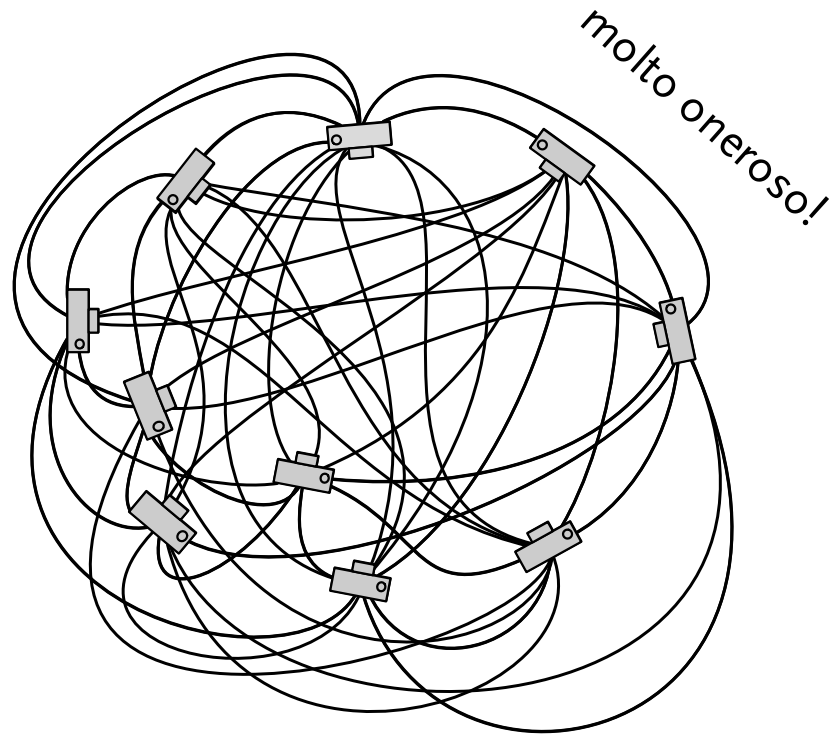


I punti estremanti nello spazio-scala rispetto a opportuni operatori vengono considerati punti salienti. Ad ogni punto saliente viene associato:

- un *descrittore* (analizzando le orientazioni del gradiente nel suo vicinato) codificato in forma vettoriale. È possibile confrontare tra loro i descrittori.
- uno *score* che ne quantifica la bontà.

Accoppiamento

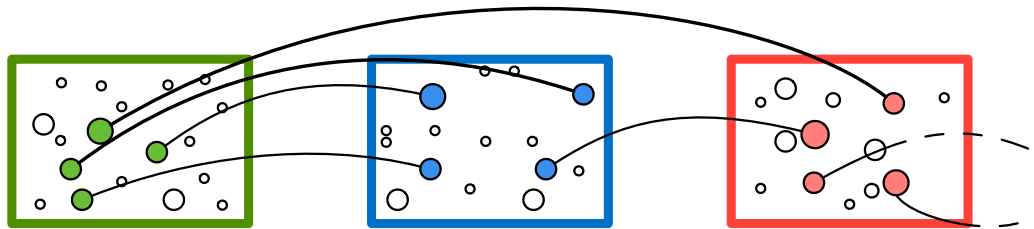
Le corrispondenze andrebbero cercate tra tutte le possibili coppie di immagini.



Invece di cercare le corrispondenze tra tutte le possibili coppie di immagini, cerchiamo di sfolire il numero di coppie plausibili...

Accoppiamento – grafo epipolare

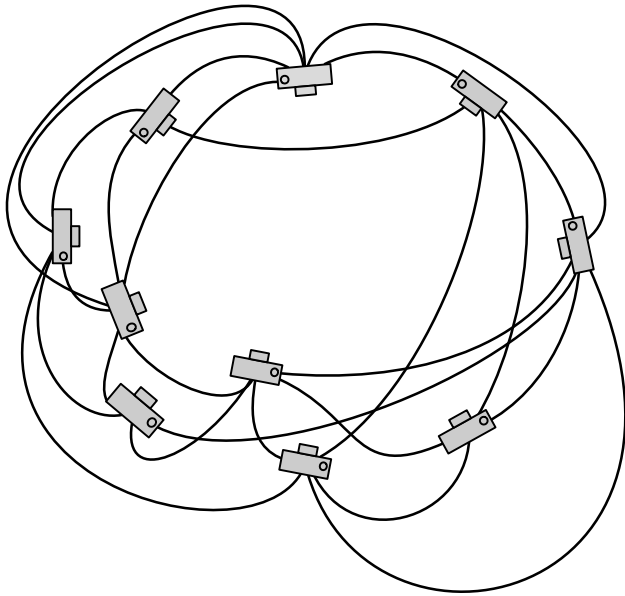
- Prendiamo in esame solo i migliori h punti salienti per ogni immagine.
- Accoppiamo in modo efficiente ogni punto saliente con il più simile (approximate nearest neighbour).



