RECONSTRUCTION VIA DETECTION
HIGHLY ACCURATE RECONSTRUCTION FROM UNORGANIZED 3D SCANS

Tolga Birdal

BRIEF RÉSUMÉ

Education

- **TUM**, Technical University of Munich
- **Sabanci University**, Istanbul

Experience

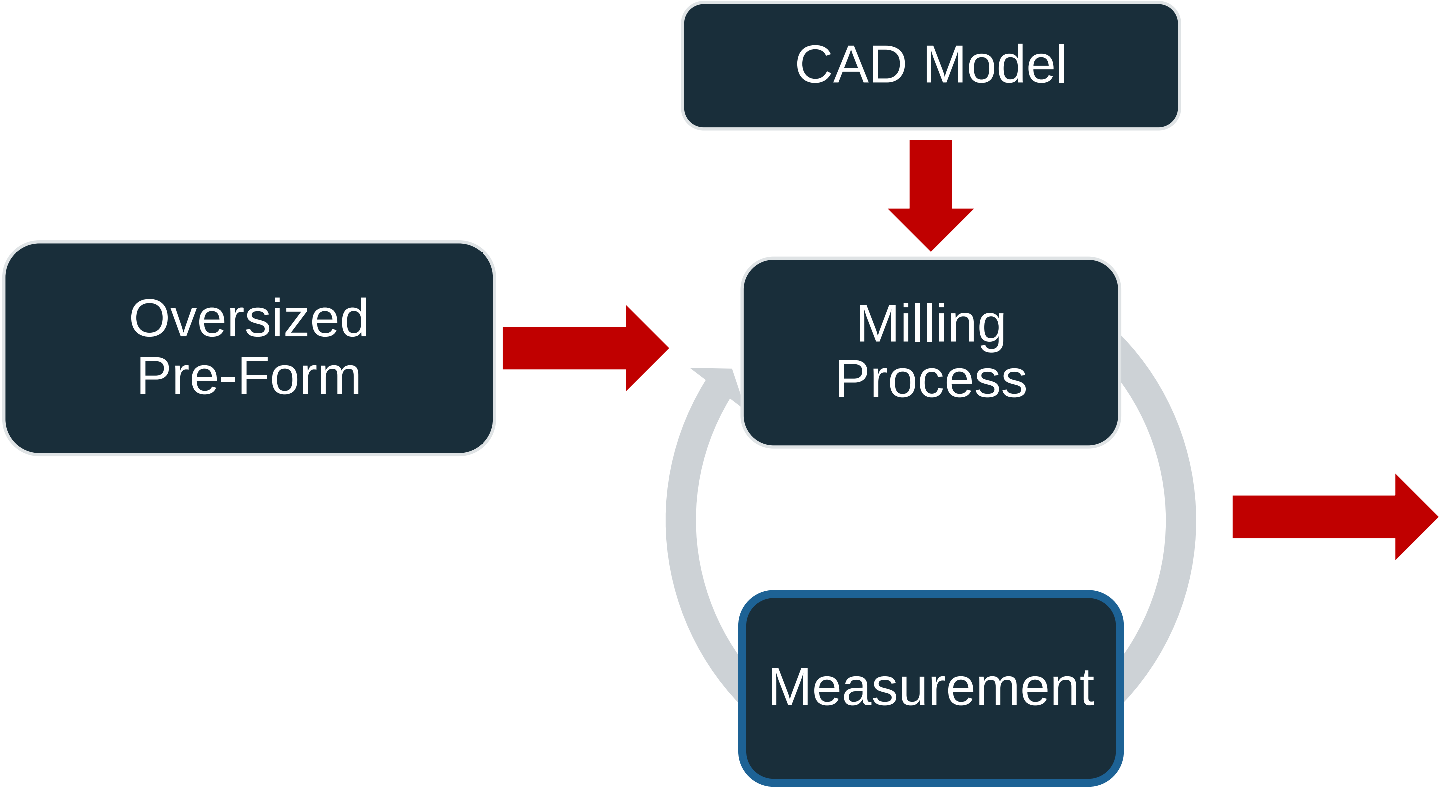
- **MERL**: Mitsubishi Electric Research Labs, Cambridge
- **Intel research**, Pittsburgh
- **CMU**: Carnegie Mellon University, Pittsburgh
- **Isra Vision**, Istanbul
- **OpenCV**, Google Summer of Code 2014
- **Siemens AG**, Munich, 2014 - Present

Entrepneurship

- **Befunky**: www.befunky.com
- **Gravi** Information Technologies

MOTIVATION

- Aid the accurate manufacturing of carved industrial parts.



CAD Model

Oversized Pre-Form

Milling Process

Measurement

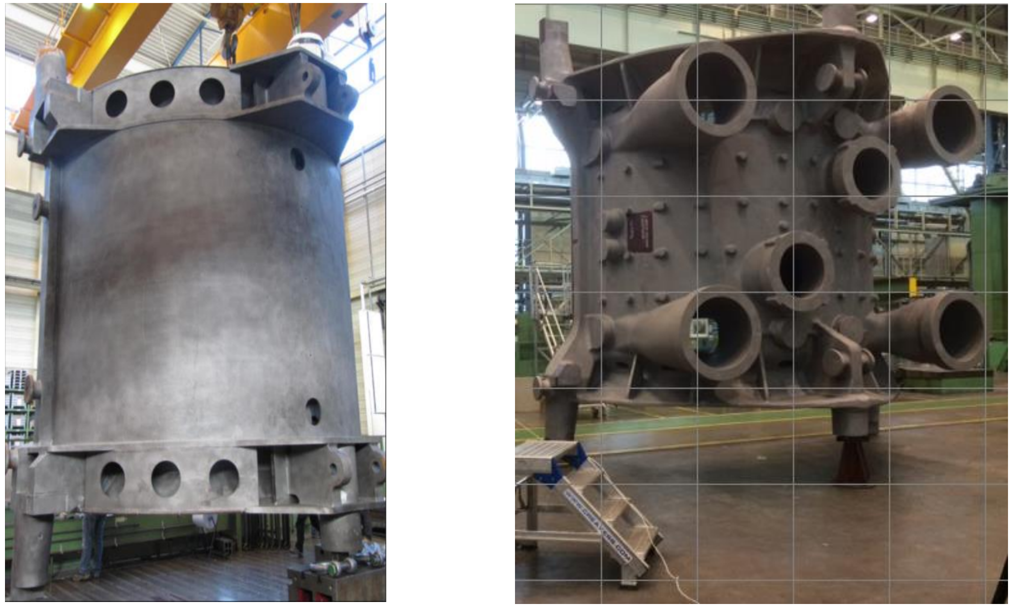
MOTIVATION

Digitization of industrial 3D objects:

- High precision
- Full automation
- Real-time performance
- User feedback
- Marker-less operation
- Segmentation

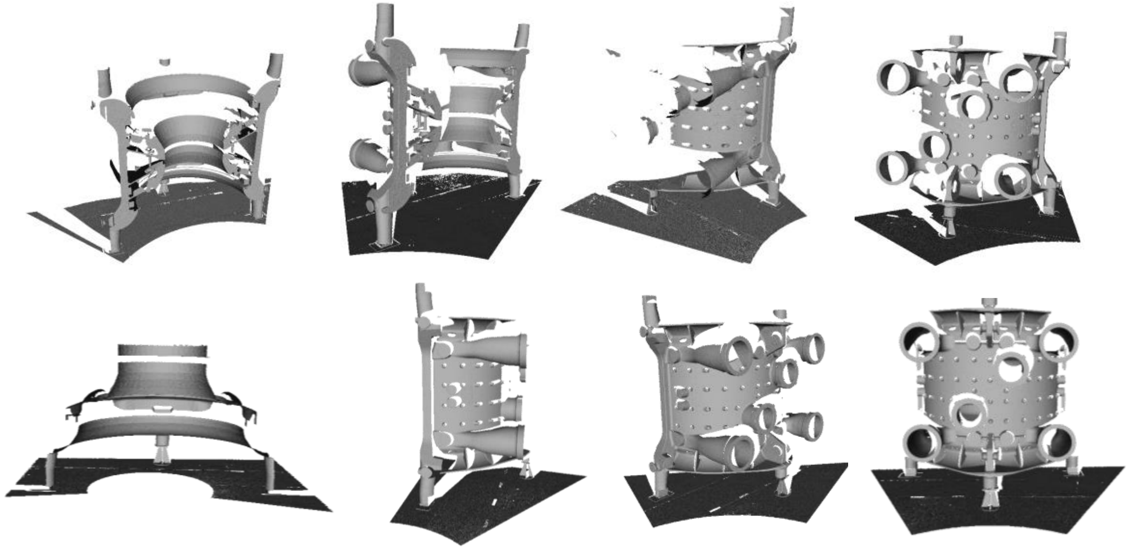
Challenges:

- Size (large volumes)
- Accuracy
- Specular materials
- Unordered sequence
- Clutter & occlusions
- Small number of scans



Images of Gas Turbine Casings

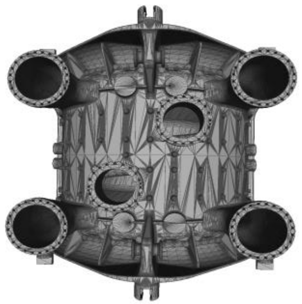
INTRODUCTION



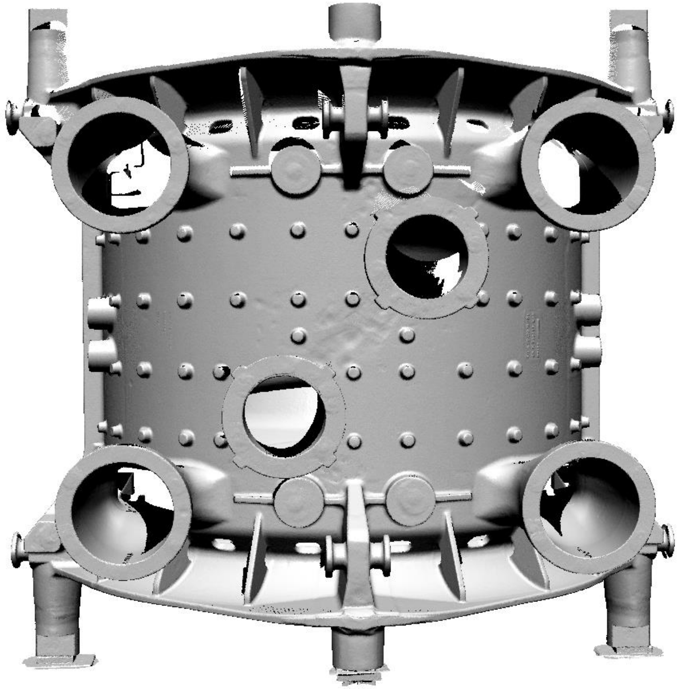
Given **3D Point Clouds**

- In Clutter
- Occlusions
- Unordered
- Large Scale

CAD Model

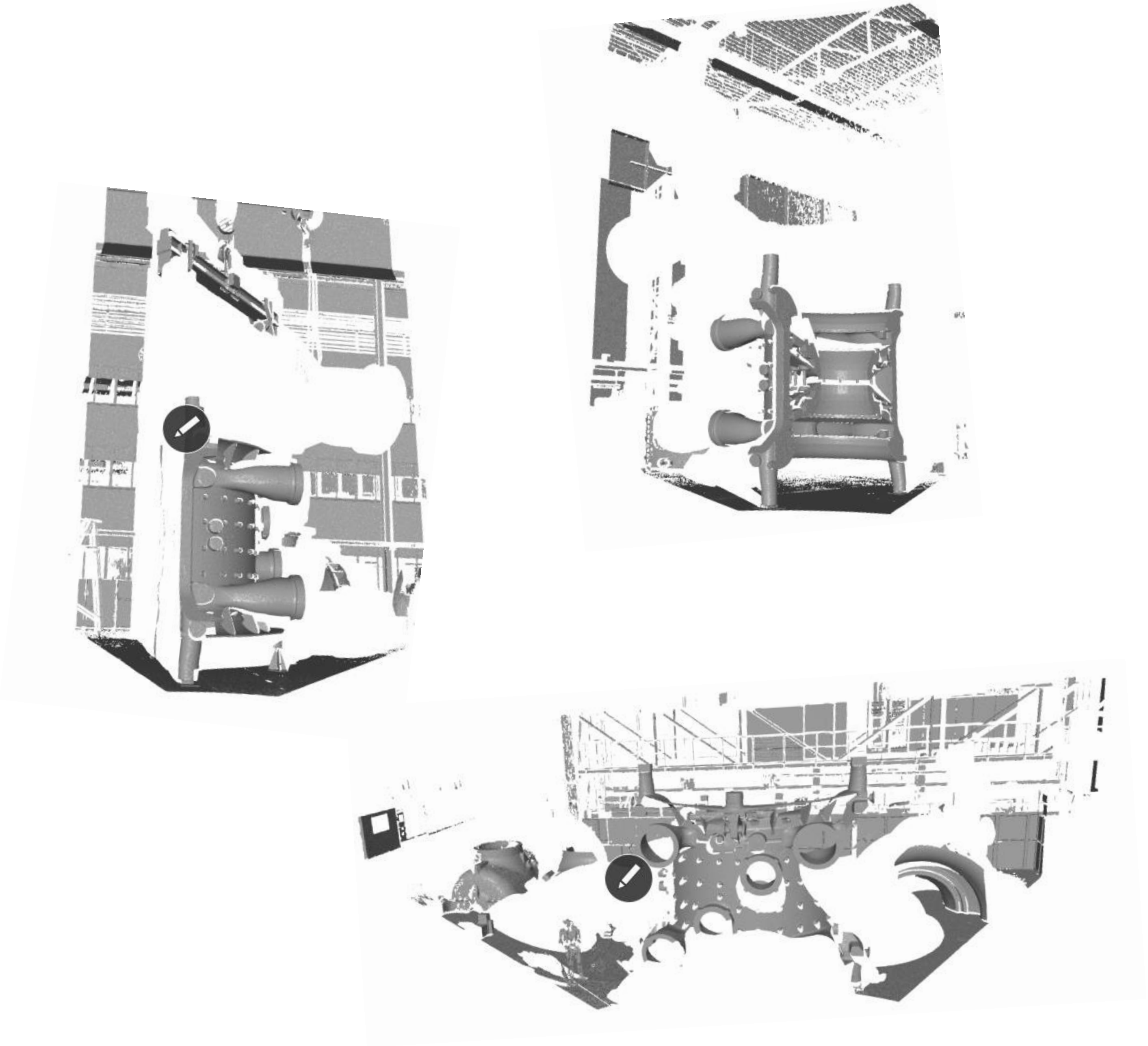


Accurately reconstruct 3D mesh.



