DOTTORATO DI RICERCA IN INGEGNERIA DELLE TELECOMUNICAZIONI
Dipartimento di Ingegneria dell'Informazione
UNIVERSITÀ DEGLI STUDI DI BRESCIA

Robust and computationally effective multi-view alignment techniques for high-end 3D modeling pipelines

Francesco Bonarrigo

Advisor: Dr. Alberto Signoroni
Coordinatore: Prof. Riccardo Leonardi

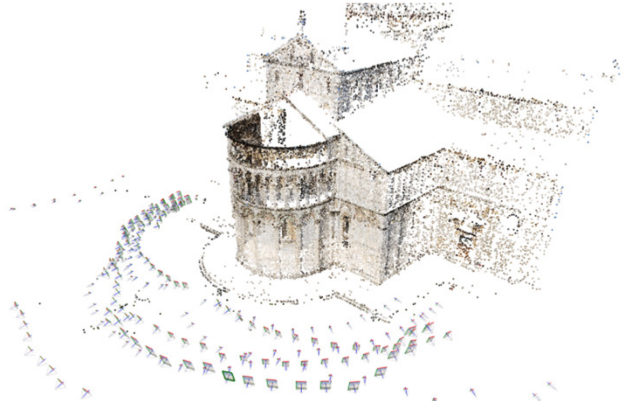
Applications for 3D data today

- *3D data today is more and more **widespread**, finding useful applications in a plethora of very different **application fields**:*
  - *Manufactory*
  - *Medical*
  - *Entertainment*
  - *Cultural heritage*
  - *Topography*

Acquisition of 3D data

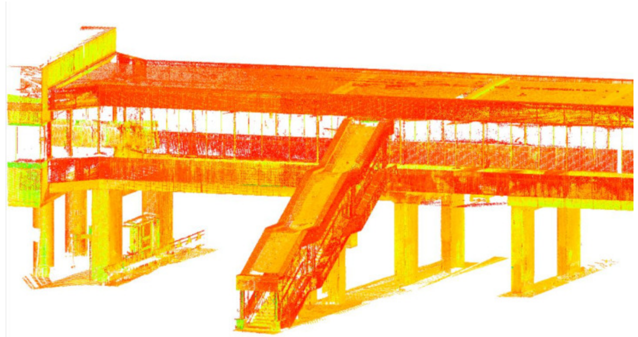
Approaches

Passive → *Photogrammetry*



Active

Time of flight



Triangulation

3D data alignment

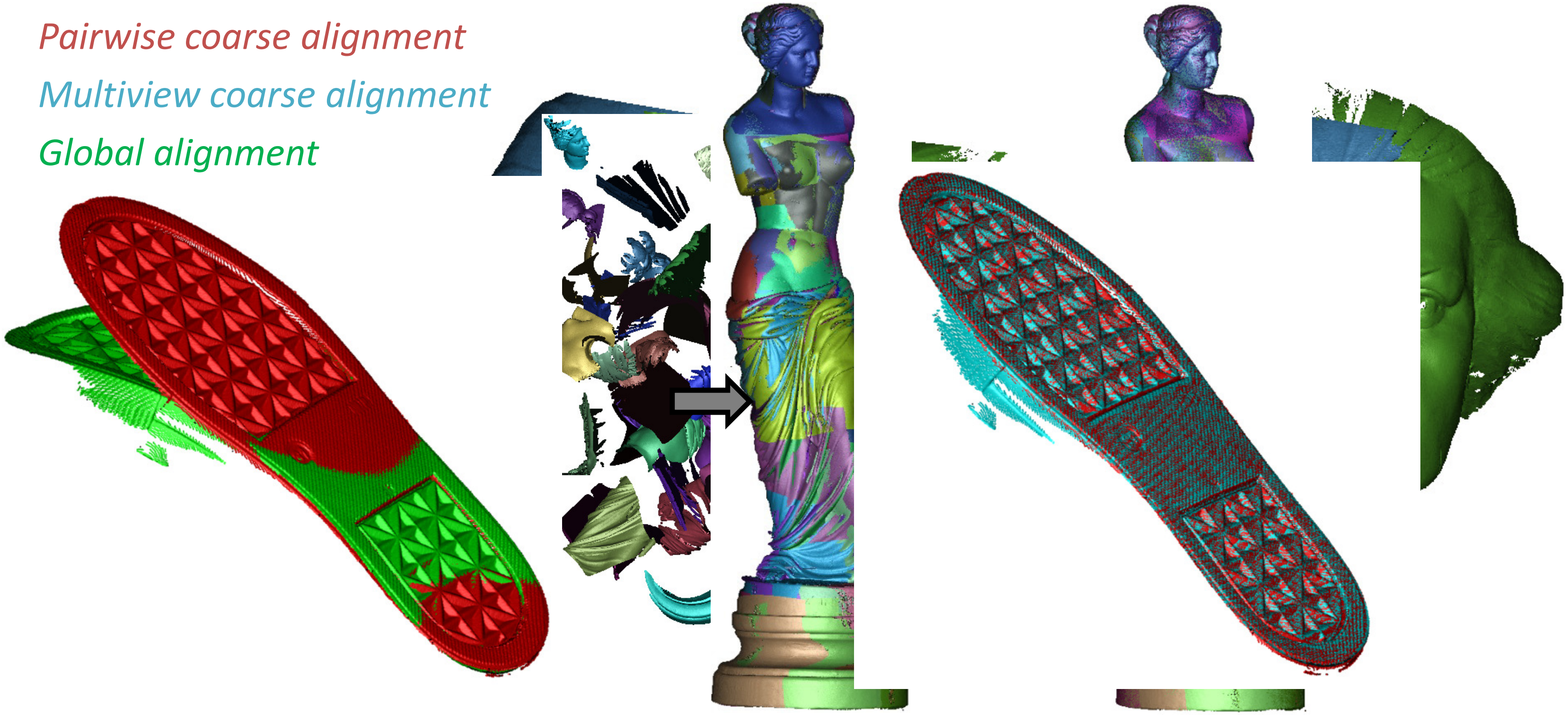
- *Whatever the acquisition approach, multiple data need to be acquired in order to cover the entire area of interest*

- *The scans need to be accurately realigned before generating the 3D model of the area of interest*

Thesis outline

The thesis presents:

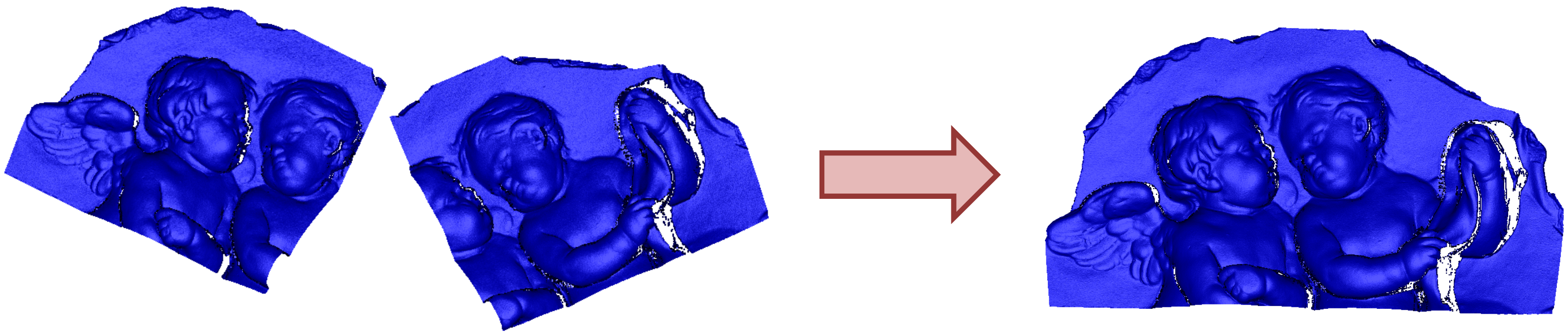
- *A pipeline to robustly and efficiently reconstruct sets of dense 3D scans:*
  - *Pairwise coarse alignment*
  - *Multiview coarse alignment*
  - *Global alignment*