- Per ogni coppia di immagini contiamo quante corrispondenze così ottenute hanno in comune.

Questa matrice può essere interpretata come la matrice di adiacenza di un grafo pesato...

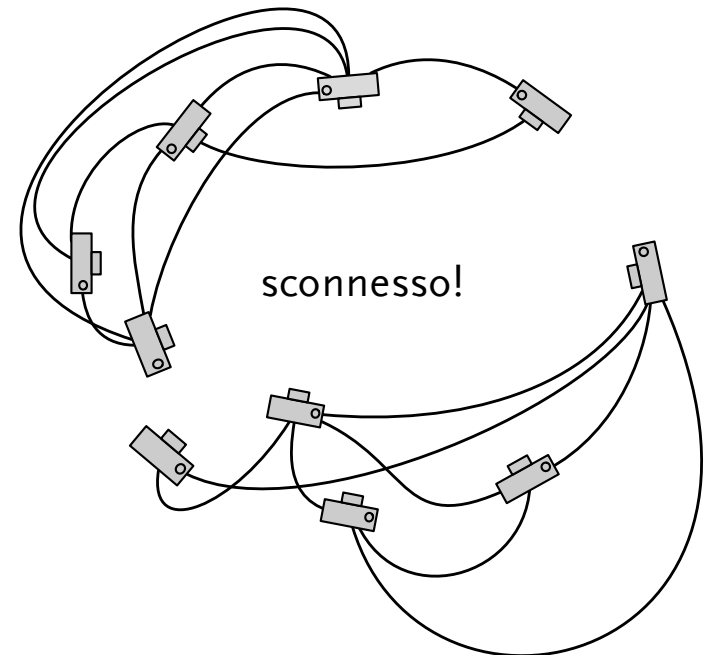
Accoppiamento – grafo epipolare



Gli archi sono ancora tanti...
ne vorremmo rimuovere altri, ma dobbiamo
stare attenti a non creare gruppi di immagini
simili disconnessi tra di loro.

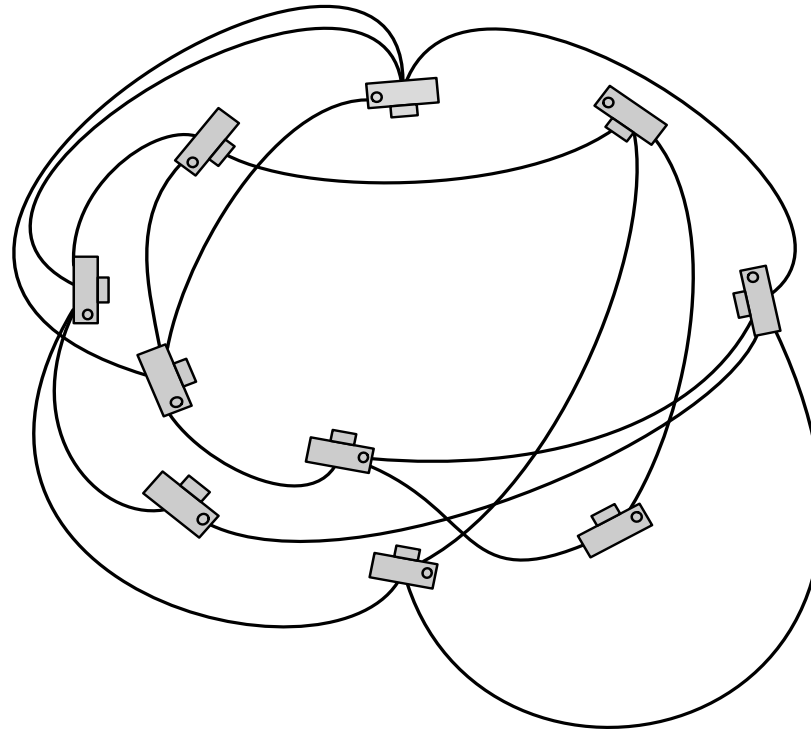
L'obiettivo diventa quello di cercare un

grafo k -connesso per archi:
comunque vengano tolti k archi, il grafo
rimane connesso.