DATA MODALITY

- Multiple, unordered scans of gas turbine casings
- Between 10M – 100M 3D data points
- Accuracy around 1-2mm in 10m working distance
- Objects of interest are contained in ~5^3m^3 3D volumes
- No depth or RGB images, only unstructured point clouds
- CAD models do not exactly match reality and are very different



PRIOR ART

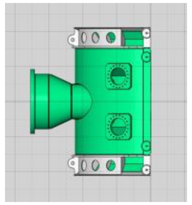
- Point Cloud Based
  - Relies on scan to scan registration
  - Cannot deal well with clutter or occlusions
  - Suffers from high complexity
  - Too slow for online operation and real-time feedback

- Volumetric Methods (KinFU-like)
  - Require sequential input data
  - Require depth images (due to SDF)
  - Resolution is limited to the voxel size
  - Suffers from drifts and tracking error

We alleviate these problems via the introduction of a proxy 3D CAD model.

OUR FORMULATION

CAD Model

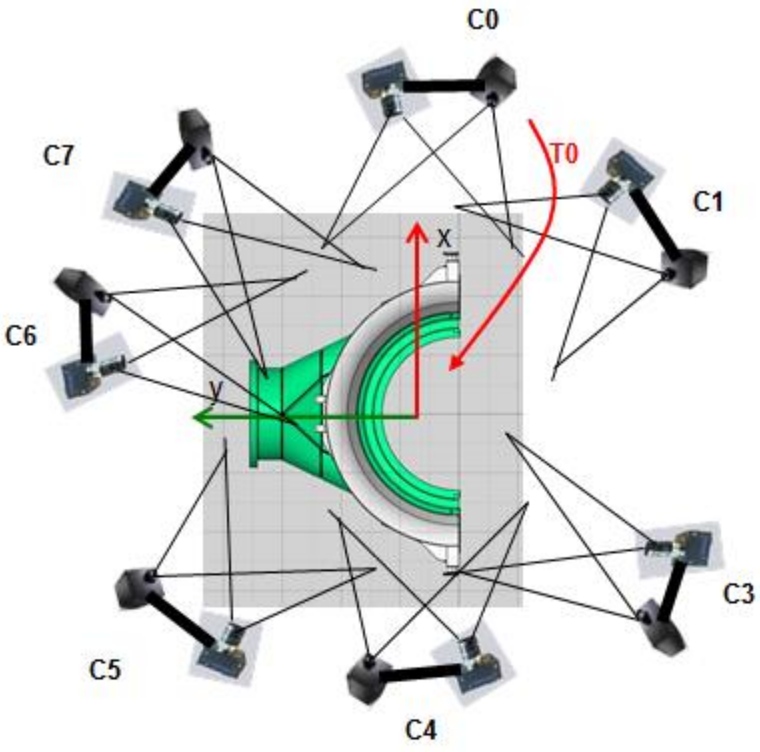


Given CAD model $\mathbf{M}$, find $\{T_i^M\}$ that best align the scenes $\{S_i\}$:

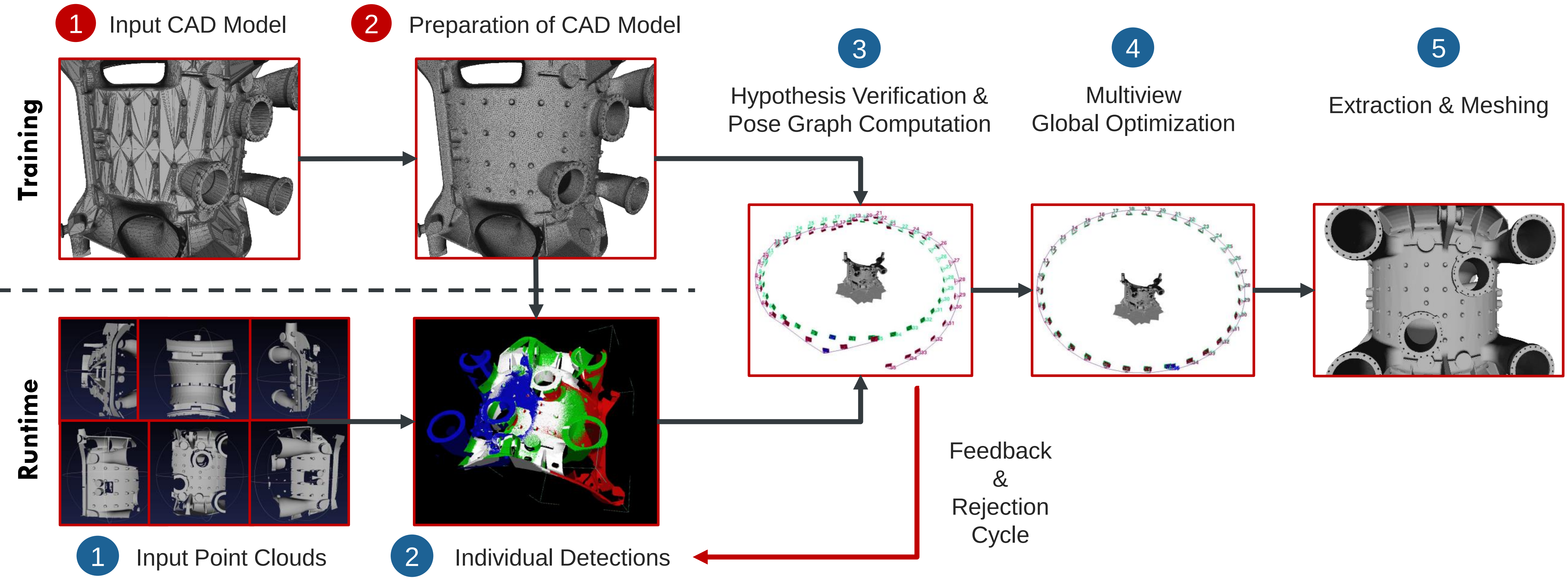
$$\mathbf{S}_G = \bigcup_{i=1}^{N} T_i^M(S_i|\mathbf{M})$$

$$T_i^M(x|M) = \mathbf{R}_i^M x + \mathbf{t}_i^M$$

T_i^M are w.r.t. model coordinate frame.