- *In addition, a technique to perform a deformable alignment between two scans is presented*

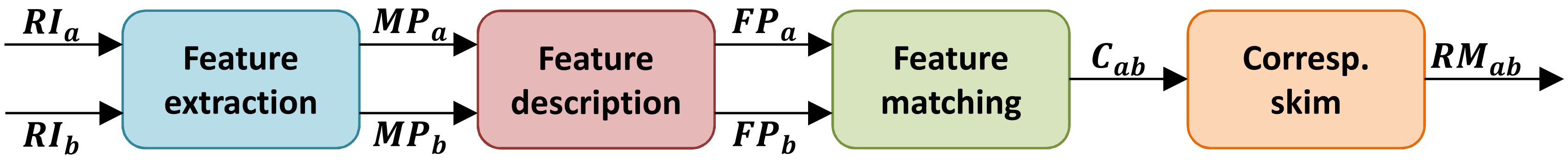
Pairwise coarse alignment

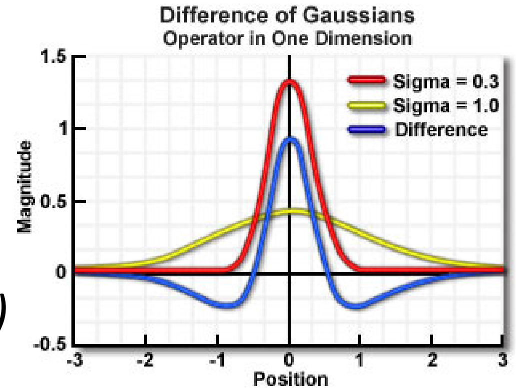
- *Once each scan has been acquired, it needs to be correctly aligned into a global reference system in order to reconstruct the final 3D model*



- ***Goal:** develop an automatic pairwise coarse alignment technique which:*
  - *make no assumption on the initial position*
  - *assume a portion of overlap between the scans*

Insight





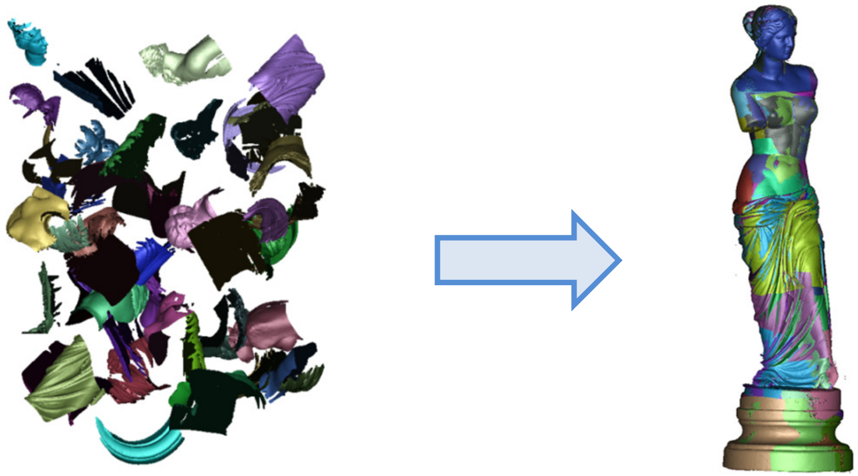
- *Feature-based alignment pipeline:*
  1. ***feature extraction**: performed through a 3D extension of the Difference of Gaussians (DoG)*
  2. ***feature description**: performed through differential information between feature and neighbour data*
  3. ***feature matching**: compute all matches between the two feature sets*
  4. ***correspondence skim**: until a small set of correct elements is found*

Key features

- *The approach is **robust**, **fast**, and **reliable**:*
  - *w.r.t. different datasets and different degrees of feature numerosity*
- *Multi-scale data analysis:*
  - *the detection of **features at multiple scales** allow to restrict the matching process to features detected at the same scale (speedup)*
- *Multi-stage correspondence skim procedure:*
  - *correspondence skim is performed through a **3-step procedure**, so to balance the computational requirements of each stage*
  - *Such approach **greatly outperforms RANSAC**, since the number of inliers can be very limited (down to 5%)*

Multi-view coarse alignment

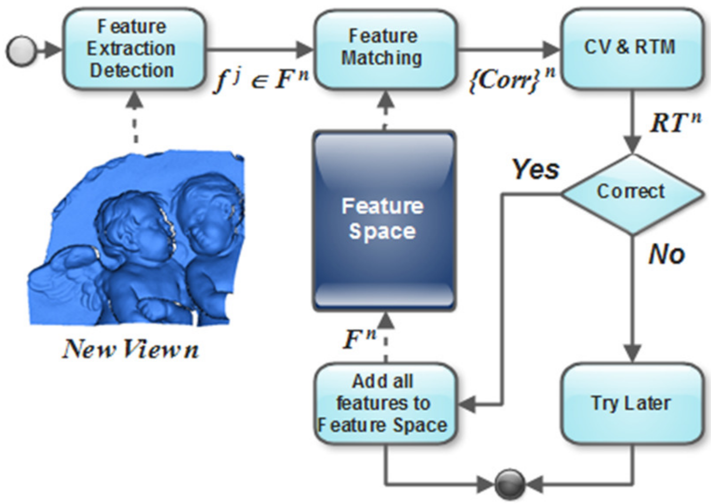
- *Many feature-based coarse alignment techniques are designed as:*
  - *pairwise approaches (limiting assumptions on acquisition path)*
  - *direct extensions of pairwise approaches (exhaustive search)*

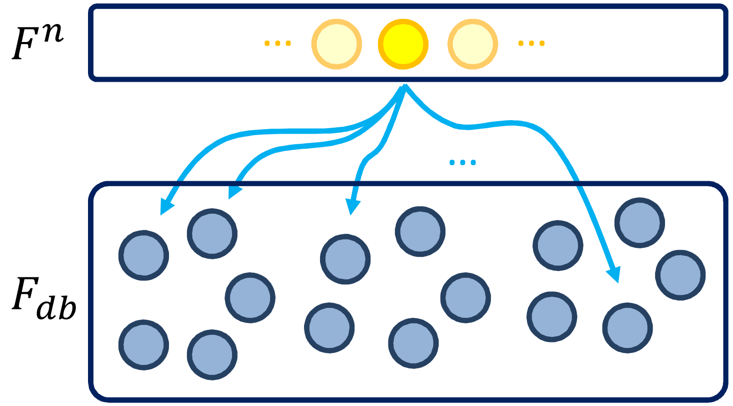


- ***Goal:*** *a clustering approach for feature-based multi-view alignment that:*
  - *increase matching speed (and nearly constant w.r.t. scans number)*
  - *does not degrade performance w.r.t. exhaustive search*
  - *constraint-free acquisition path*