Accoppiamento – grafo epipolare

Algoritmo per approssimare un grafo k -connesso per archi



1. Trovare un albero ricoprente con peso massimo
2. Rimuovere gli archi dell'albero
3. Ripetere k volte

Accoppiamento – raffinamento

I punti salienti vengono accoppiati cercando i punti più simili nelle immagini suggerite dal grafo epolare.

- Problema: non tutti i punti che hanno un descrittore simile sono punti proiezione dello stesso punto 3D (es: strutture ripetute).
- Validazione geometrica: per ogni coppia di viste si calcola la matrice fondamentale F , e due punti x, x' vengono accoppiati solo se soddisfano il vincolo epolare:

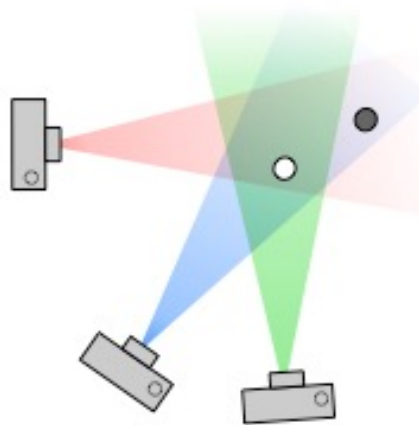
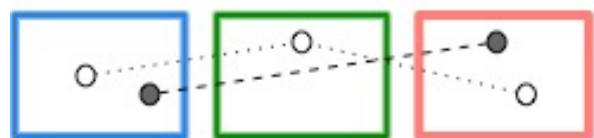
$$x^\top F x' \leq \epsilon$$