OUR PIPELINE



Training

1 Input CAD Model

2 Preparation of CAD Model

Runtime

1 Input Point Clouds

2 Individual Detections

3 Hypothesis Verification & Pose Graph Computation

4 Multiview Global Optimization

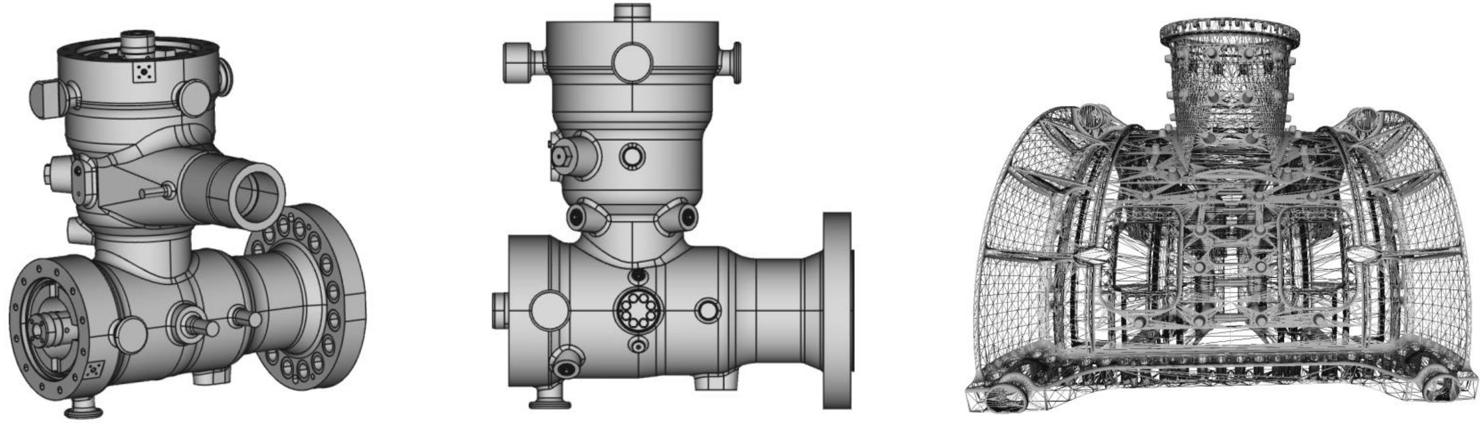
5 Extraction & Meshing

Feedback & Rejection Cycle

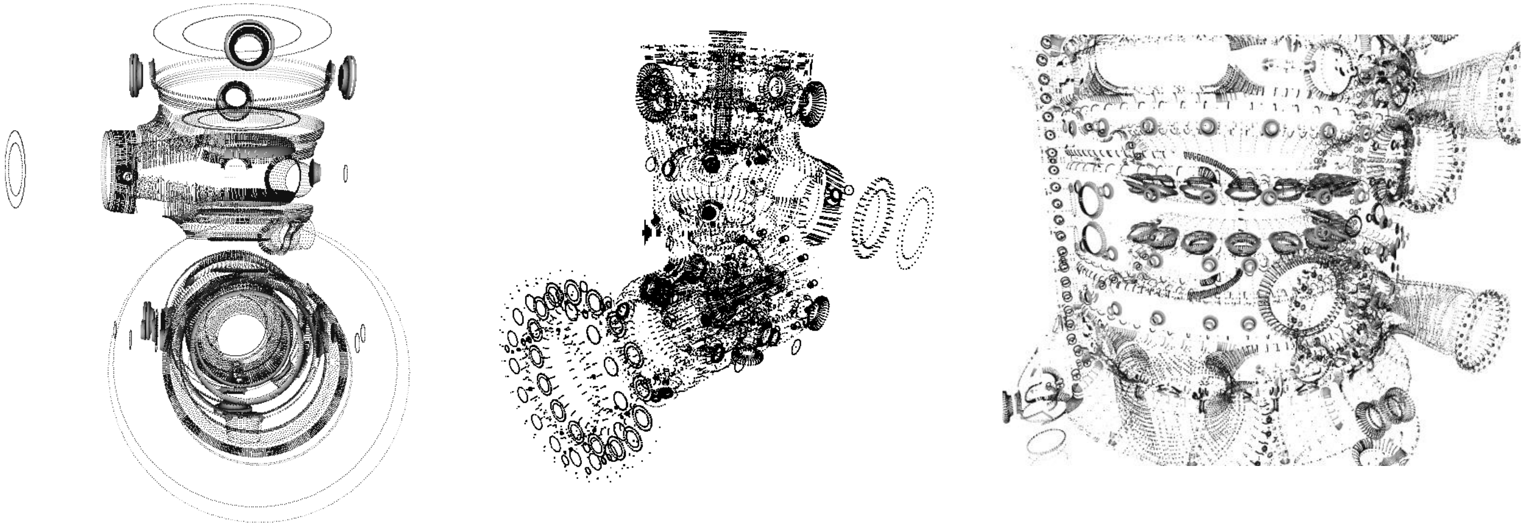
PREPARATION OF CAD MODEL

PROBLEMS WITH CAD MODELS IN INDUSTRY

Non-uniform triangle structure

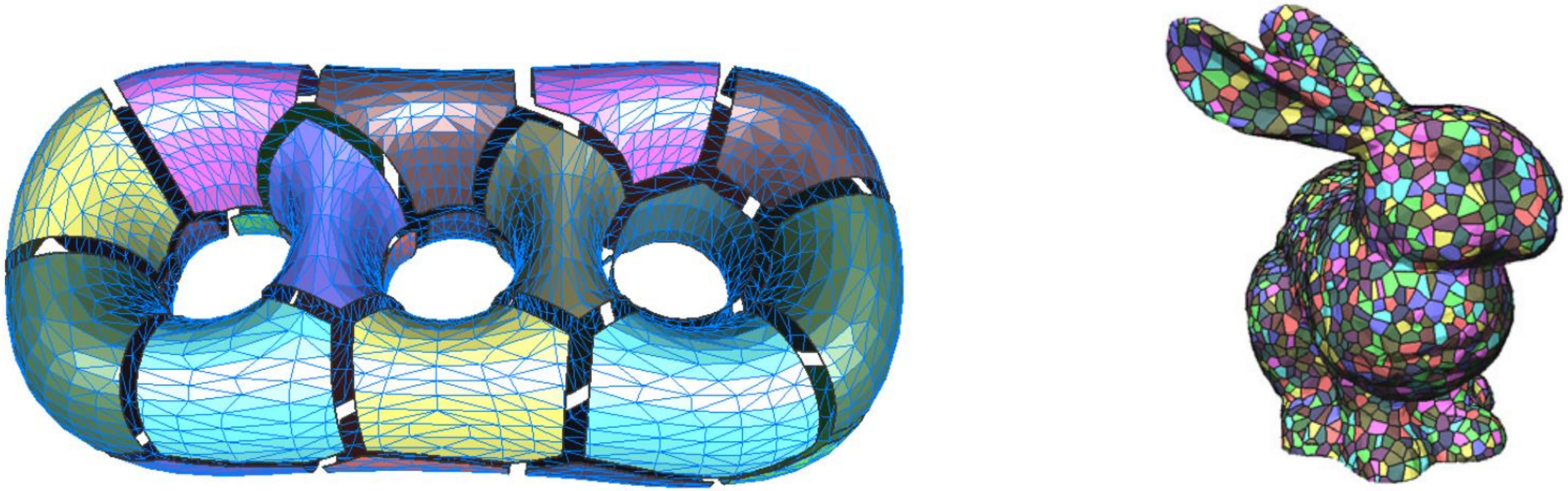


Uneven distribution of vertices

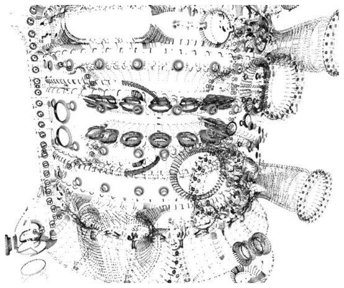
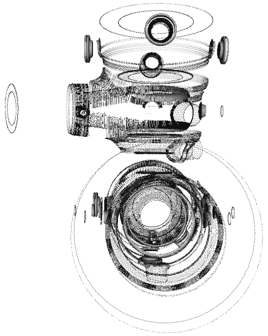
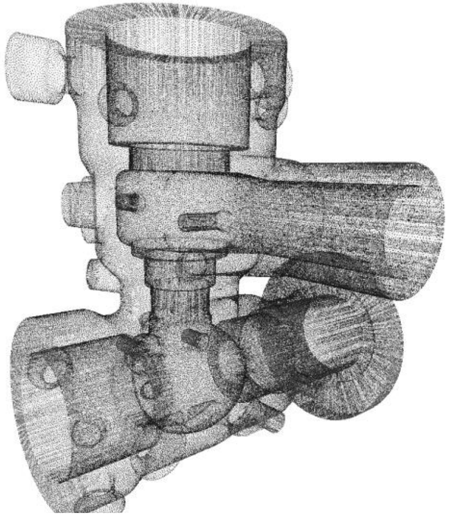
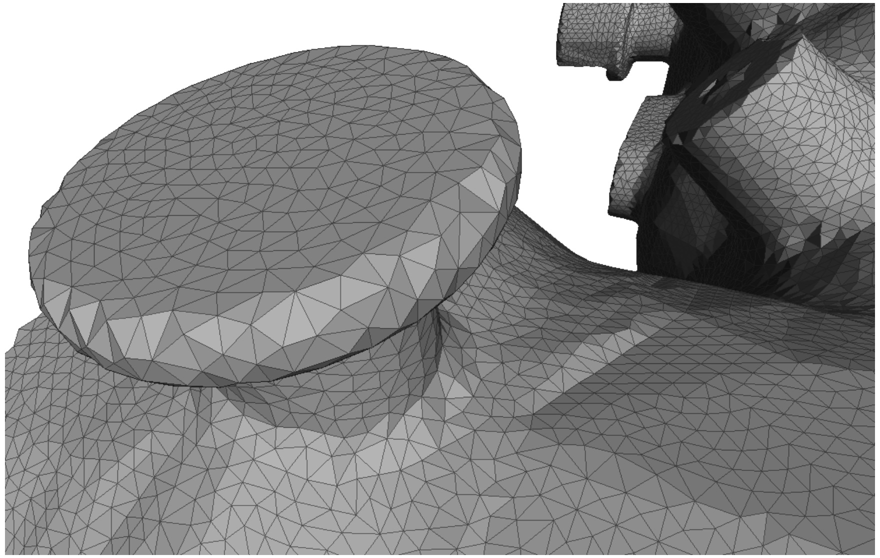
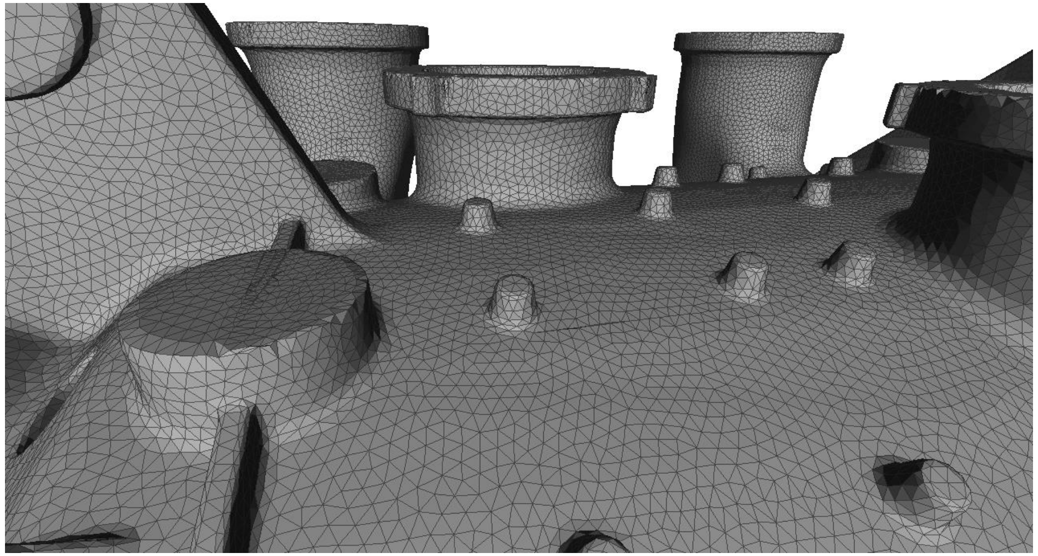
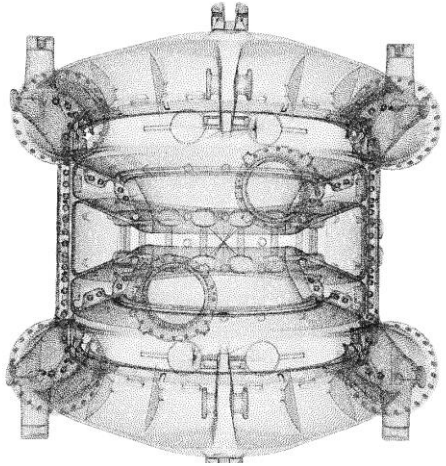
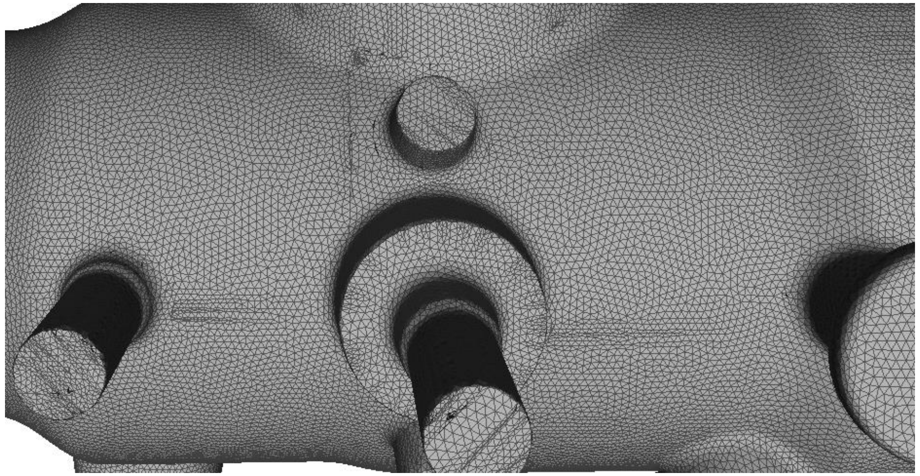
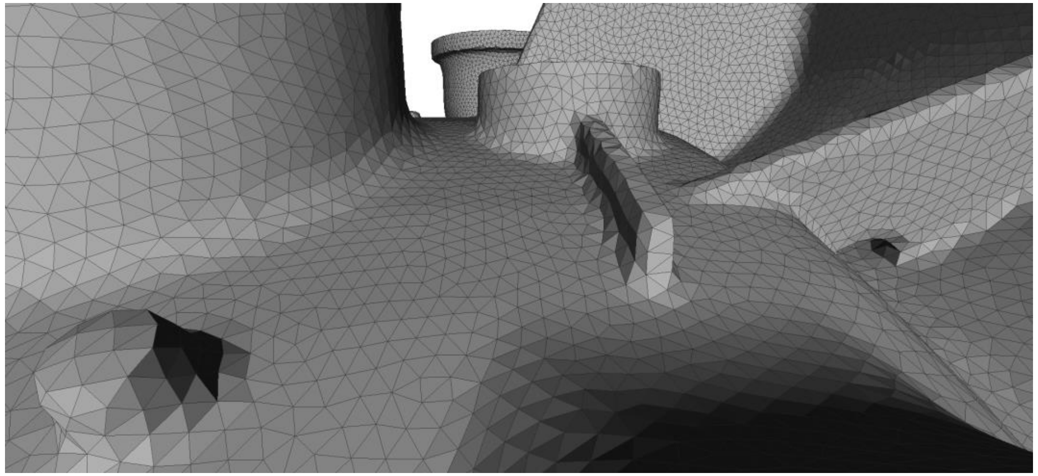


APPROACH : RE-MESHING

- ***Process of converting meshes to a more suitable discrete representation***
- Method of choice : *Restricted Voronoi Diagrams (RVD) (Yan et. al.)*
- With Centroidal RVD, we can achieve an isotropic re-meshing.



APPROACH : RE-MESHING

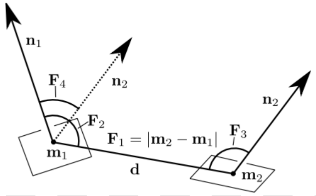









3D DETECTION OF CAD MODELS

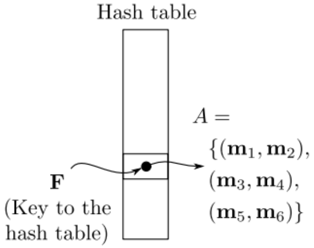
OUTLINE OF THE ALGORITHM

Load / Generate 3D Model

PPF Feature Extraction



Compute Hashtable & Precompute α_{model}

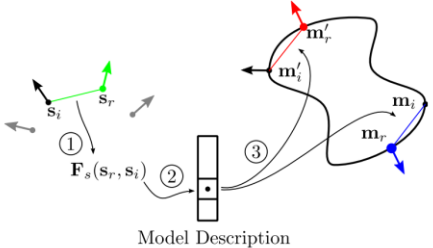


Trained Model

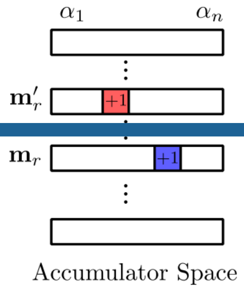
Training

Runtime

Load / Generate 3D Model

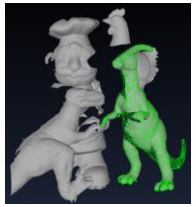


Hashtable Lookup & Vote for Pose



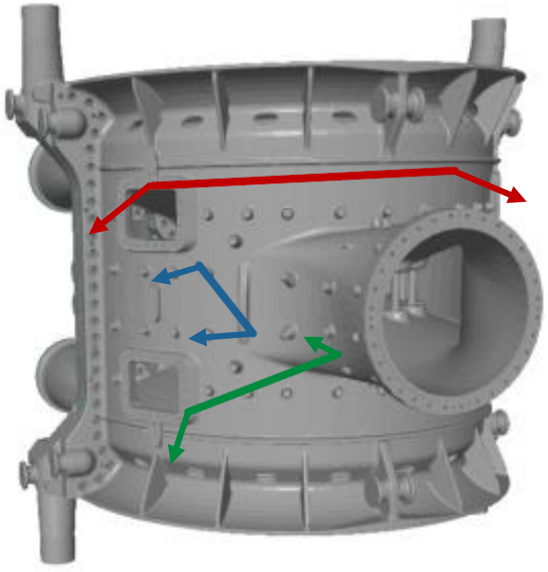
Pose Clustering

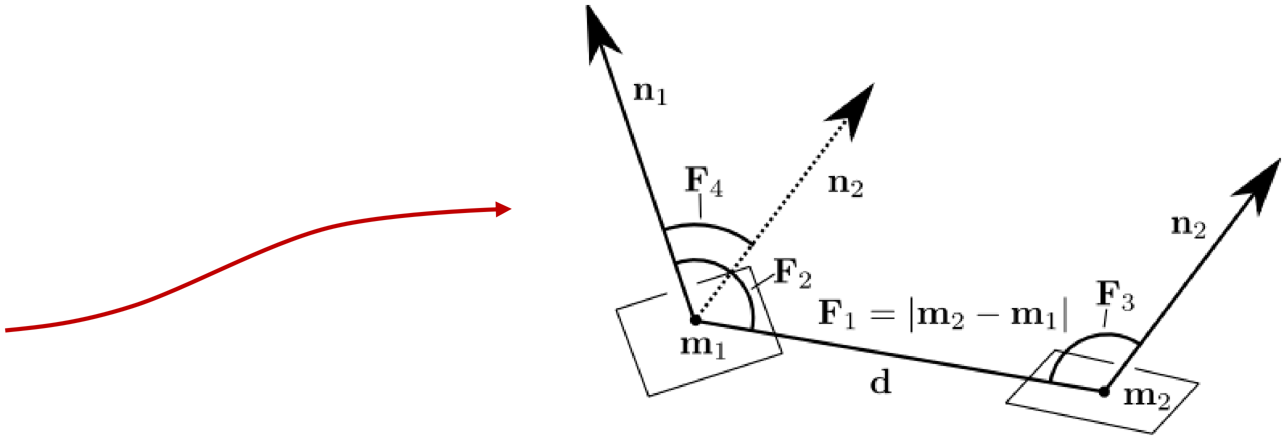
ICP Refinement & Final Pose



B. Drost, M. Ulrich, N. Navab, and S. Ilic, “Model globally, match locally: Efficient and robust 3d object recognition.” in *CVPR*. IEEE, 2010, pp. 998–1005. [Online]. Available: http://dblp.uni-trier.de/db/conf/cvpr/cvpr2010.html#DrostUNI10

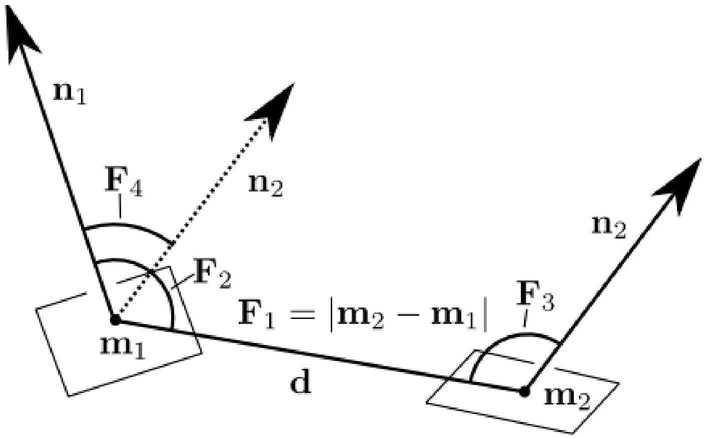
POINT PAIR FEATURES



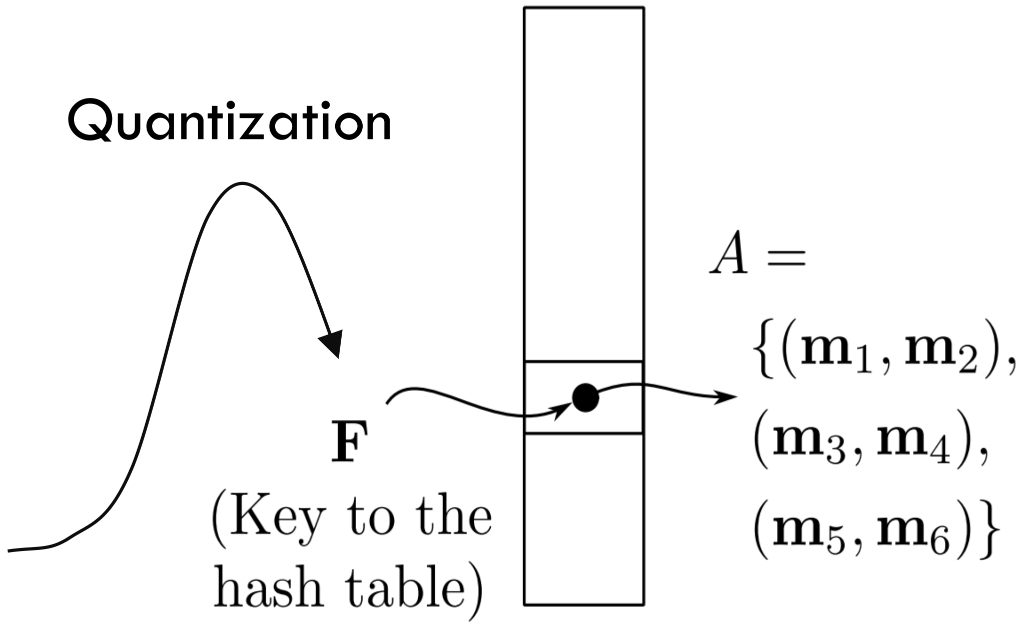


$$\mathbf{F}(\mathbf{m_1},\mathbf{m_2}) = (\|\mathbf{d}\|_2, \angle(\mathbf{n_1},\mathbf{d}), \angle(\mathbf{n_2},\mathbf{d}), \angle(\mathbf{n_1},\mathbf{n_2}))$$

INDEXING

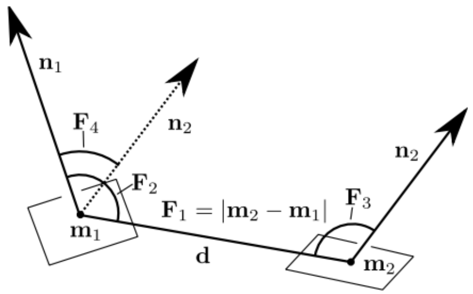


$$\mathbf{F}(\mathbf{m_1}, \mathbf{m_2}) = (\|\mathbf{d}\|_2, \angle(\mathbf{n_1}, \mathbf{d}), \angle(\mathbf{n_2}, \mathbf{d}), \angle(\mathbf{n_1}, \mathbf{n_2}))$$



- **F** is quantized and used as a key to Hashtable.
- Buckets store the reference points, and a rotation angle around the normal.
- We need to sample points to reduce complexity!