Insight

- *A general exhaustive search feature-based coarse alignment scheme:*





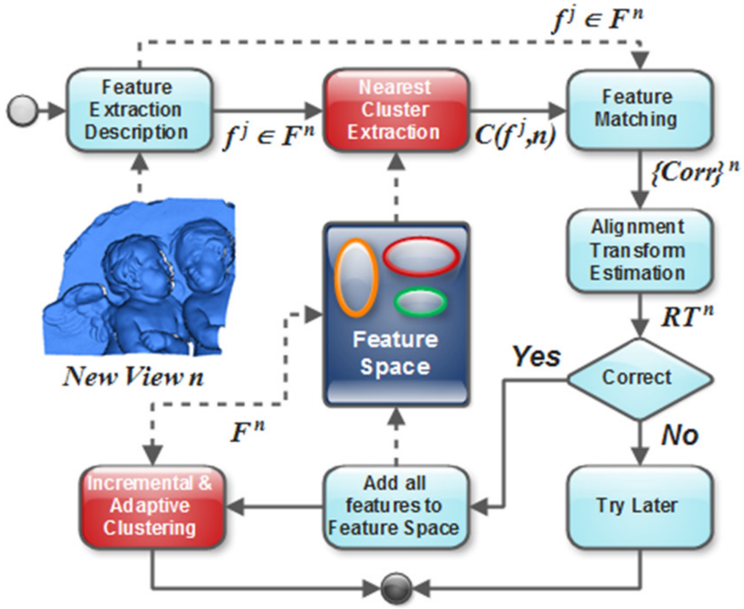
- *Exhaustive matching time:*

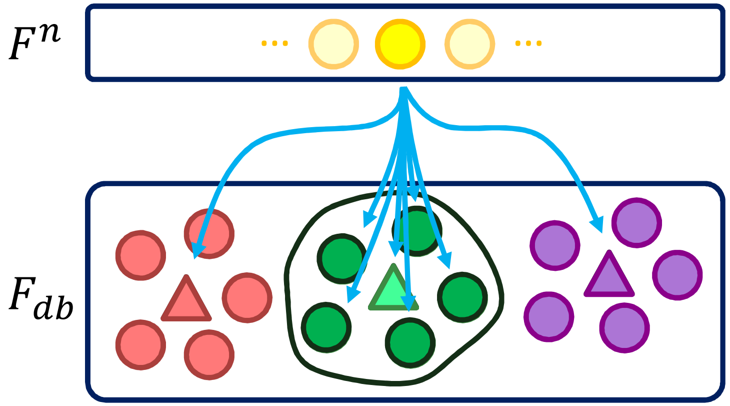
$$T_{F^n \rightarrow F_{db}} \approx \alpha \cdot \overbrace{\tilde{f}}^{\approx |F^n|} \cdot \overbrace{(n-1)\tilde{F}}^{\approx |F^n_{db}|}$$

  - *time grows* ***linearly*** *w.r.t.* $|F_{db}|$

Insight

- *General scheme of (sequential) k-means feature-based coarse alignment:*





- *K-means matching time:*

		Intra-cluster matching time		Inter-cluster matching time		Clusters update time
$T_{F^n \rightarrow F_{db}}$	$\approx$	$\alpha \cdot \tilde{f} \cdot C_{size}$	$+$	$\beta \cdot \tilde{f} \cdot k^n$	$+$	t_c^n
		constant w.r.t. n		linear growth w.r.t. k^n		grows w.r.t. n

Key features

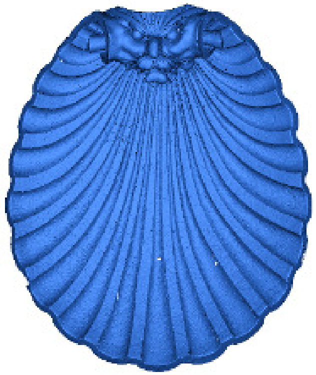
- *iat-k-means:*
  - ***I**ncremental: dynamically add features in F_{db} (online processing)*
  - ***A**daptive: dynamically increase the number of clusters when 'needed'*
  - *early **T**ermination: interrupt the cluster update before convergence*

- *For the specific class of **circular** features, we can obtain $\beta \ll \alpha$:*
  - *cost of feat-centroid matching $\ll$ feat-feat matching*

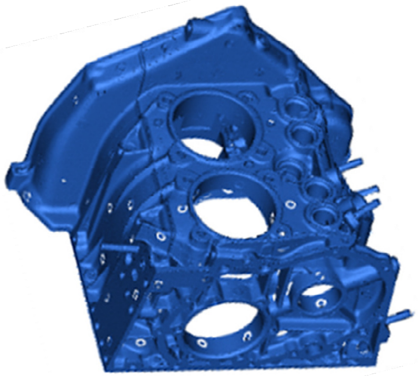
- *Performance comparison demonstrates iat-k-means effectiveness:*
  - *computation time greatly reduced and nearly constant w.r.t. n*
  - *no performance loss w.r.t. exhaustive search*

Datasets

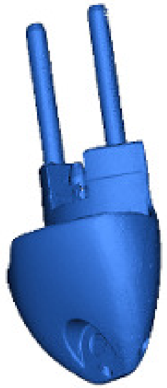
Cupid
45 RIs, 13M pts



Shell
98 RIs, 74M pts

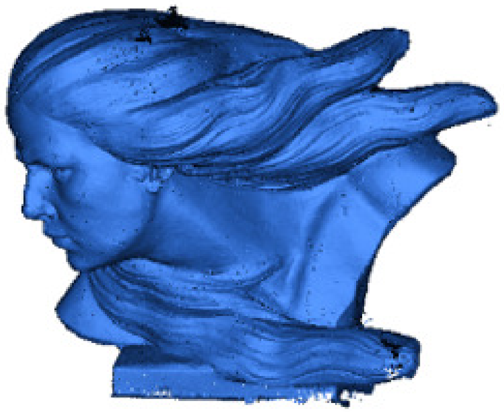


Carter
70 RIs, 38M pts



Slide
48 RIs, 7M pts

Neptune
79 RIs, 54M pts



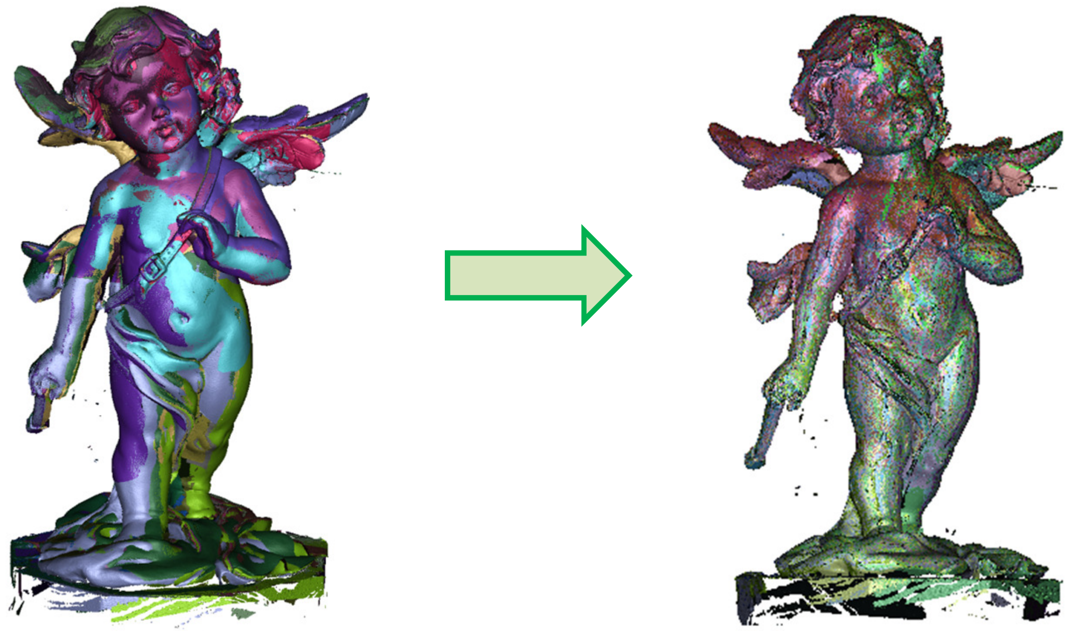
Hurricane
32 RIs, 22M pts

Capital
36 RIs, 20M pts

Venus
61 RIs, 51M pts

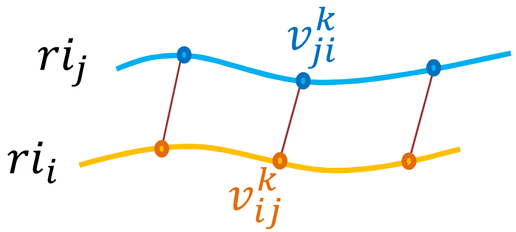
Global alignment

- *Once coarse alignment has been reached (and further refined through an Iterative Closest Point), a global alignment is required to distribute and further reduce the residual error*