Accoppiamento – raffinamento

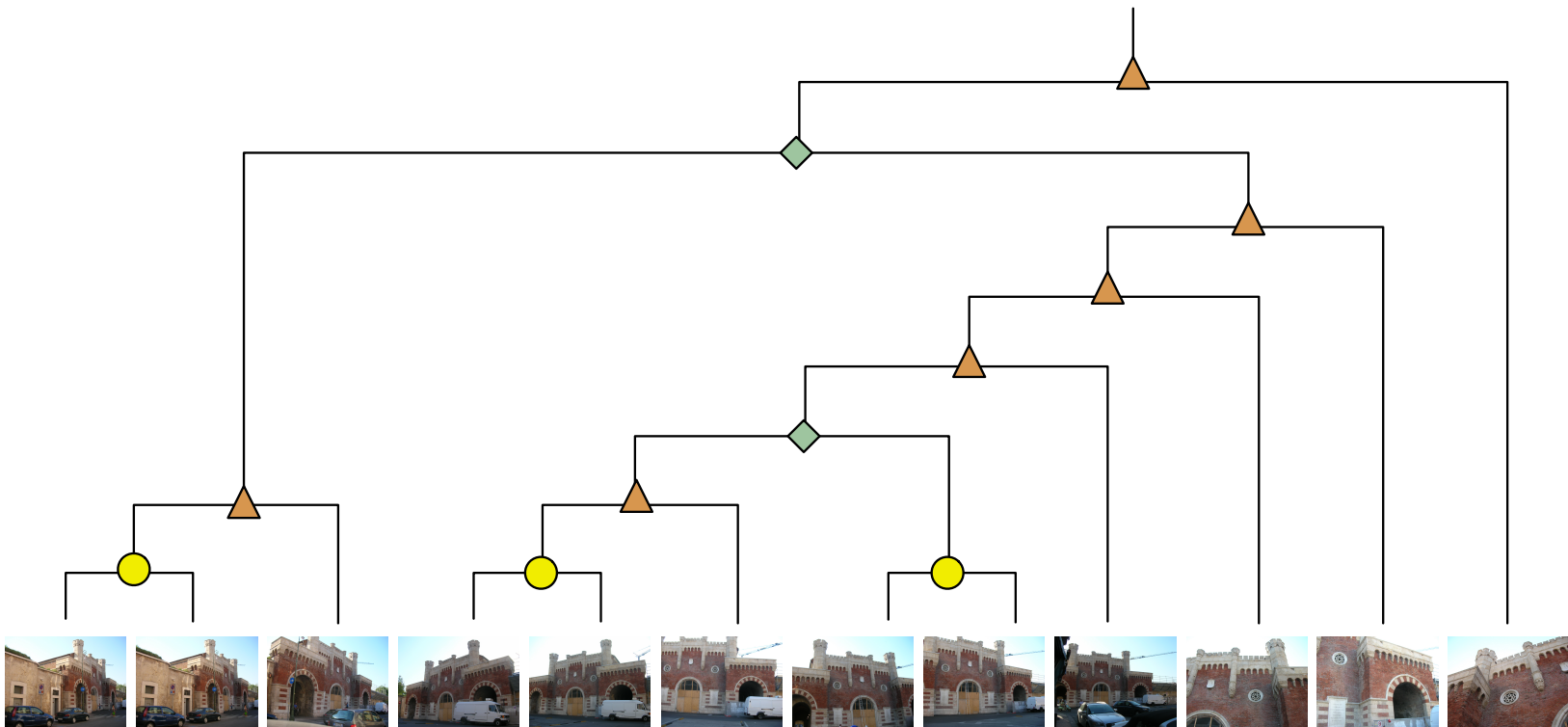


Le coppie di punti accoppiati in più immagini vengono concatenate in *tracce*: insiemi di punti corrispondenti.



Clustering gerarchico

Le immagini vengono raggruppate gerarchicamente per similarità (tenendo conto del numero di corrispondenze e di come sono distribuite nell'immagine) dando origine ad un albero in cui le foglie sono le immagini ed i nodi ne rappresentano raggruppamenti (*cluster*) via via più grandi.



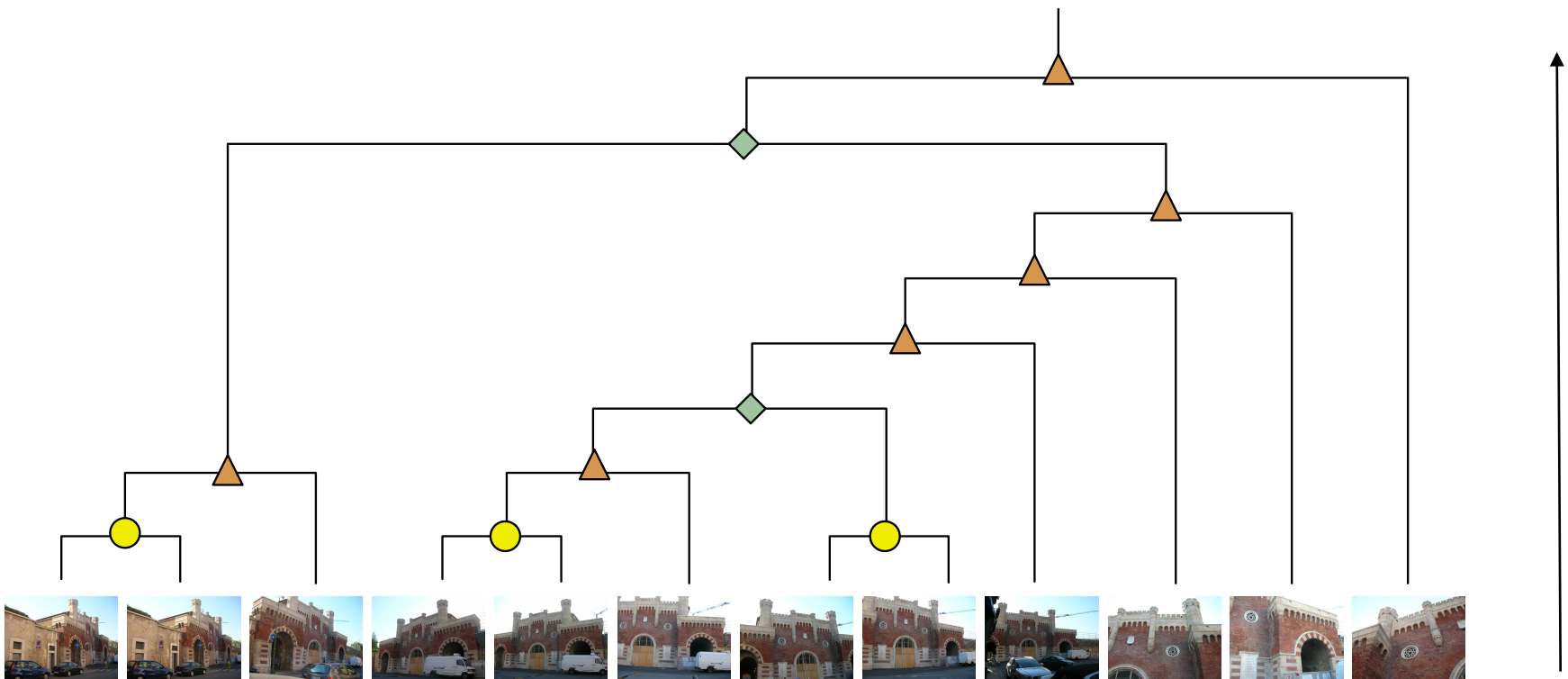
Clustering gerarchico

Partendo dal basso ad ogni nodo viene eseguita una particolare azione:

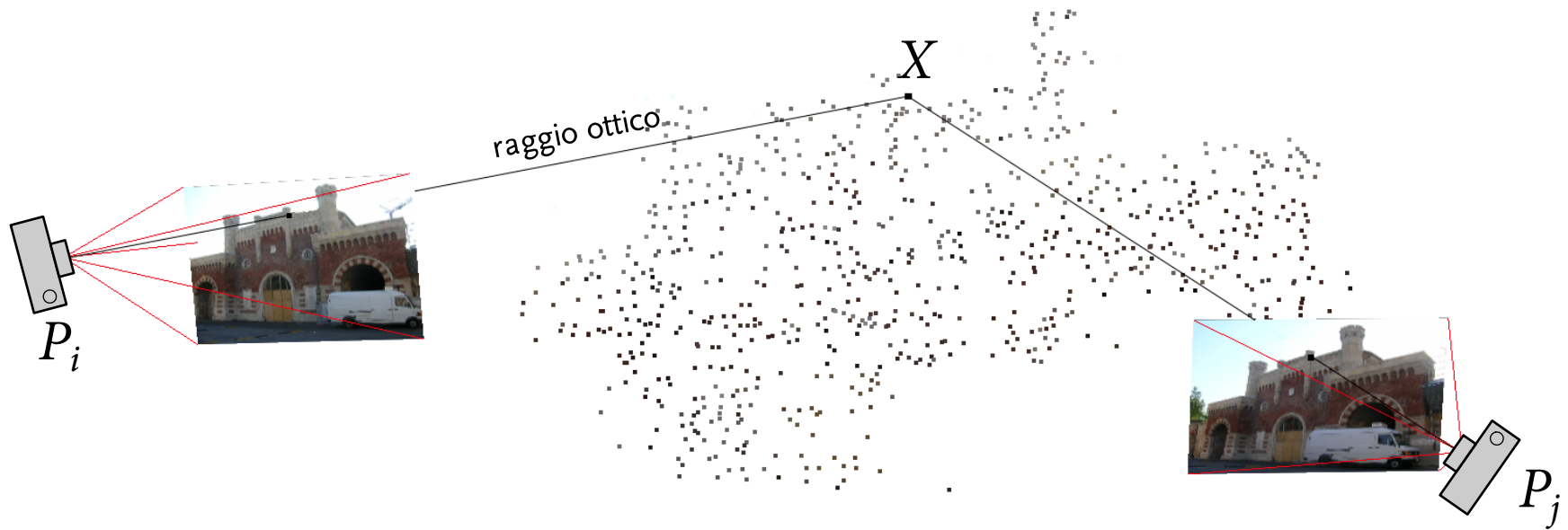
● Ricostruzione stereo

▲ Aggiunta di una vista

◆ Fusione di due modelli



Clustering gerarchico – ricostruzione stereo ●



- calcolo della geometria epipolare tra due viste
- stima delle camere
- stima dei punti 3D per triangolazione.