FEATURE QUANTIZATION

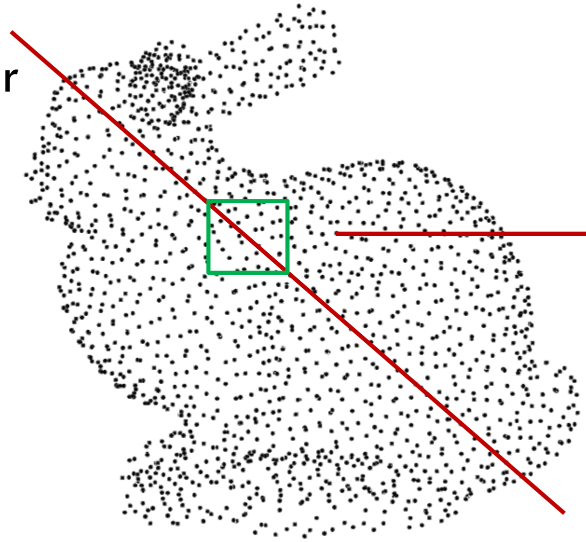


$$\mathbf{F}_q(\mathbf{m_1}, \mathbf{m_2}) = \left(\frac{\|\mathbf{d}\|_2}{d_q}, \frac{\angle(\mathbf{n_1}, \mathbf{d})}{a_q}, \frac{\angle(\mathbf{n_2}, \mathbf{d})}{a_q}, \frac{\angle(\mathbf{n_1}, \mathbf{n_2})}{a_q}\right)$$

Angle quantization

Points within similar radius
are mapped to same bins.

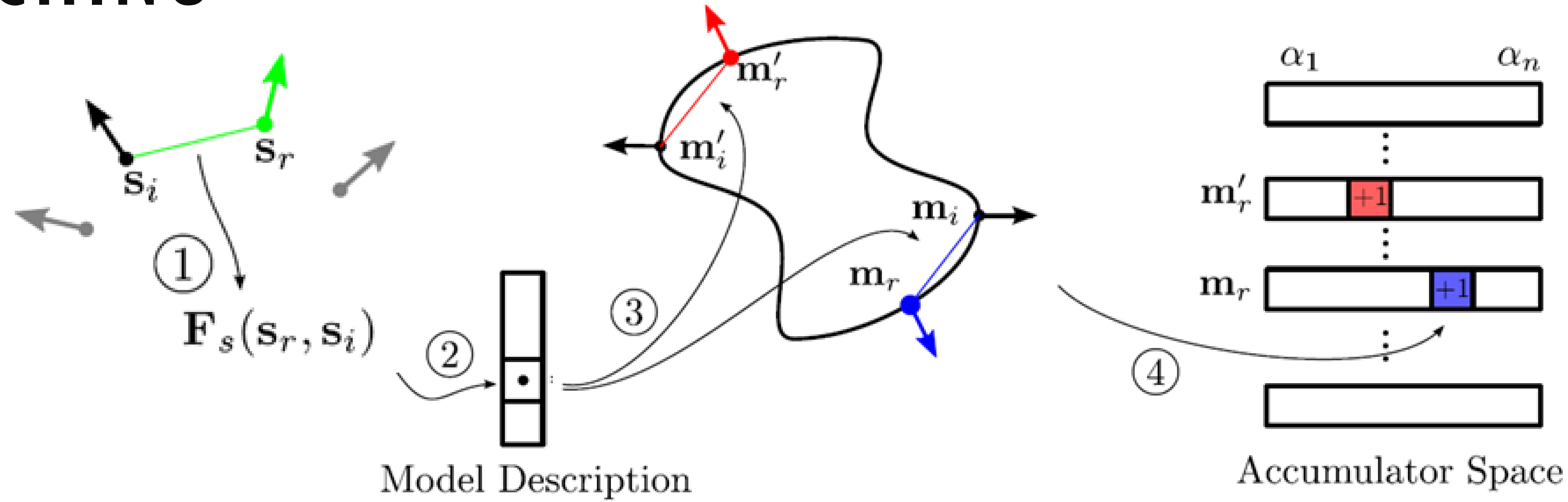
d: diameter



$$d_q = \tau d$$

Relative quantization coefficient

SEARCHING



- The reference point $\mathbf{s}_r$ is assumed to lie on the model and paired with every other point $\mathbf{s}_i$ in the scene
- The global model description is used to get all similar point pairs from the model
- Each point pair from the model votes for one sample of the local parameters

HYPOTHESIS VERIFICATION & RANKING

- Many candidates remain to be evaluated
- Simple Idea:
  - Retain all meaningful hypothesis
  - Register each one to the CAD model
  - Select the best
- Ranking:
  - A model based score function

How do we make ICP so fast that we could verify all candidates in real-time ?

Birdal, T., Ilic, S.: Point pair features based object detection and pose estimation revisited. In: 3D Vision (3DV), 2015 2nd International Conference on, IEEE (2015) 527–535

ANSWER: DISTANCE TRANSFORMS

- Make use of the available model prior

$$\forall x \in \mathbb{V}(\mathbf{M}) \qquad D(x) = \min_i \|\mathbf{m}_i - x\| \qquad I(x) = \arg\min_i \|\mathbf{m}_i - x\|$$

Voxel Grid

Model point

Index of closest model point

- Pre-compute and store in voxel grid:
  - The distance to the closest model point
  - The index of the closest model point
- Approximate but very fast distance computations

EFFICIENT ICP USING DISTANCE TRANSFORMS

- Only a sparse set of scene points are used in verification.
- Naturally, we would like to minimize point-to-plane error:

$$E(S, M) = \sum_{i=1}^{N} ||(R \cdot \underset{\substack{\downarrow \\ \text{Scene}}}{s_i} + t - \underset{\substack{\downarrow \\ \text{Model}}}{m_i})^T \underset{\llcorner\rightarrow \text{Model Normal}}{n_{m_i}}||^2$$

- With distance transforms:

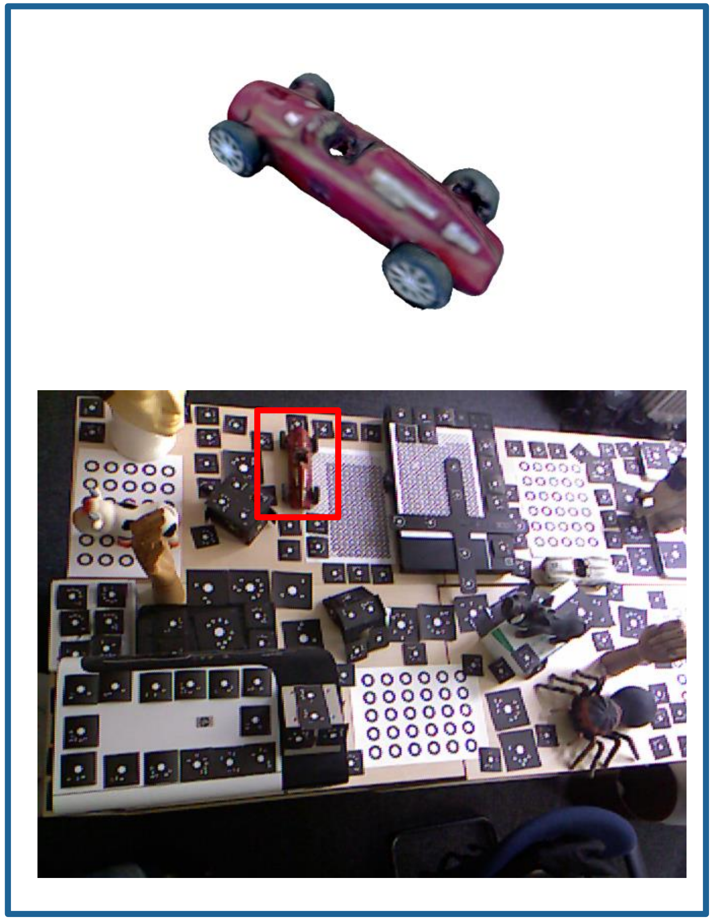
$$q_i = \mathbf{M}(I(T \circ p_i))$$

~0.8ms per hypothesis

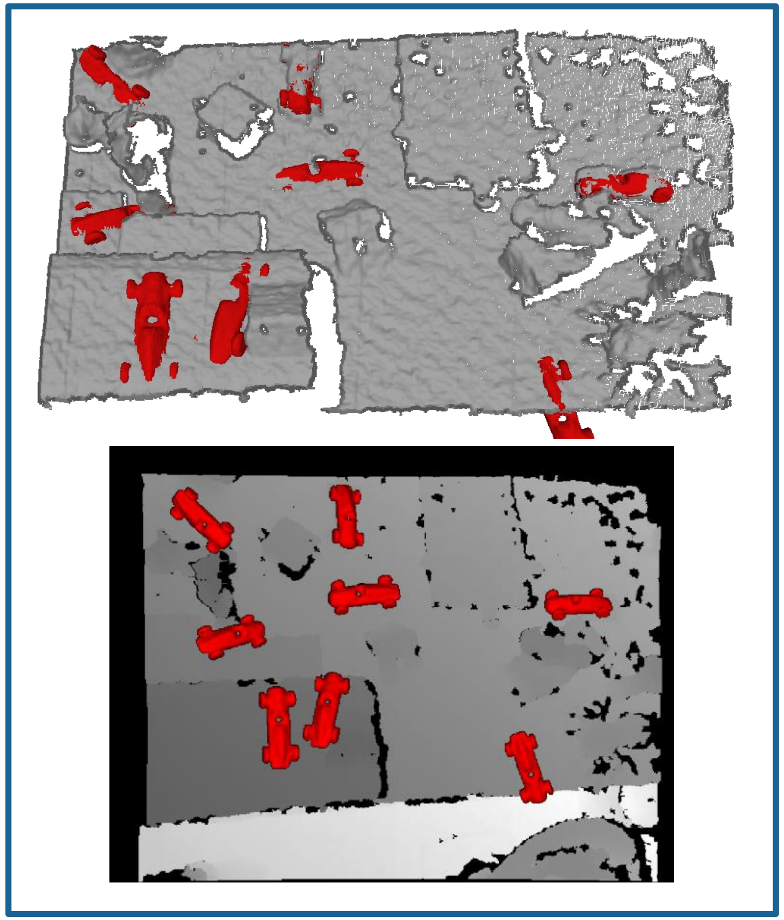
Birdal, T., Ilic, S.: Point pair features based object detection and pose estimation revisited. In: 3D Vision (3DV), 2015 2nd International Conference on, IEEE (2015) 527–535

ILLUSTRATION ON REAL KINECT DATA

Model & Data

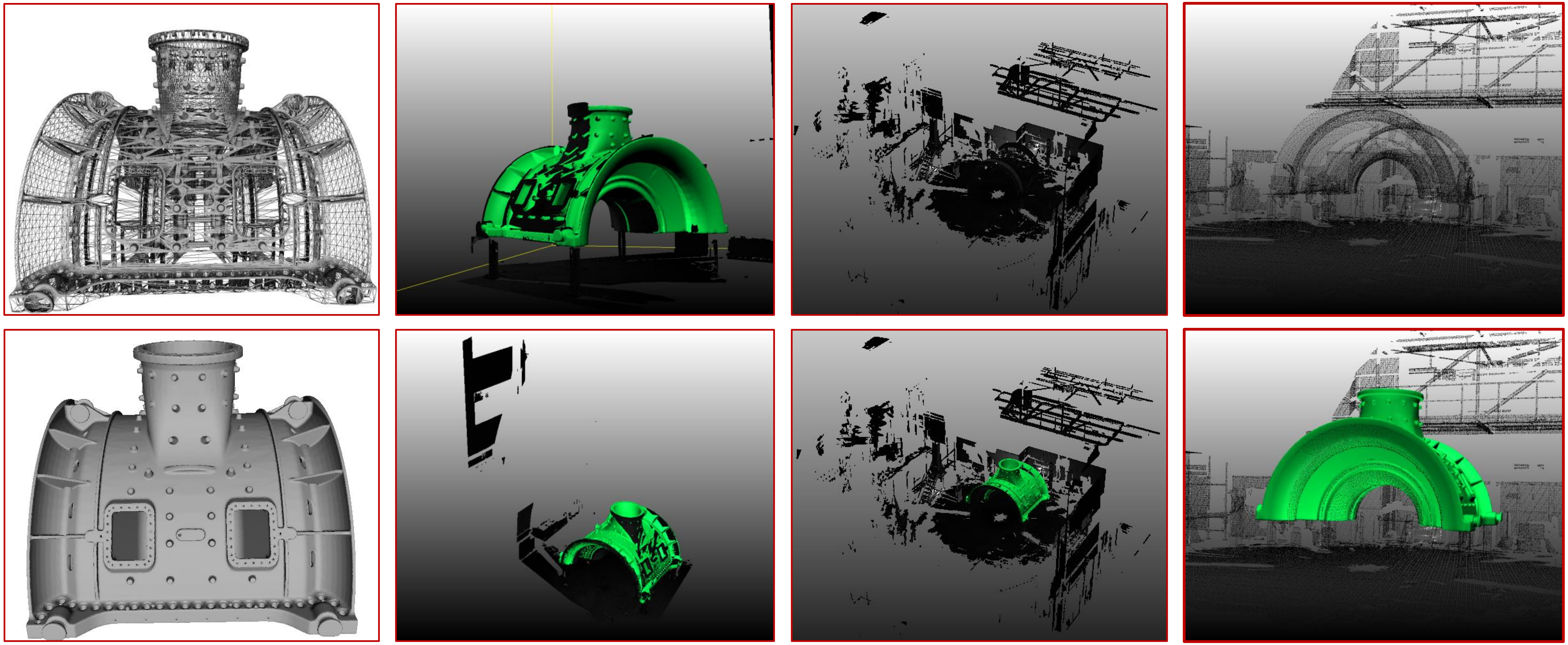


Detections



Best Hypothesis

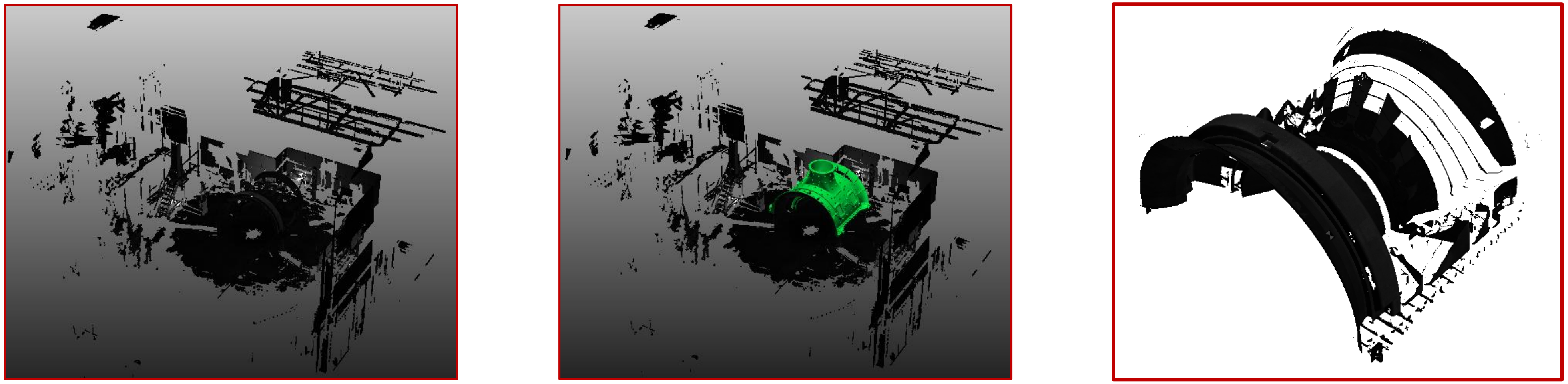
OBJECT DETECTIONS ON LASER SCANS



(a) (b) (c) (d)