- ***Goal:*** *develop an alignment approach capable to handle an entire set of views simultaneously, thus reducing a global error figure*

Insight



$$g(\mathcal{R},\mathcal{T}) = \min_{(\mathcal{R},\mathcal{T})} \sum_{i=1}^{N} \sum_{j=i+1}^{N} \sum_{k=1}^{n_{ij}} \left\| \left(R_i v_{ij}^k + t_i\right) - \left(R_j v_{ji}^k + t_j\right) \right\|^2$$

$$e_i \coloneqq i_{th}\ column\ of\ I_N$$

$$a_{ij}^k = (e_i \otimes I_3) v_{ij}^k - (e_j \otimes I_3) v_{ji}^k$$

$$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix} = \sum_{i=1}^{N} \sum_{j=i+1}^{N} \sum_{k=1}^{n_{ij}} \begin{bmatrix} a_{ij}^k \\ e_{ij} \end{bmatrix} \begin{bmatrix} a_{ij}^k & e_{ij} \end{bmatrix}$$

$$g(\mathcal{R},\mathcal{T}) = tr\left(\begin{bmatrix} \mathcal{R} & \mathcal{T} \end{bmatrix} \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \begin{bmatrix} \mathcal{R}^T \\ \mathcal{T}^T \end{bmatrix} \right)$$

$$\mathcal{T}(\mathcal{R}) = -\mathcal{R} B C^{\dagger}$$

$$M \coloneqq A - B C^{\dagger} B^T$$

$$Schur\ complement\ of \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$$

$$g(\mathcal{R}) = tr\left(\mathcal{R} M \mathcal{R}^T\right)$$

1. Algebra

Insight

$$g(\mathcal{R}): \quad SO_3^N \longrightarrow \mathbb{R}$$
$$g(\mathcal{R}) = tr(\mathcal{R}M\mathcal{R}^T)$$

$$g \circ \varphi(\boldsymbol{\omega}): \mathbb{R}^{3N} \longrightarrow \mathbb{R}$$
$$g \circ \varphi(\boldsymbol{\omega}) = tr\left(\mathcal{R}e^{\widetilde{\Omega}(\boldsymbol{\omega})} M {e^{\widetilde{\Omega}(\boldsymbol{\omega})}}^T \mathcal{R}^T\right)$$

$\forall\, \mathcal{R} \in SO_3^N\ \exists$ *smooth exponential map*

$$\varphi(\boldsymbol{\omega}): \qquad \mathbb{R}^{3N} \qquad \longrightarrow \quad SO_3^N$$
$$\boldsymbol{\omega} = \begin{bmatrix} \omega_x^1 & \omega_y^1 & \omega_z^1 \\ \omega_x^2 & \omega_y^2 & \omega_z^2 \\ & \vdots & \\ \omega_x^N & \omega_y^N & \omega_z^N \end{bmatrix} \longmapsto \mathcal{R}e^{\widetilde{\Omega}(\boldsymbol{\omega})}$$

2. Local parametrization

Second order Taylor approximation of $g \circ \varphi$ *for* $\overline{\boldsymbol{\omega}} = 0 \in \mathbb{R}^{3N}$ *yields:*

$$tr(\mathcal{R}M\mathcal{R}^T) + 2\boldsymbol{\omega}^T \nabla_{g\circ\varphi}(0) + \boldsymbol{\omega}^T\left(\widehat{H}_{g\circ\varphi}(0) + \widetilde{H}_{g\circ\varphi}(0)\right)\boldsymbol{\omega}$$

Descent direction for the error function determined through Gauss-Newton:

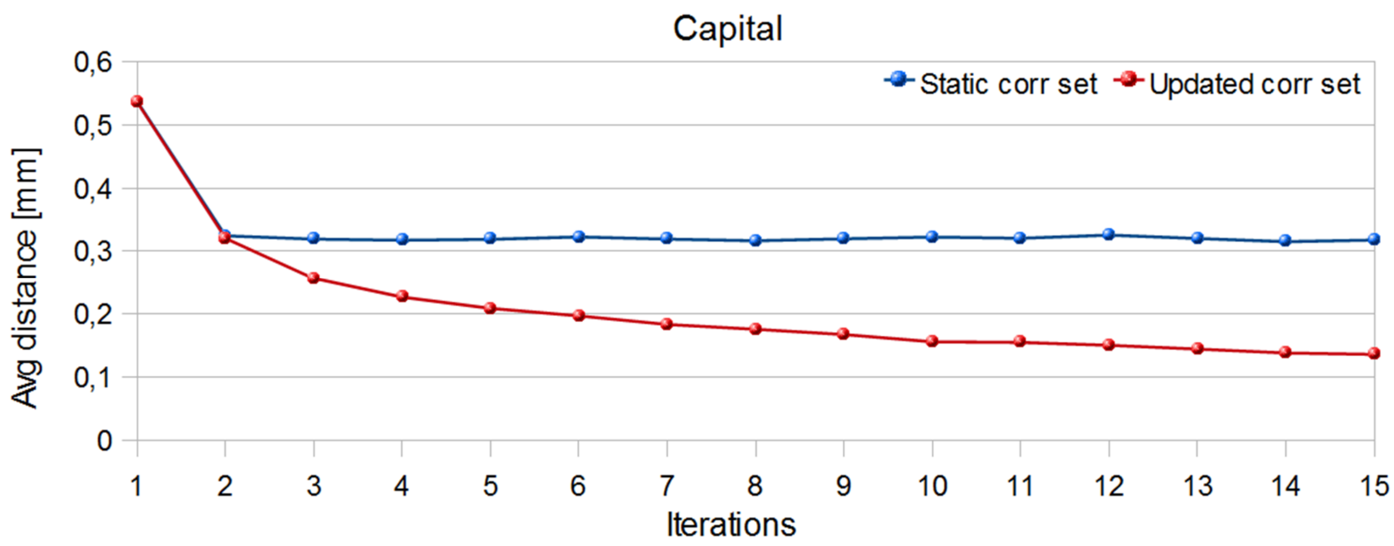
$$\boldsymbol{\omega}_{opt} = \begin{cases} -\left[\widehat{H} + \widetilde{H}\right]^{-1}\nabla & if\ \ \widehat{H} + \widetilde{H} > 0 \\ -\left[\widehat{H}\right]^{-1}\nabla & otherwise \end{cases}$$

$$\mathcal{R}^{i+1} = \mathcal{R}^i e^{\widetilde{\Omega}(\lambda_{opt}\boldsymbol{\omega}_{opt})}$$

3. Gauss-Newton Optimization

Key features

- *Correspondence set update:*
  - *OOM employed static correspondence selection*
  - *IOM updates the correspondence set at every iteration*
    - *which in turn pose serious computational issues…*
    - *… but grants a remakable performance boost*



Key features

- *Point-to-plane metric adoption grants a slight precision improvement*
- *Modified descent direction computation:*

$$\boldsymbol{\omega}_{opt} = \begin{cases} -\left[\widehat{H} + \widetilde{H}\right]^{-1}\nabla & if \quad \widehat{H} + \widetilde{H} > 0 \\ -\left[\widehat{H}\right]^{-1}\nabla & otherwise \end{cases}$$