Triangolazione

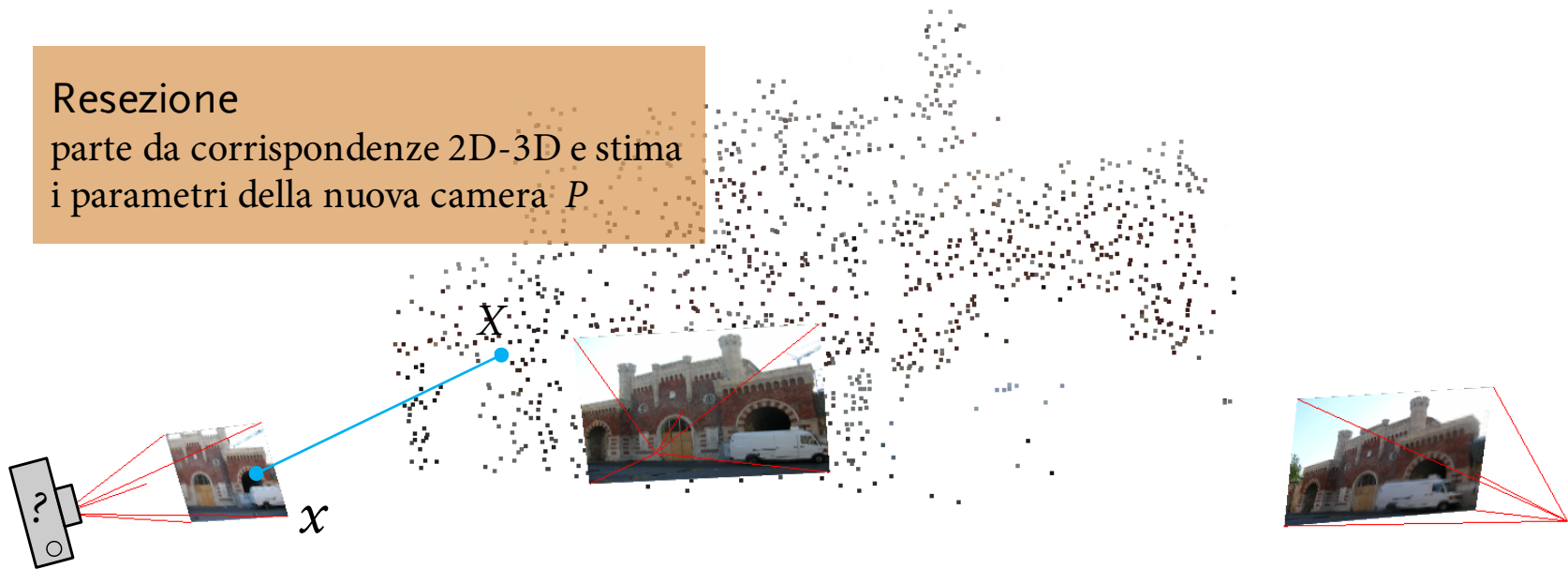
a partire da corrispondenze 2D e dalle camere stima la posizione dei punti 3D



Clustering gerarchico – aggiunta di una vista

Resezione

parte da corrispondenze 2D-3D e stima i parametri della nuova camera P



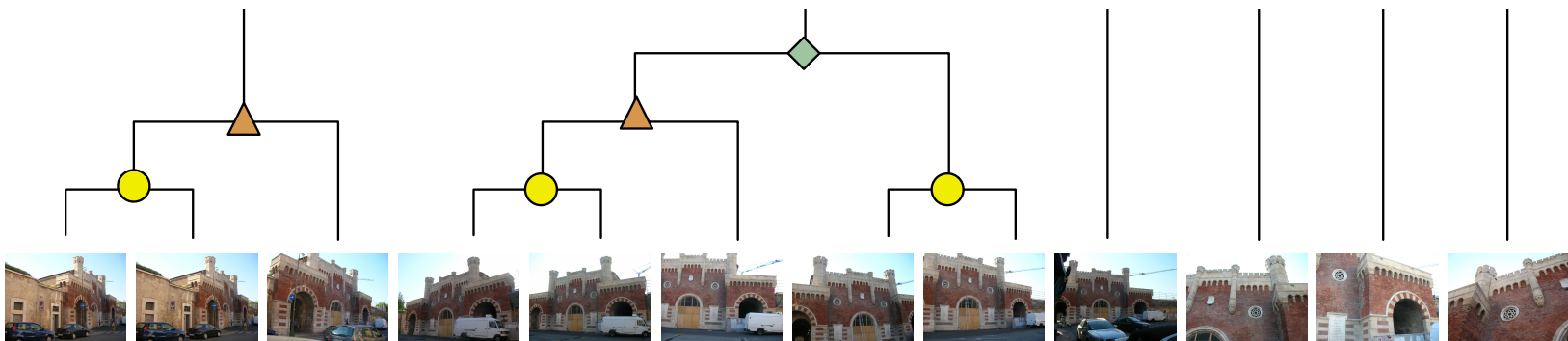
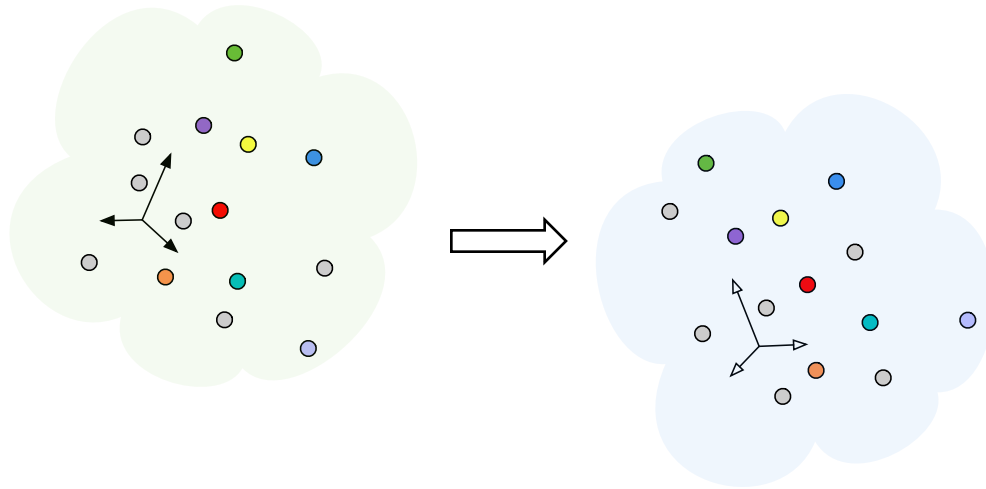
- Le corrispondenze relative ai punti già ricostruiti sono utilizzate per inferire corrispondenze 2D-3D. Sulla base di queste, sono stimate le P rispetto alla ricostruzione esistente.
- Le tracce sono eventualmente aggiornate e punti 3D vengono aggiunti.