SEGMENTATION

- Retrieve only points with close proximity to vertices of the CAD model



- Segmentation of Planes : Embedded into the voting stage

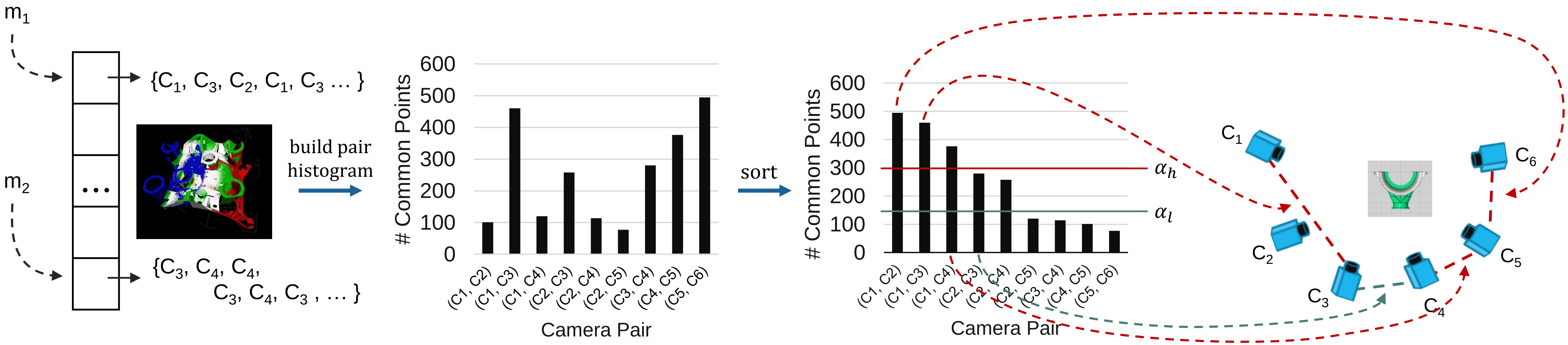
$$vote_{s_r,s_i} = \begin{cases} w_{plane} & |\angle(\mathbf{s_r},\mathbf{d})| \le \frac{\pi}{2} + \epsilon \wedge |\angle(\mathbf{s_i},\mathbf{d})| \le \frac{\pi}{2} + \epsilon \wedge |\angle(\mathbf{s_r},\mathbf{s_i})| \le 2\epsilon \\ 0, & \text{otherwise} \end{cases}$$

FINAL STITCHING

- So far:
  - Scans are transferred to the CAD model space
  - They are segmented
  - ICP registration is performed
- What remains:
  - Everything is related to CAD model, which might be far from reality
  - No pose relationships established
  - Global consistency is not enforced

COMPUTING THE POSE GRAPH

- Standard pipelines require exhaustive methods : e.g. Minimum spanning trees
- CAD Model eases this process:



(a) Index over Sampled Model Points **(b)** Histogram of Pairwise Overlaps **(c)** Sorted Histogram of Pairwise Overlaps **(d)** Iterative Insertion of Edges

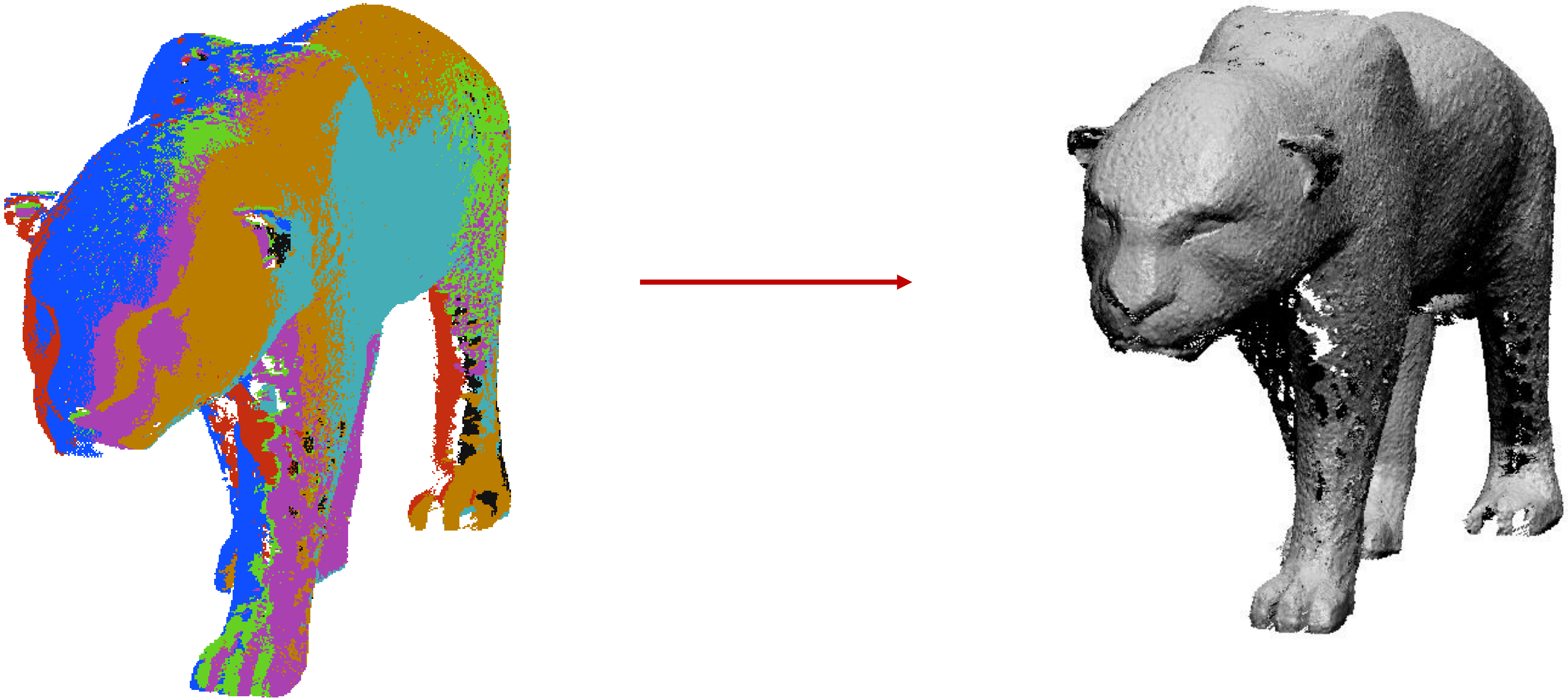
MULTIVIEW REGISTRATION

- Having the pose graph, and ignoring the CAD model, we globally optimize camera poses to bring scans into best alignment.

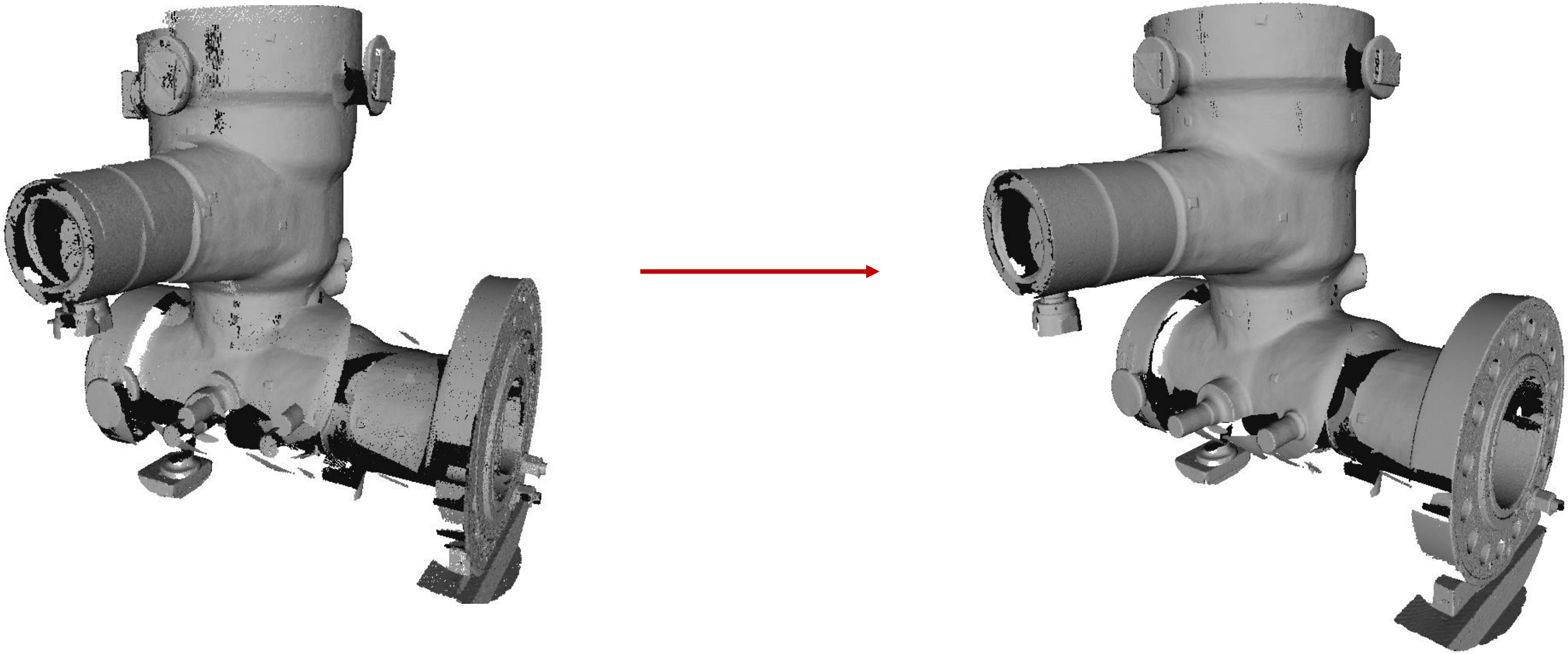
$$\{\theta_1, ..., \theta_M\} = \arg\min_{\theta}(E)$$

$$\underbrace{E(\theta_1, ..., \theta_M)}_{\text{Global Energy}} = \underbrace{\sum_{h=1}^{M}\sum_{k=1}^{M}}_{\text{Pose Pairs}} \underset{\substack{\downarrow \\ \text{Adjacency Matrix}}}{A(h,k)} \underset{\substack{\downarrow \\ \text{Scan Points}}}{\sum_{i=1}^{N_h}} \underset{\substack{\downarrow \\ \text{Robust Weights}}}{\rho}\left(\|d(\underset{\substack{\downarrow \\ \text{Pose Scan } h}}{\theta_h}(p_i^h), \underset{\substack{\downarrow \\ \text{Pose Scan } k}}{\theta_k}(q_i^h))\|^2\right)$$