// Line search with Newton direction

// Damped Gauss-Newton

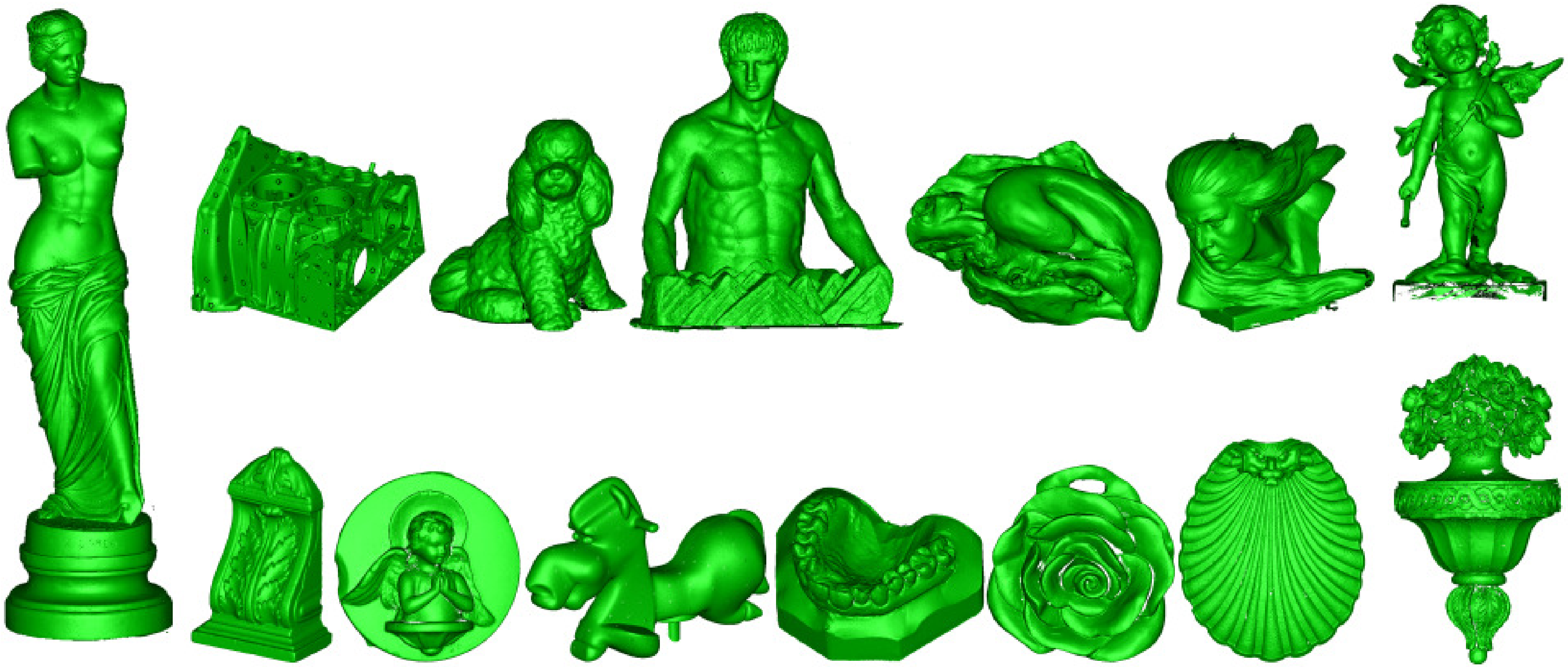
where:

$$\underbrace{\widehat{H}}_{} = \underbrace{J^T}_{[3N \times 9N]} \underbrace{(I_3 \otimes M)}_{[9N \times 9N]} \underbrace{J}_{[9N \times 3N]}$$

$$\underbrace{\widetilde{H}}_{[3N \times 3N]} = -\underbrace{\tilde{Q}^T}_{[3N \times 9N^2]} \underbrace{(I_{3N} \otimes M \boldsymbol{\mathcal{R}}^T \boldsymbol{\mathcal{R}})}_{[9N^2 \times 9N^2]} \underbrace{\tilde{Q}}_{[9N^2 \times 3N]}$$

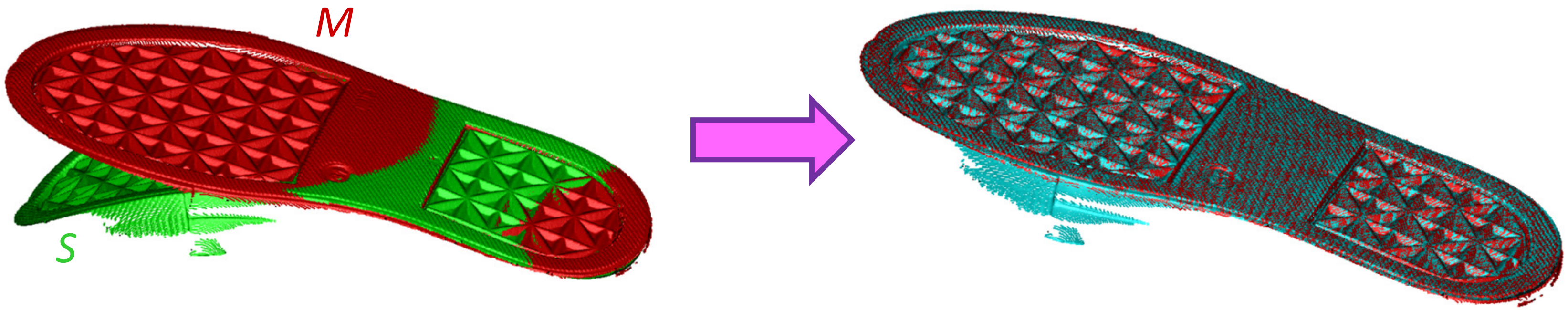
  - *Avoiding* $\widetilde{H}$ *computation greatly reduces the computational burden:* $\widehat{H} + \widetilde{H}$ *is never semi-positive definite*

Datasets



Deformable alignment

- *Ideally the residual alignment error between two scans should be zero, however this cannot be reached due to:*
  - *Imperfect calibration, uncooperative surfaces, limited depth resolution*
  - *Moving / deforming object*



- ***Goal:** given two 3D surfaces, a scan S and a model M, deform S on M applying the 'least' possible deformation*