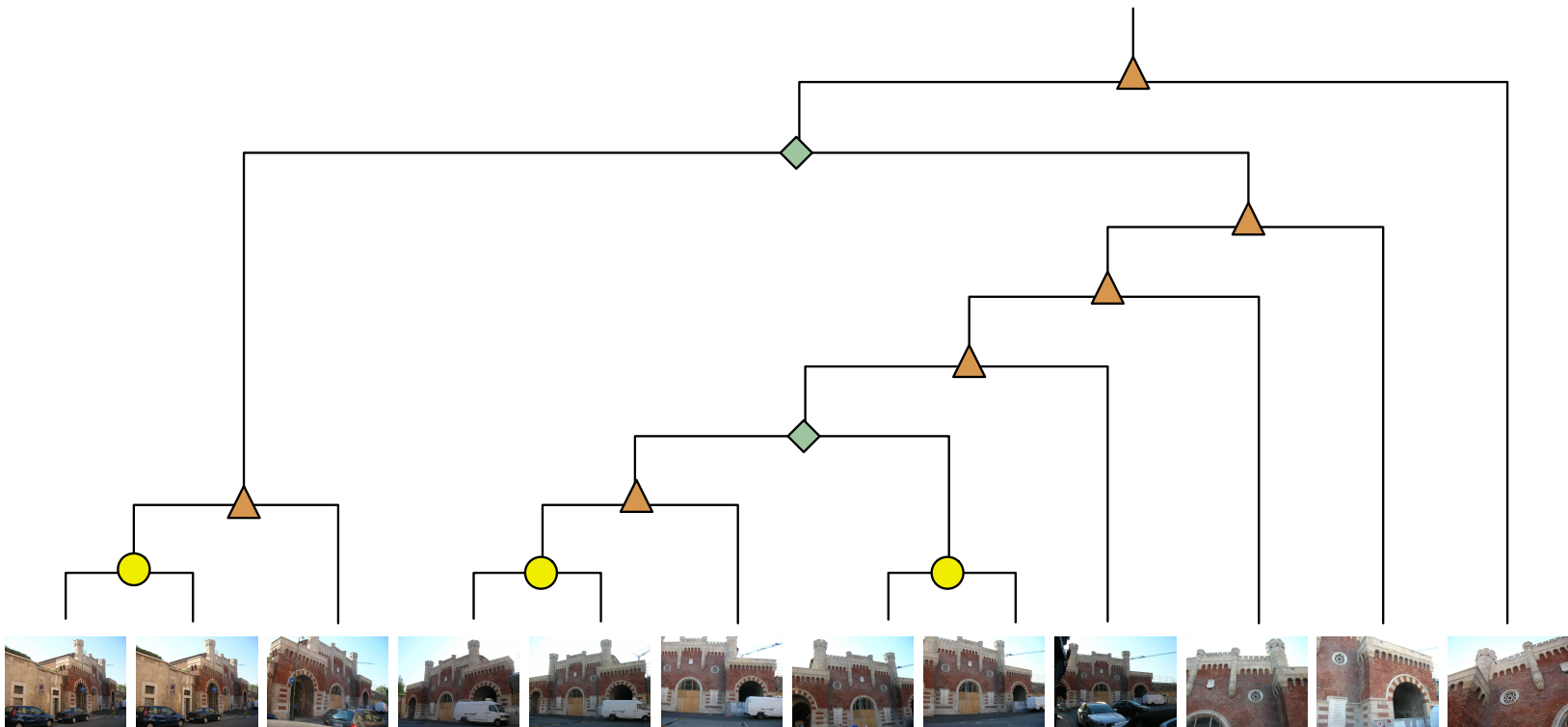
Clustering gerarchico – fusione di due modelli