MULTIVIEW REGISTRATION

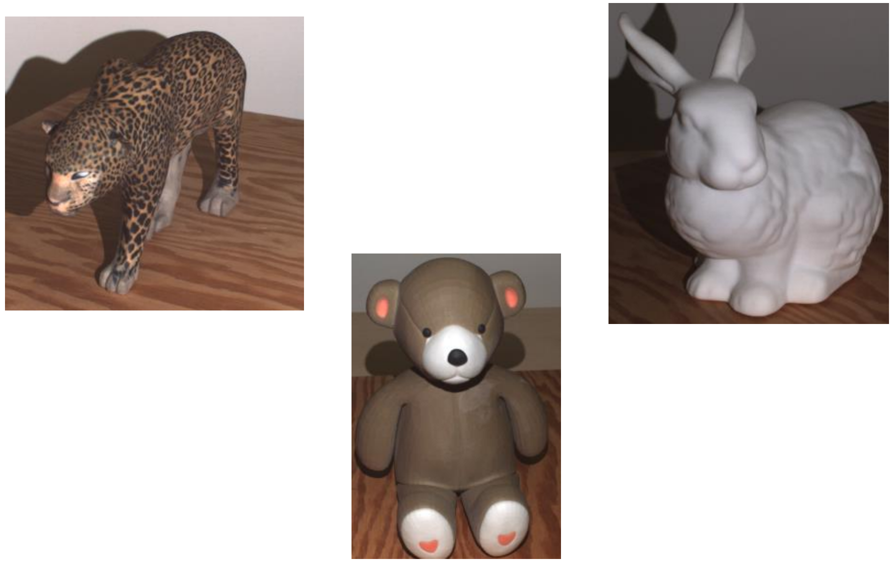


MULTIVIEW REGISTRATION



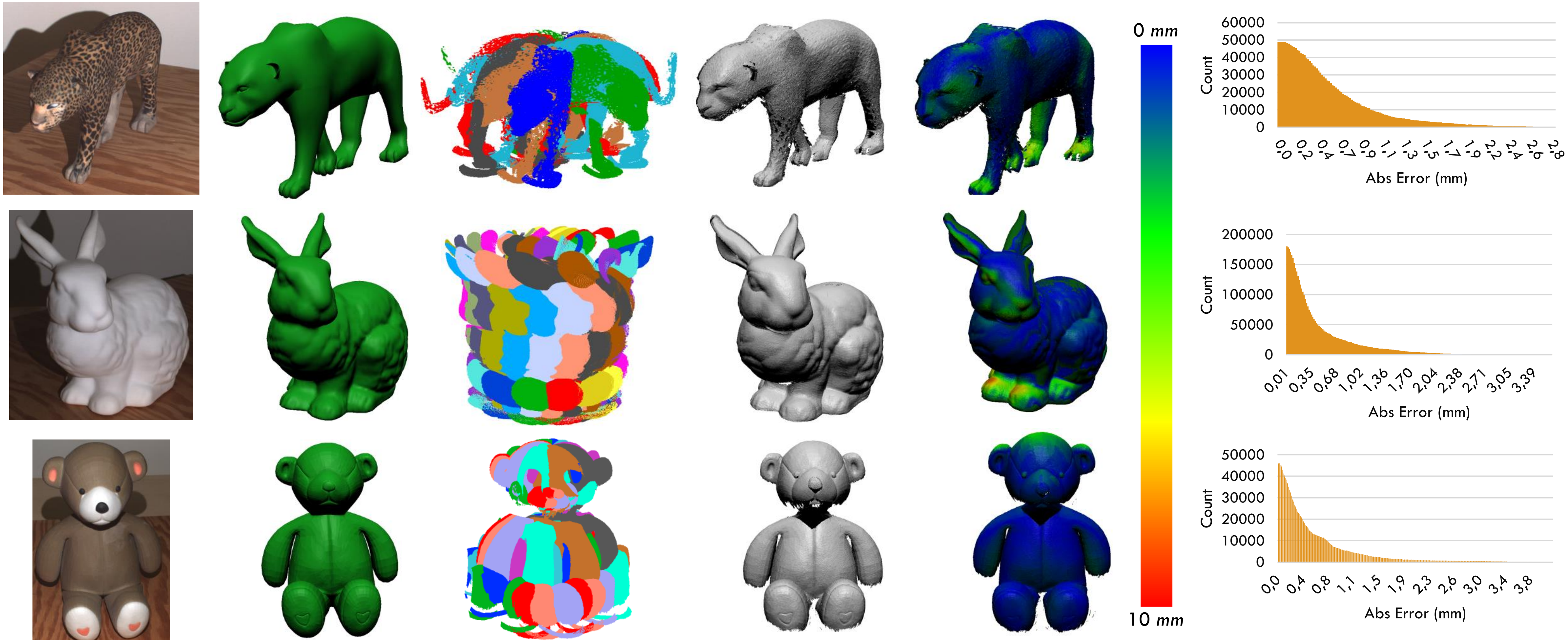
RESULTS

TOY OBJECTS



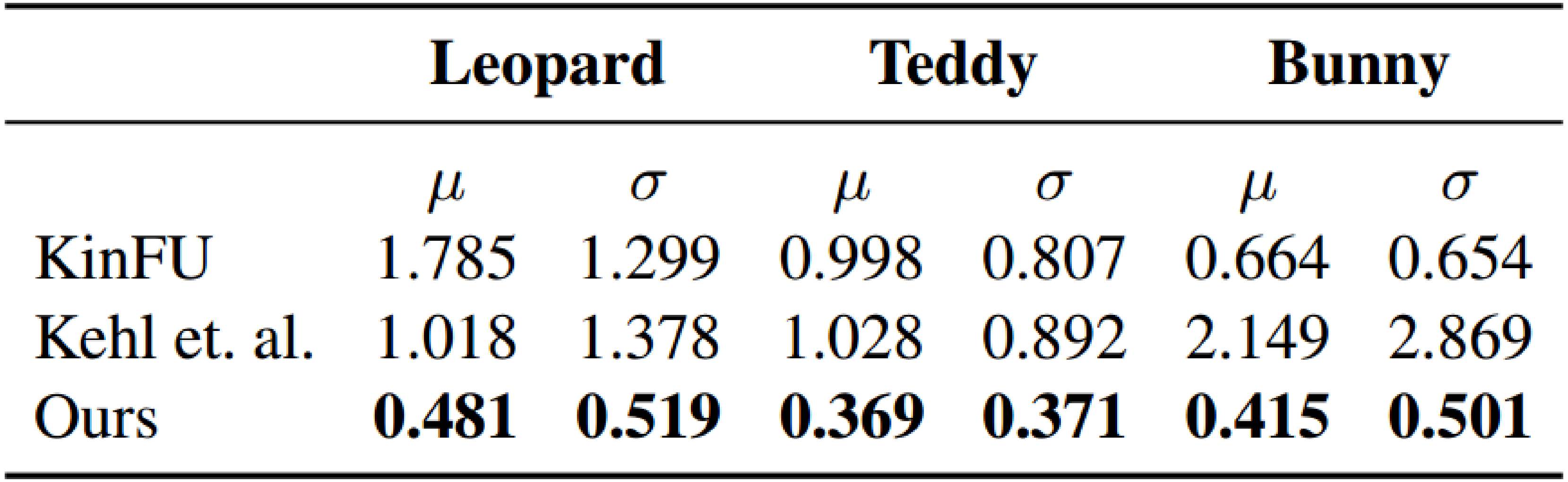
- 3 small objects: Decorative toys
- 3D printed from CAD models
- Printing accuracy <10µ
- 15cm – 30cm in diam.
- Captured with industrial structured light scanner (<0.25mm)
- 3D points are reconstructed from depth scenes (calibration errors are there)

RESULTS TOY OBJECTS



RESULTS ON TOY OBJECTS

- Accuracy assessment in comparison to standard volume based methods (in *mm*)



	Leopard		**Teddy**		**Bunny**	
	μ	σ	μ	σ	μ	σ
KinFU	1.785	1.299	0.998	0.807	0.664	0.654
Kehl et. al.	1.018	1.378	1.028	0.892	2.149	2.869
Ours	**0.481**	**0.519**	**0.369**	**0.371**	**0.415**	**0.501**

- Might not be a fair comparison as we have perfect CAD models

INDUSTRIAL OBJECTS

VENTIL

- 2^3 m^3 of volume
- 8 Surphaser laser scan
- 10M points / scan
- Taking a scan : ~20min.
- Availability of photogrammetry
- Availability of Surphaser reconstruction

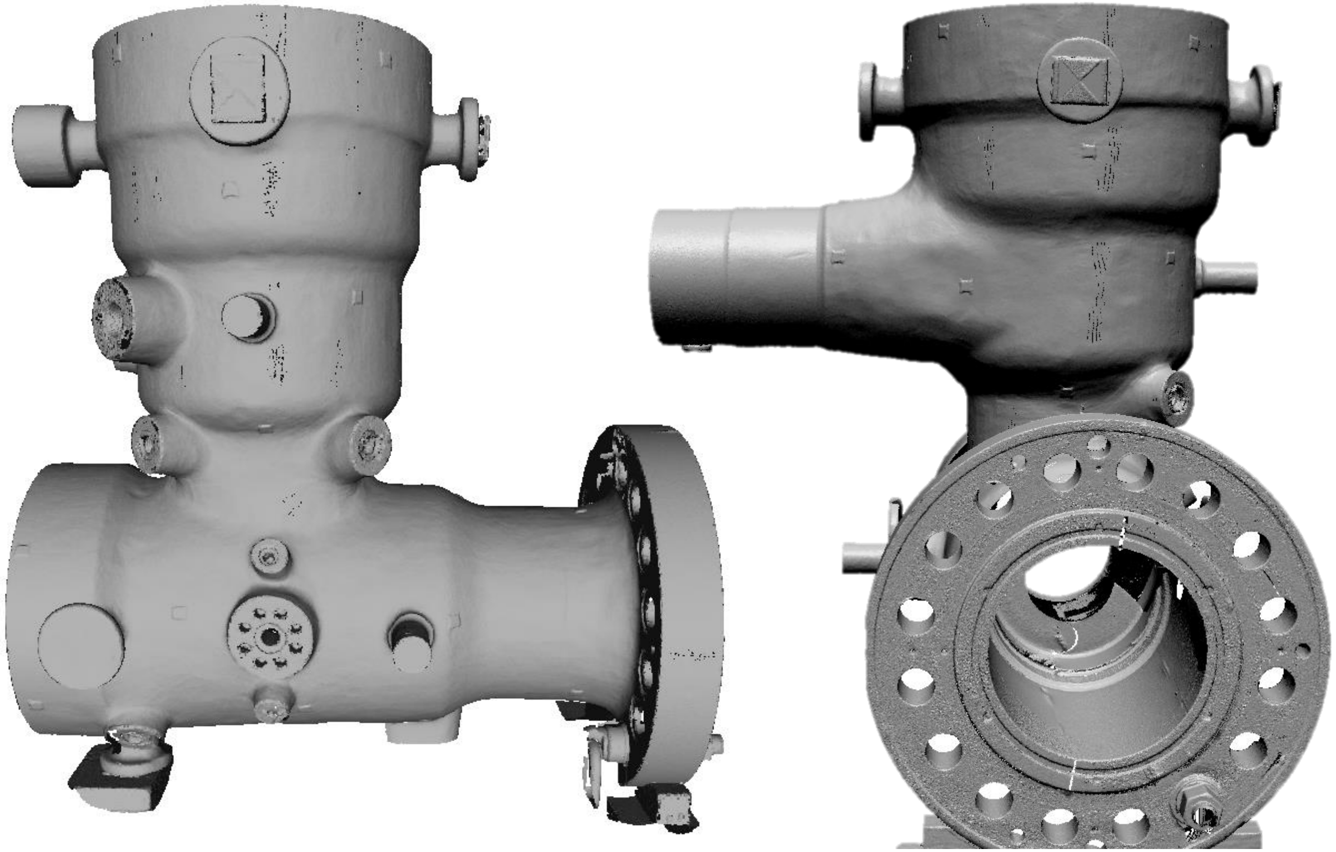
TURBINE

- 5^3 m^3 of volume
- 10 laser scans with Surphaser
- 10M – 100m points / scan
- Taking a scan : ~30min.
- Availability of photogrammetry