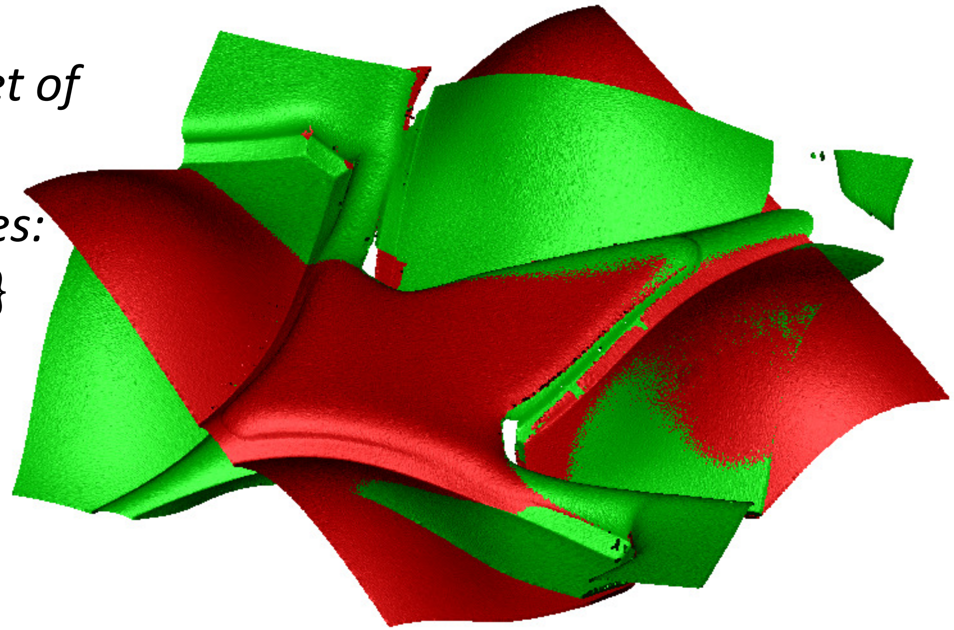
Insight

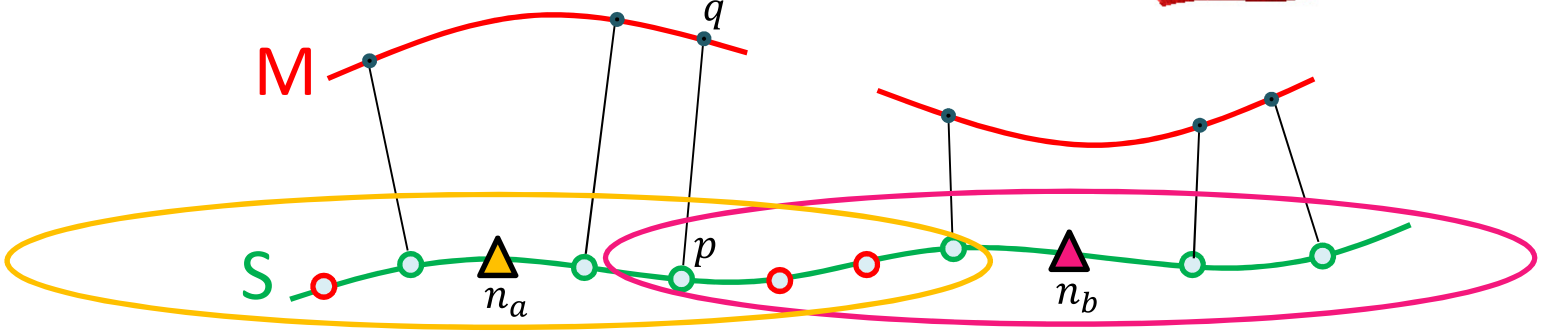
Core idea of the approach:

- *solve the problem globally for a subset of nodes N*
- *drive deformation via correspondences:*

$$C = \{(P, Q) \quad | \quad P \in S, Q \in M\}$$

- *extend the deformation from N to S*



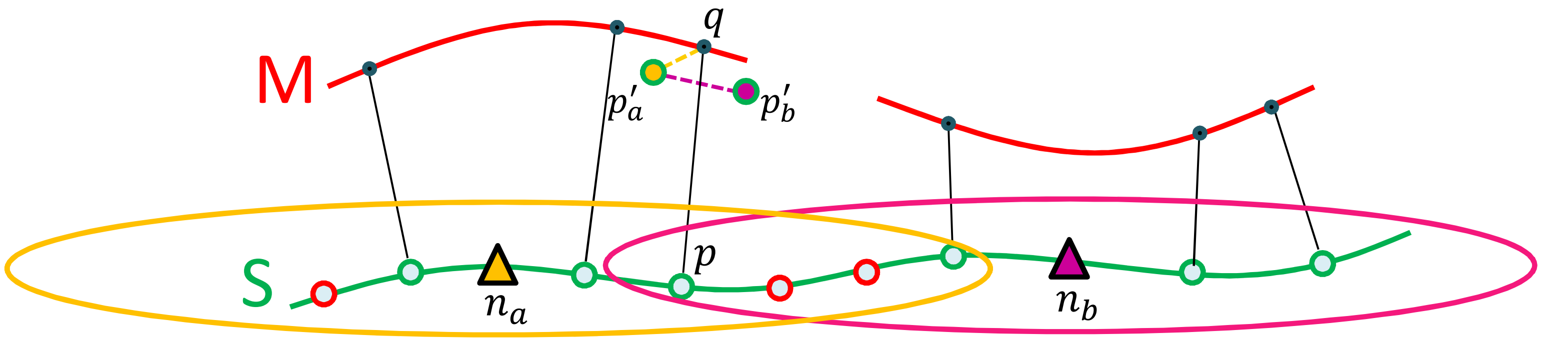


Key features

- *Energy to minimize composed by a **fitting** and a **smoothing** term:*

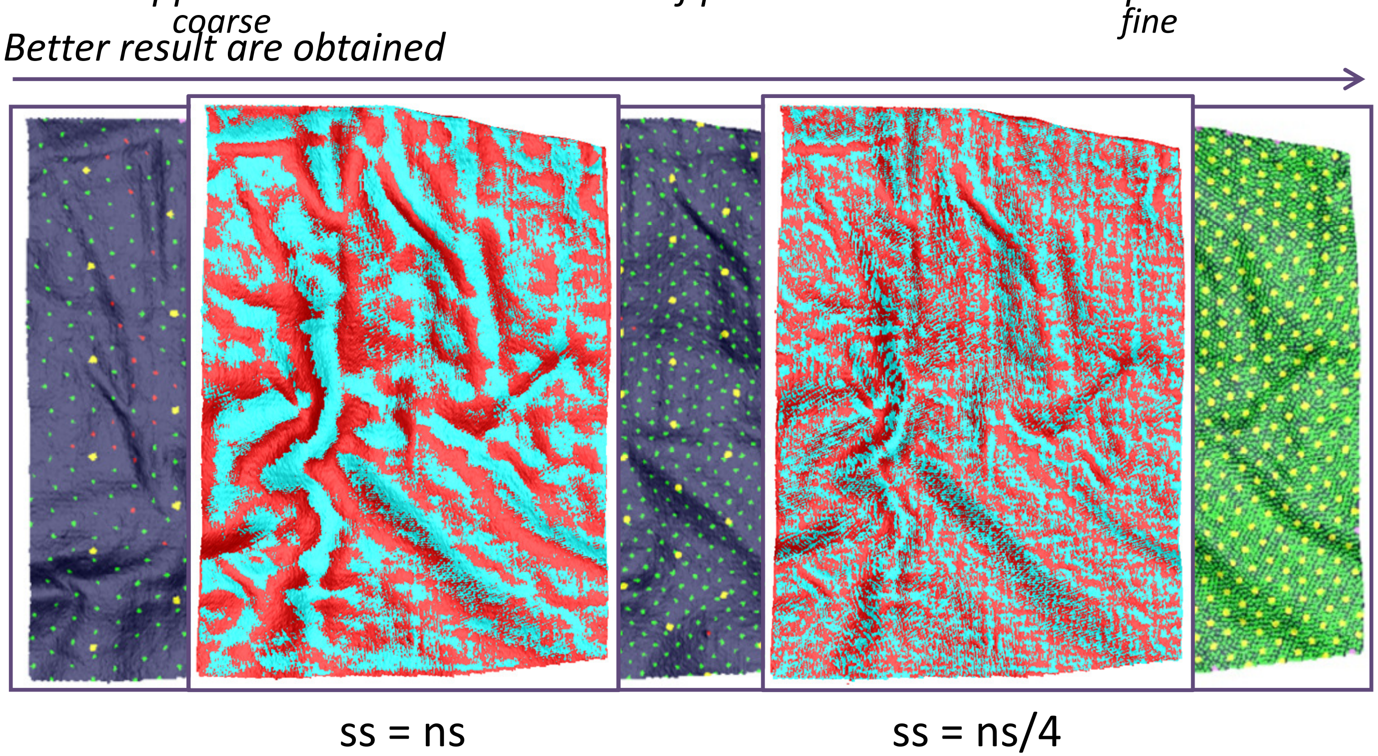
$$E = w_f \cdot \sum_{n=1}^{|fN|} \sum_{c=1}^{|vC_n|} \frac{\|(R_n p_c + t_n) - q_c\|^2}{|fN| \cdot |vC_n|} + w_s \cdot \sum_{n_a=1}^{|N|} \sum_{n_b=1}^{|N_{n_a}|} \sum_{p=1}^{|P_{n_a} \cap P_{n_b}|} \frac{\left\|(R_{n_a} p + t_{n_a}) - (R_{n_b} p + t_{n_b})\right\|^2}{|N| \cdot |N_{n_a}| \cdot |P_{n_a} \cap P_{n_b}|}$$