Registrare due modelli parziali nello stesso sistema di riferimento.

Trovare la trasformazione (similitudine per via dell'ambiguità di scala) che porta il primo modello nel secondo.



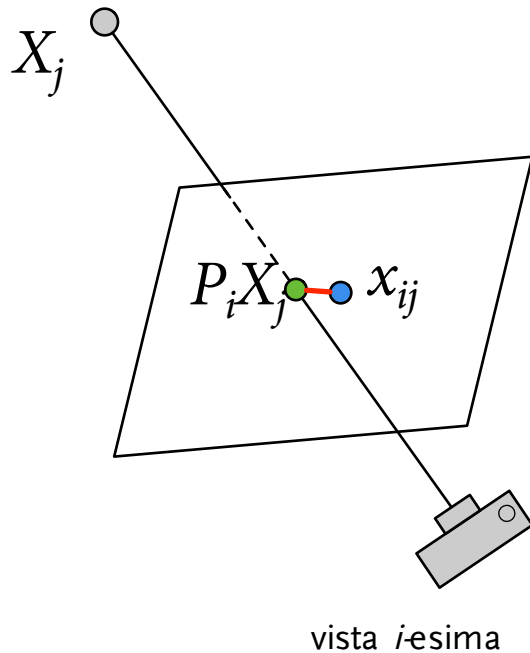
Clustering gerarchico



Bundle Adjustment

Problema: possibile accumulazione degli errori, che porta ad una deriva del risultato: errori sui parametri stimati delle prime fotocamere e punti influenzano gli elementi che vengono aggiunti successivamente.

Bundle adjustment: si cerca di minimizzare l'errore globale di riproiezione di tutti i punti 3D nelle immagini in cui sono visibili.



$$\min \sum_i \sum_j d(P_i X_j, x_{ij})^2$$

Autocalibrazione

Per semplicità abbiamo supposto di conoscere sempre i parametri interni della camera in modo da ottenere una ricostruzione euclidea della scena.

In pratica questa informazione non è sempre disponibile ma Zephyr riesce a ricavarla:

- direttamente dai dati EXIF delle fotografie
- oppure, sfruttando una procedura di autocalibrazione, stimano i parametri di K

In questo modo il modello che si ottiene è definito a meno di un fattore di scala.

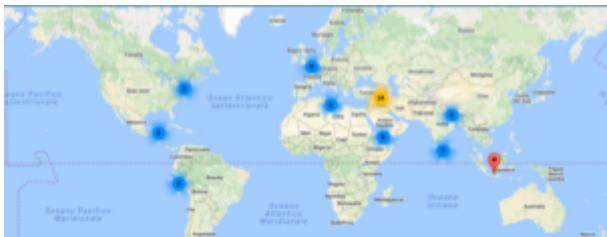
Rekrei – (Progetto Mosul)



Febbraio 2015 - Mosul (Iraq): distruzione di reperti archeologici mesopotamici conservati nel Mosul Museum of artifacts ad opera dell'ISL

Museo virtuale

- Usare computer vision e fotogrammetria per la conversazione patrimonio culturale iracheno e siriano distrutto dallo stato islamico.
- Matthew Vincent e Chance Coughenour creano una piattaforma crowdsourcing pensata per raccogliere e utilizzare le immagini di reperti archeologici e monumenti, rendendole disponibili per la ricostruzione digitale.



Riferimenti e risorse

- <http://www.3dflow.net/it/> - pagina web di 3DFlow
- <http://www.diegm.uniud.it/fusiello/demo/samantha/> - pagina web su Samantha
- Andrea Fusiello, *Visione computazionale: Tecniche di ricostruzione tridimensionale*, Franco Angeli, 2013
- <https://rekrei.org> - mosul project

Grazie per l'attenzione :-)