VENTIL

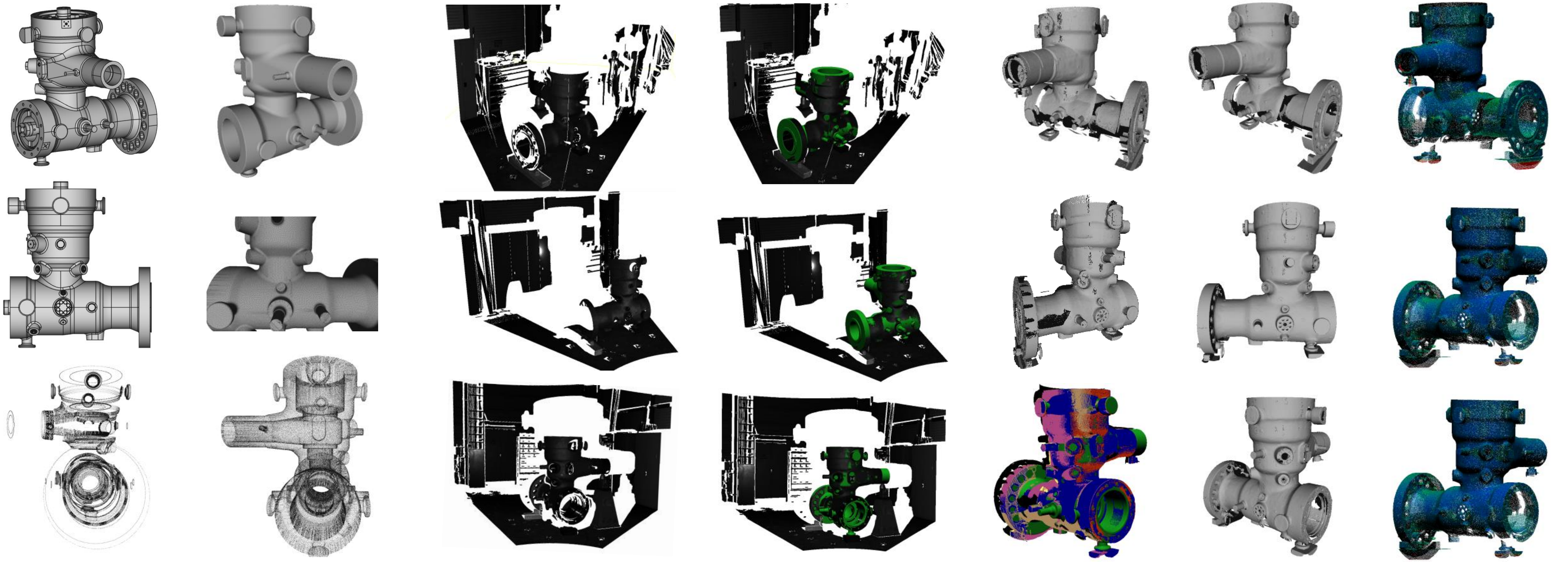
Surphaser Scanner

Ventil Part



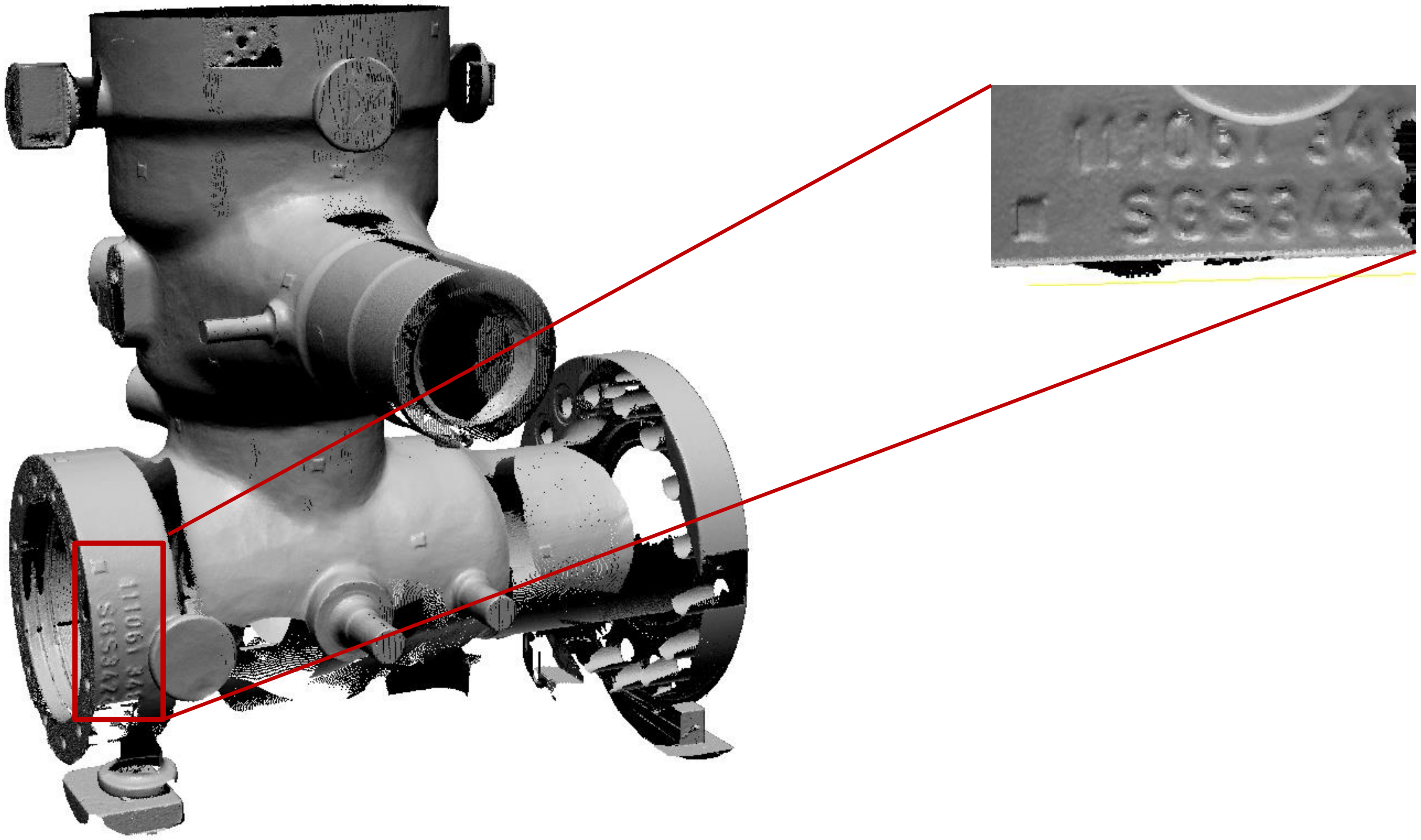
Our Reconstructions

WRAP-UP

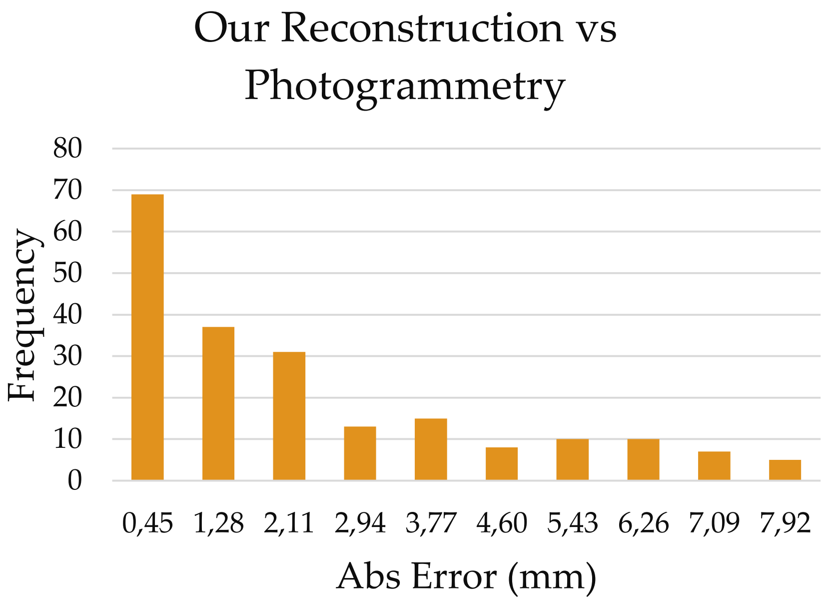


a) CAD Model b) CAD-Prep c) Input Scans d) Detections e) w/o Opt f) Optimized g) Surphaser vs Ours