Fitting *Smoothing*



Key features

- *Hierarchical node spacing:*
  - *Node spacing decreases during the optimization (coarse-to-fine refinement)*
- *Uneven grid / sample spacing:*
  - *Similar approaches use the same sets of points as nodes and samples*
  - *Better result are obtained*



Key features

- To extend the deformation from the node set N to the scan S, for each $s \in S$ a rototranslation matrix is computed through dual quaternion skinning

- All rototranslation matrices are applied in a single step

Iter = 10

Conclusions

- *A pipeline to reconstruct sets of dense 3D scans has been introduced:*
  - *A **coarse alignment** approach that demonstrated to be:*
    - ***robust** with respect to a variety of input data (dental, cultural heritage, …)*
    - *capable to align scans that possess a **restricted overlap area***
    - *incrementally aligns **big datasets** without performance degradation*
  - *A **global alignment** technique that guarantees:*
    - *to compute a solution considering **all scans simultaneously***
    - ***precise alignment** even for **big datasets** (tested up to 170 scans, 127 M pts)*
    - *computational performance can be **improved through sparse matrices***

- *An algorithm to perform a **deformable alignment** between pairs of scans:*
  - ***no assumption on the model***
  - ***twice as fast as current state of the art***
  - ***robust**, as long as moderate deformation*

Thank you!

Conference papers:

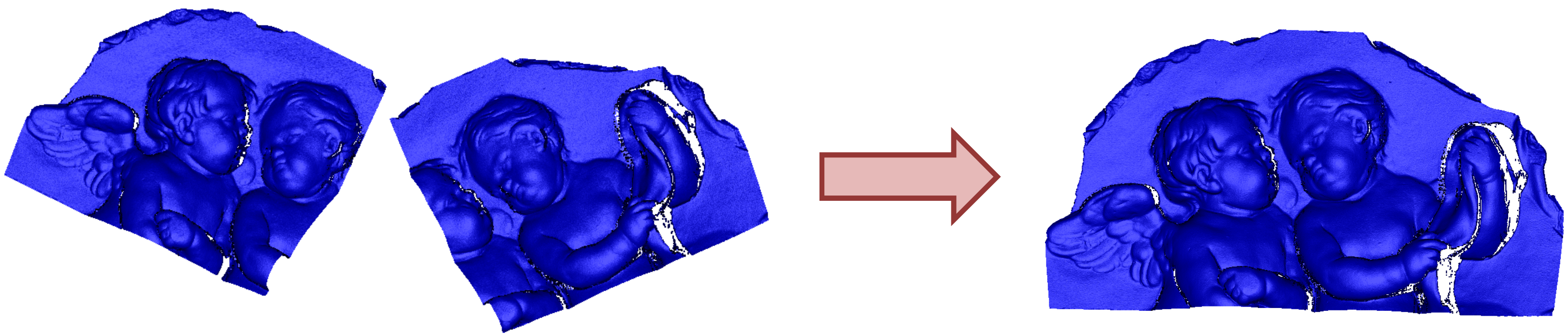
- [BSLC10] F. Bonarrigo, A. Signoroni, R. Leonardi, and M. Carocci - ***A multiscale feature extraction approach for 3D range images*** - WIAMIS 2010.
- [BS11] F. Bonarrigo and A. Signoroni - ***An enhanced Optimization-on-a-Manifold framework for global registration of 3D range data*** - 3DIMPVT conference, 2011.
- [BSL11] F. Bonarrigo, A. Signoroni, and R. Leonardi - ***A robust pipeline for rapid feature-based pre-alignment of dense range scans*** - ICCV 2011.
- [PBS12] N. Pezzotti, F. Bonarrigo, and A. Signoroni - ***Boosting the computational performance of feature-based multiple 3D scan alignment by iat-k-means clustering*** - 3DIMPVT 2012.
- [BBS12] F. Bonarrigo, M. Botsch and A. Signoroni - ***As-rigid-as-possible registration of deformable surfaces*** - (*submission pending)*

Journal publications:

- [BSL12] F. Bonarrigo, A. Signoroni, and R. Leonardi - ***Multi-view alignment with database of features for an improved usage of high-end 3D scanners*** - EURASIP 2012.
- [BS13] F. Bonarrigo and A. Signoroni - ***Efficient Global registration of large collections of range images with an improved Optimization-on-a-Manifold approach*** - (*submission pending)*

Pairwise coarse alignment

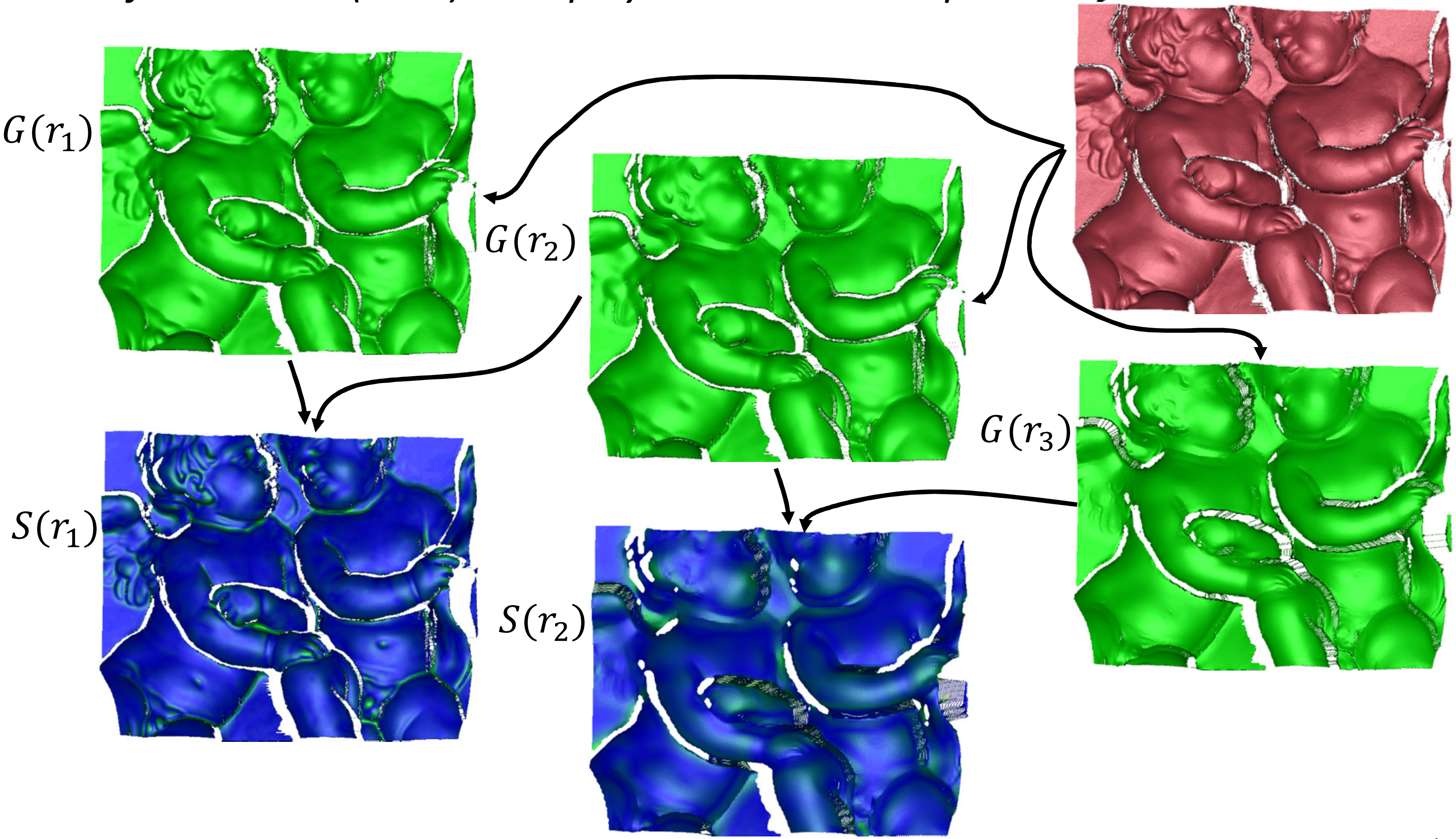
- *Once each scan has been acquired, it needs to be correctly aligned into a global reference system in order to reconstruct the final 3D model*