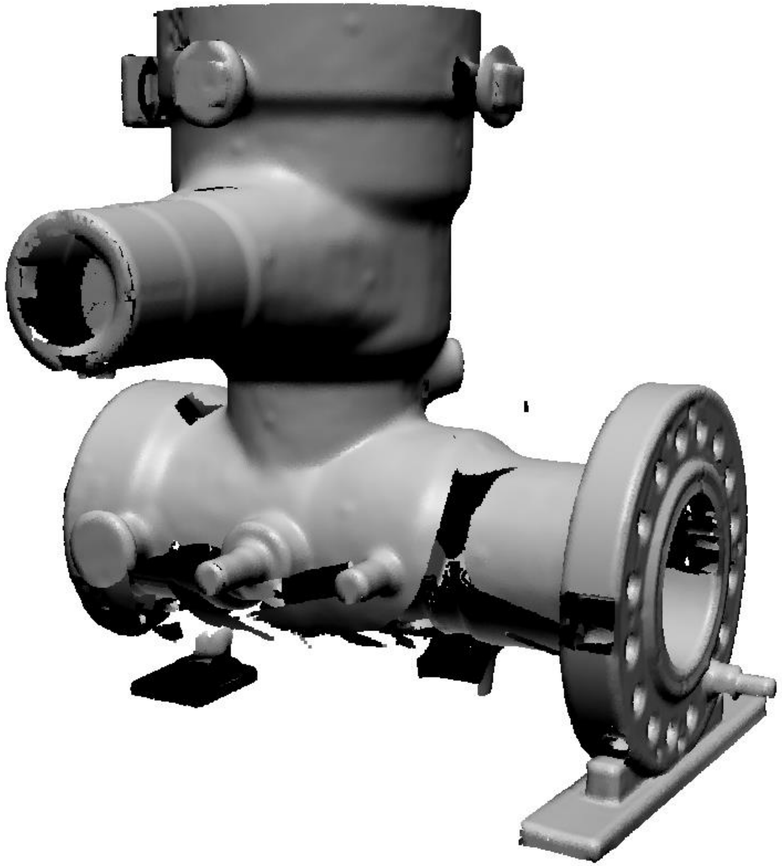
VENTIL



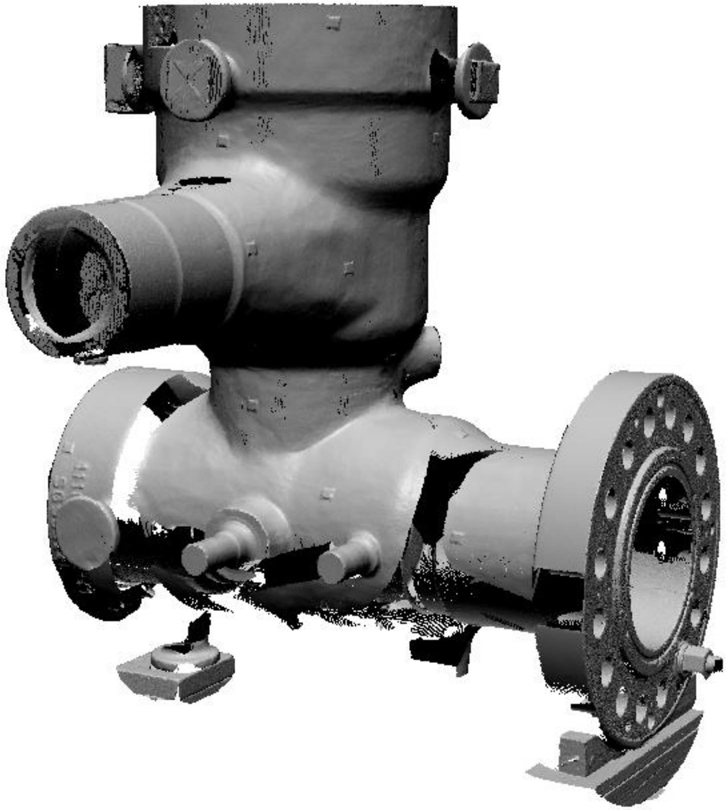
VENTIL: COMPARISON TO PHOTOGRAMMETRY



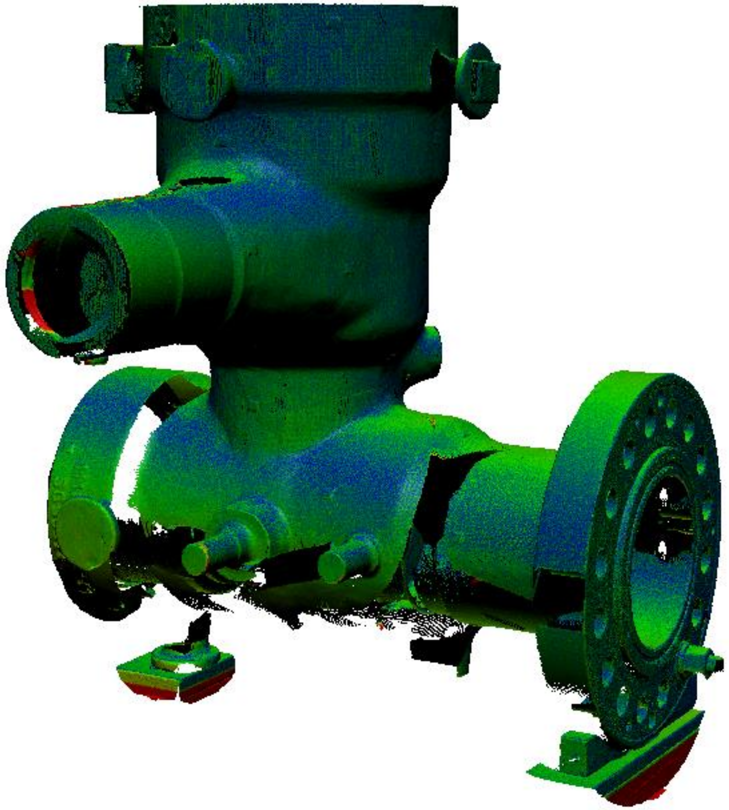
COMPARISON TO SURPHASER

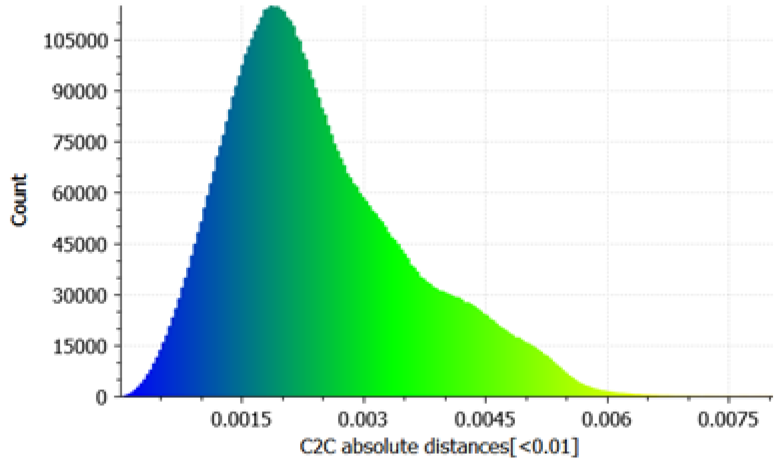


SURPHASER



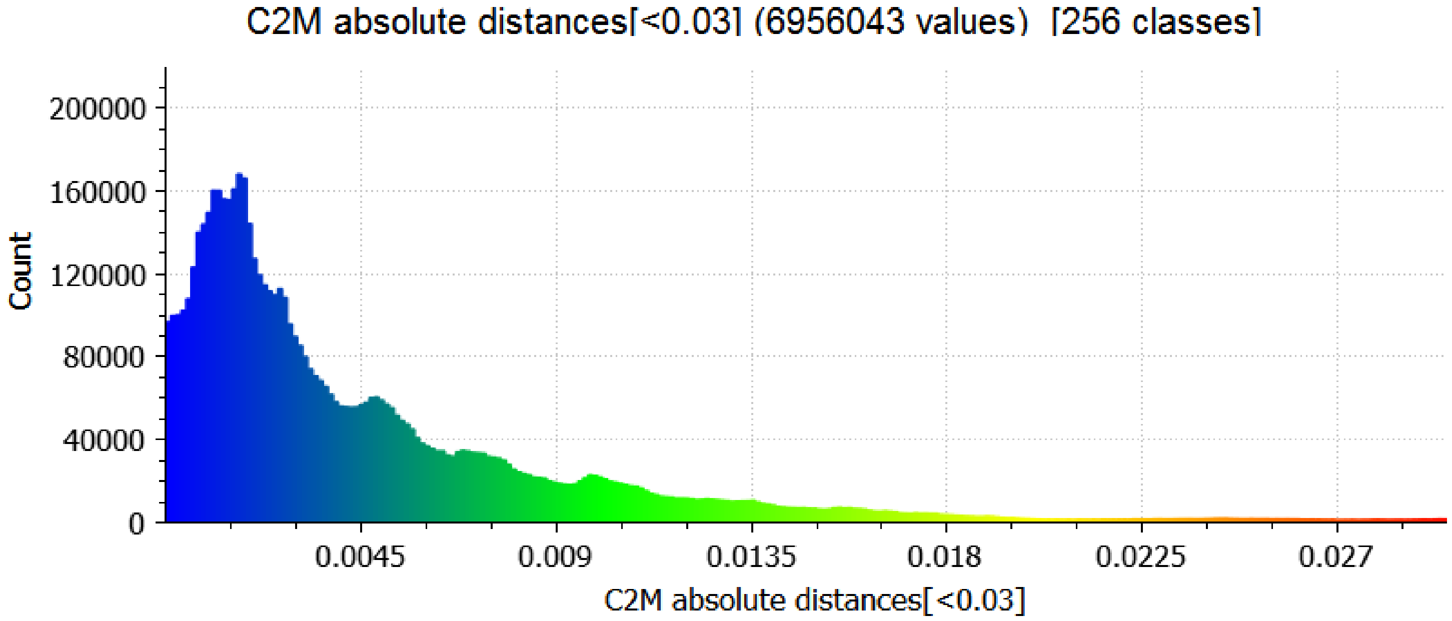
OURS

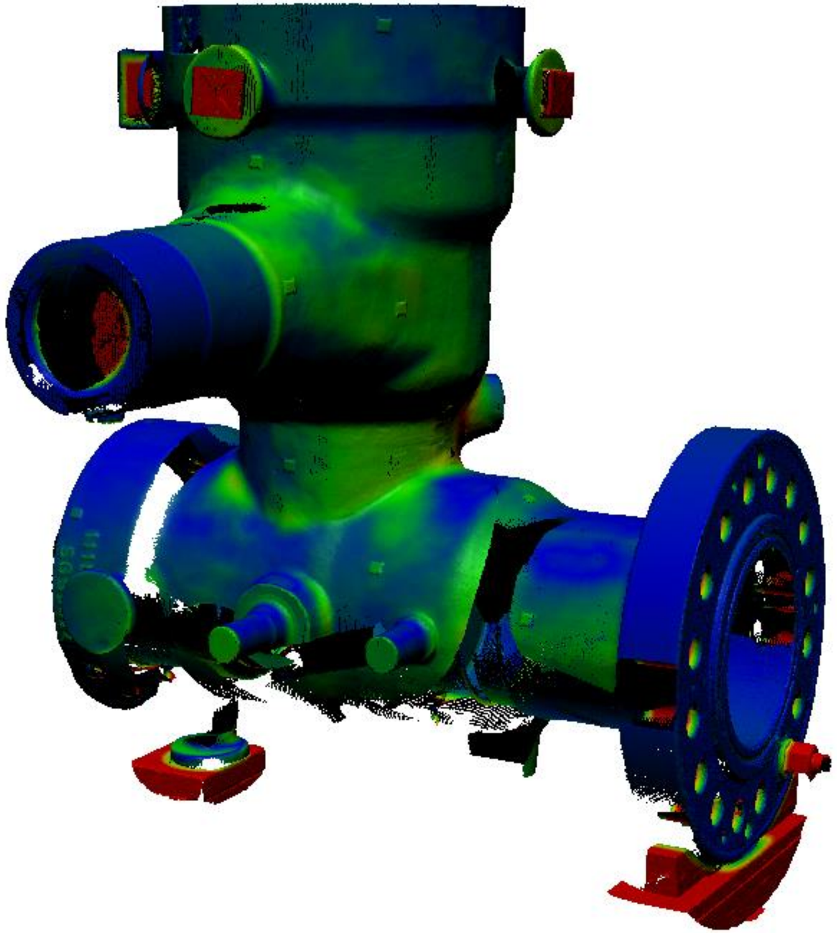




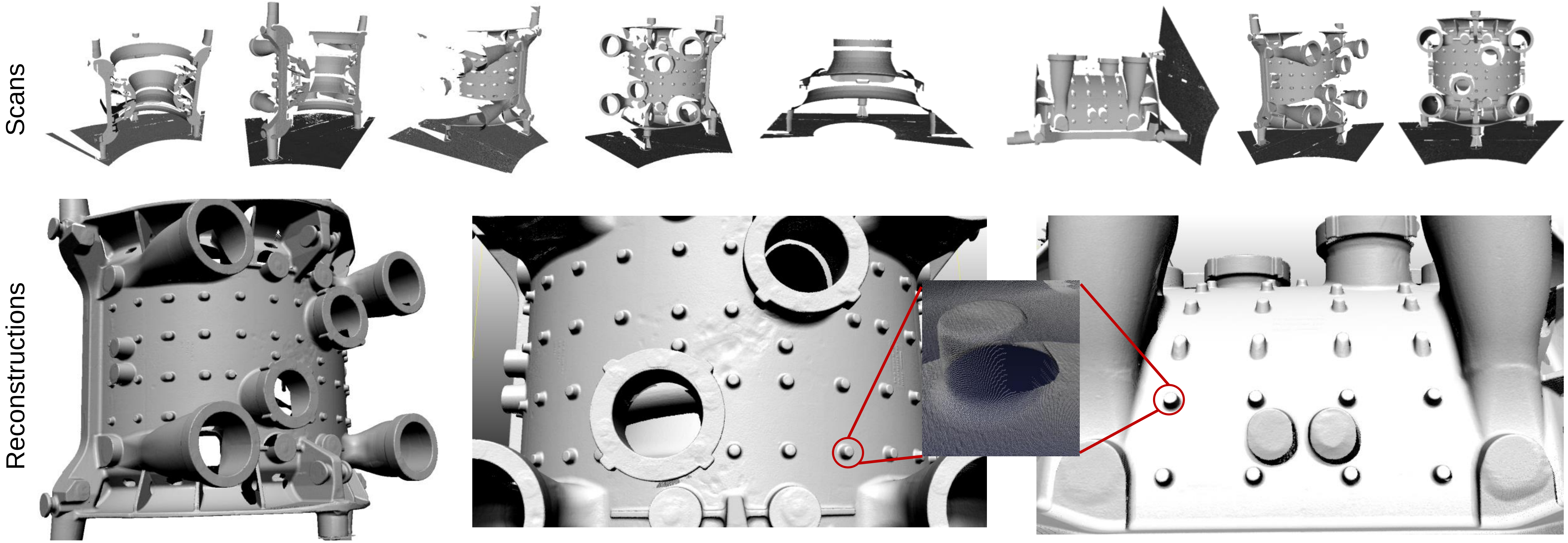
CAD MODEL VS RECONSTRUCTION

- CAD models do not correspond to the real manufactured objects.





TURBINE



TURBINE

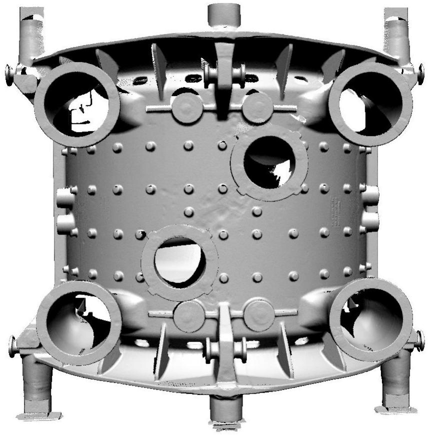
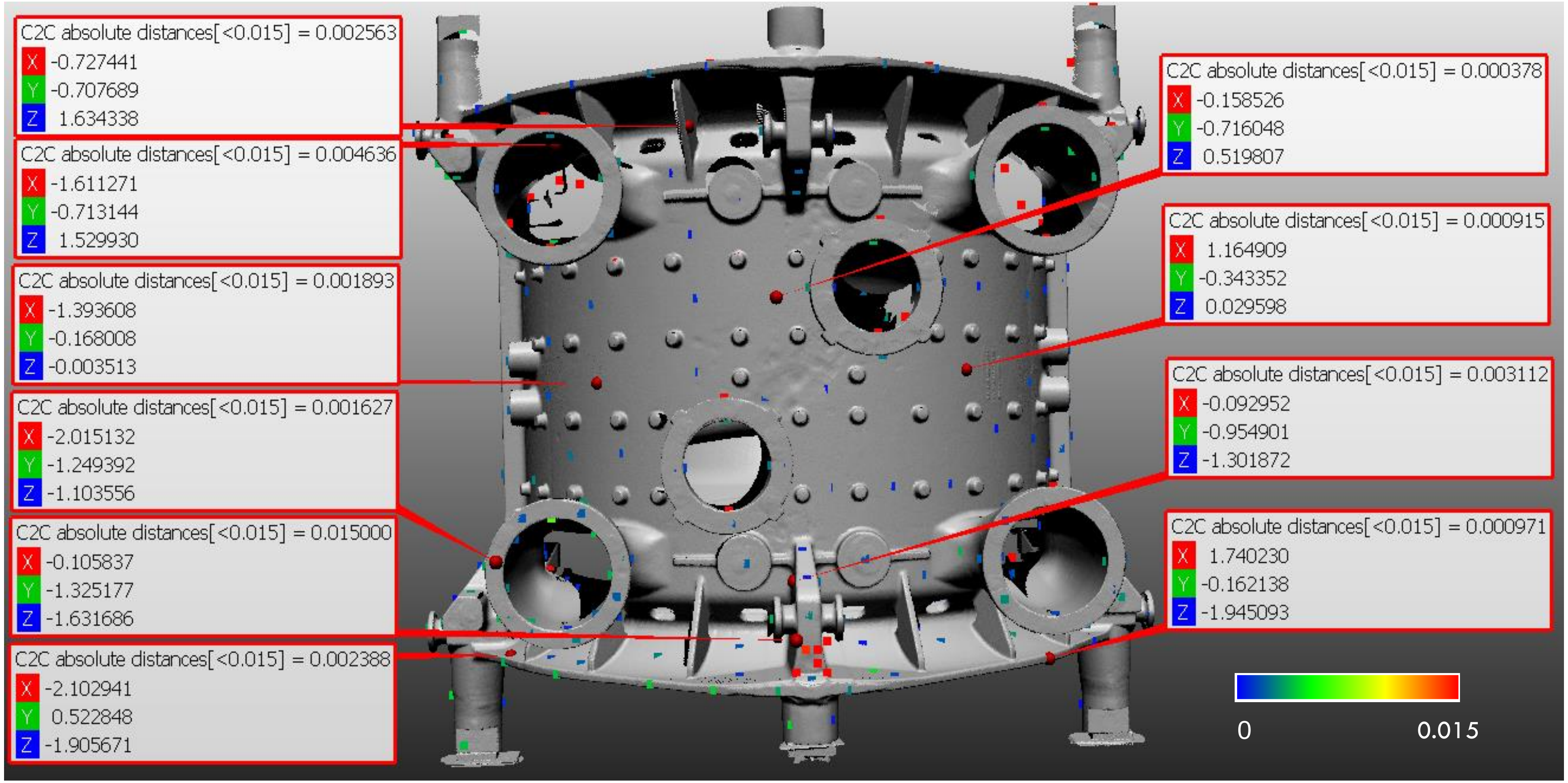


Figure shows a comparison to photogrammetry (mm)



C2C absolute distances[<0.015] = 0.002563
X -0.727441
Y -0.707689
Z 1.634338

C2C absolute distances[<0.015] = 0.004636
X -1.611271
Y -0.713144
Z 1.529930

C2C absolute distances[<0.015] = 0.001893
X -1.393608
Y -0.168008
Z -0.003513

C2C absolute distances[<0.015] = 0.001627
X -2.015132
Y -1.249392
Z -1.103556

C2C absolute distances[<0.015] = 0.015000
X -0.105837
Y -1.325177
Z -1.631686

C2C absolute distances[<0.015] = 0.002388
X -2.102941
Y 0.522848
Z -1.905671

C2C absolute distances[<0.015] = 0.000378
X -0.158526
Y -0.716048
Z 0.519807

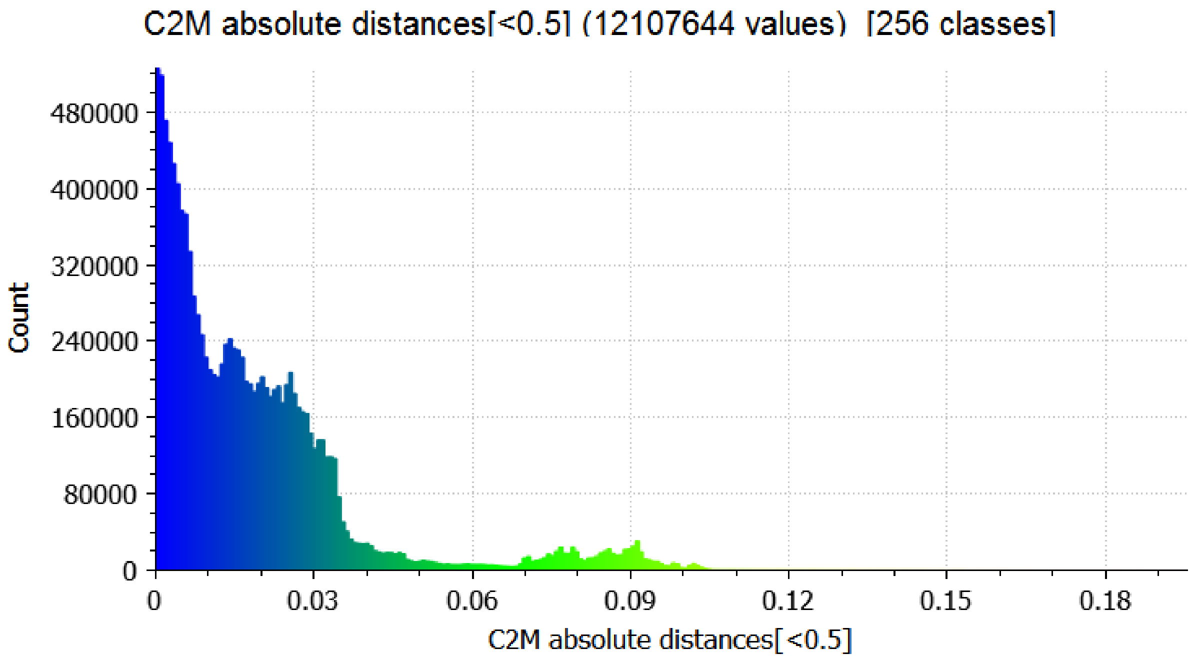
C2C absolute distances[<0.015] = 0.000915
X 1.164909
Y -0.343352
Z 0.029598

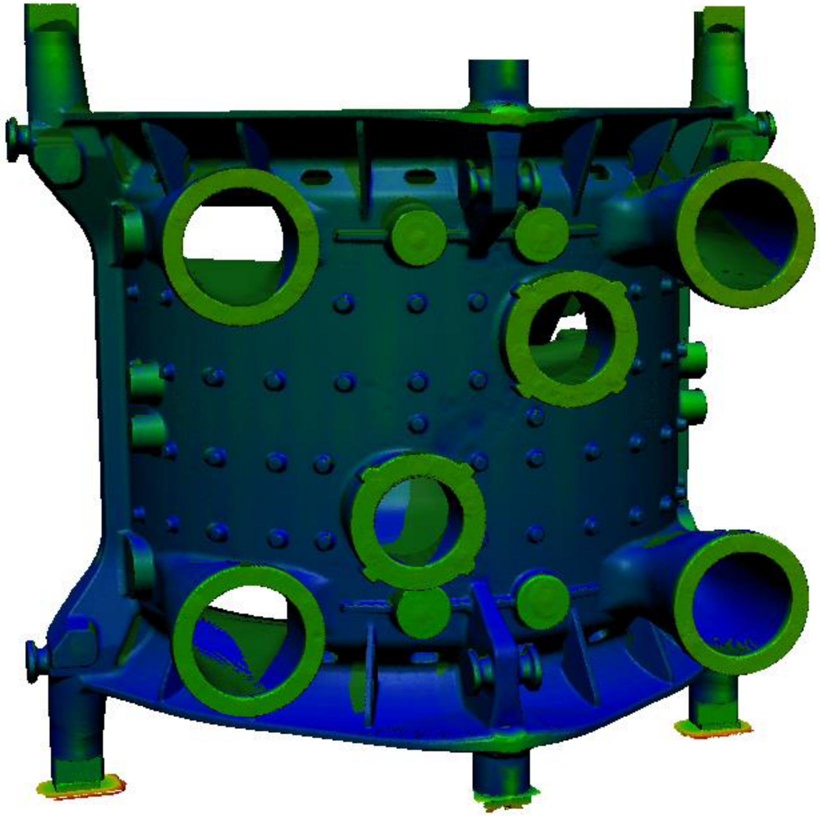
C2C absolute distances[<0.015] = 0.003112
X -0.092952
Y -0.954901
Z -1.301872

C2C absolute distances[<0.015] = 0.000971
X 1.740230
Y -0.162138
Z -1.945093

CAD MODEL VS RECONSTRUCTION

- CAD models do not correspond to the real manufactured objects.





APPROXIMATE TIMINGS

- Averaged over available objects
- Windows OS, Intel i5 CPU with 16GB of RAM
- Parameters are tuned for best performance vs speed trade-off.

	CAD-prep	Object Detection	Verification	Global MV ICP
Toy Objects	40.41	0.34	0.024	42.60
Industrial Parts	64.37	3.10	0.27	112.94

FUTURE WORK

- Symmetric objects are not seamlessly handled
- A more accurate segmentation is pluggable: Graph Cuts, CRF with model prior ?
- A vast literature on Multiview registration exists : A survey ?
- Final meshing remains to be an open problem – Marching Cubes, Poisson, Smooth Signed Distances all have their own flaws
- Extensive evaluation on other objects with photogrammetric studies

THANK YOU FOR YOUR ATTENTION

You could reach me at:

TOLGA BIRDAL

tolga.bridal@tum.de

www.tbirdal.me