- ***Goal:*** *develop an automatic pairwise coarse alignment technique which:*
  - *make no assumption on the initial position*
  - *assume a portion of overlap between the scans*

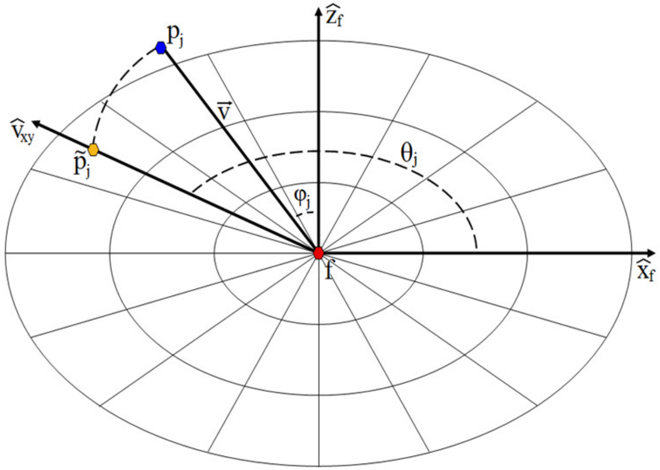
Insight

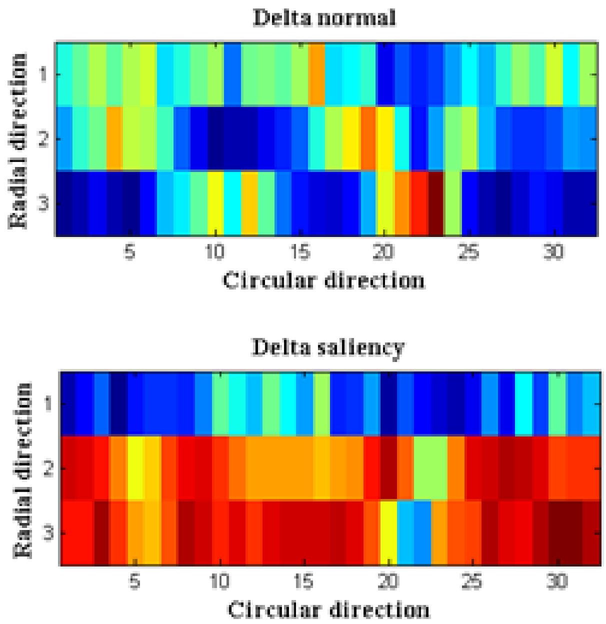
- *Feature extraction: a 3D extension of the multi-scale analysis Difference-Of-Gaussians (DOG) is employed to determine possible feature locations*



Insight

- *Feature description: a 2D circular grid (M radial and L angular sectors) is used to describe a feature neighbourhood (saliency and normal)*

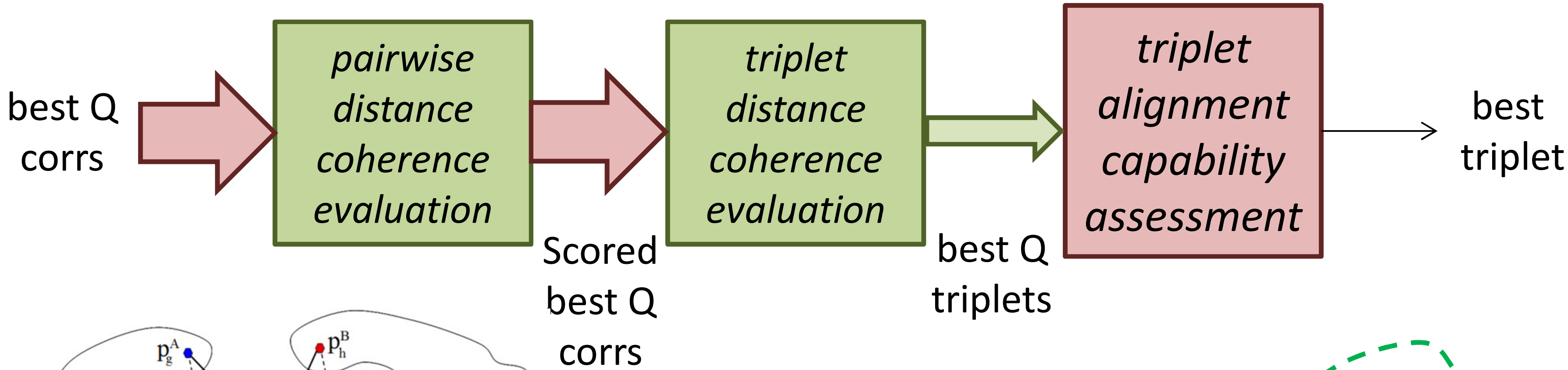


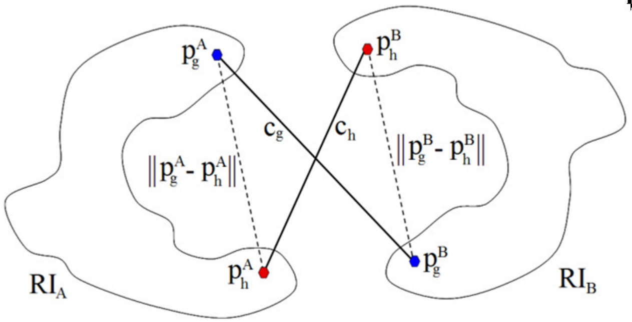


- *Feature matching: each feature of a scan is matched against all those detected from the other scan at a given scale; a score invariant w.r.t. orientation is computed for each correspondence, and the best ones are retained*

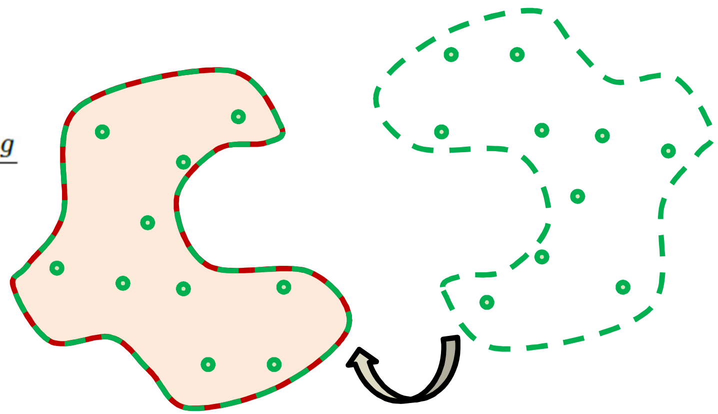
Insight

- *Alignment estimation: a 3-step procedure has been developed to further skim the correspondence set in order to retain only the correct matches*





$$t^{score} = 1 - \frac{d_{gh} + d_{hj} + d_{jg}}{3}$$



$$d_{gh} \equiv d(c_g, c_h) = \frac{\left|\, \|p_g^A - p_h^A\| - \|p_g^B - p_h^B\| \,\right|}{\max\left(\|p_g^A - p_h^A\|, \|p_g^B - p_h^B\|\right)}$$