Master in Computer Vision
Barcelona

Module: M4 3D Vision

Lecture: **4.10 Point Cloud Processing**

Lecturer: Josep R. Casas

3D data representation

- Depth map
  - **2.5D** (concept: RGBD = 2D + depth)

- Point cloud
  - **organized**: keeps relationships in sensor neighborhood
  - **unorganized**: one can ***just*** compute nearest neighbors in 3D

- Mesh
  - nice scanned/reconstructed surfaces: watertight / convex...

Human Body Analysis using Depth Data

Humans **naturally understand** the Human Body

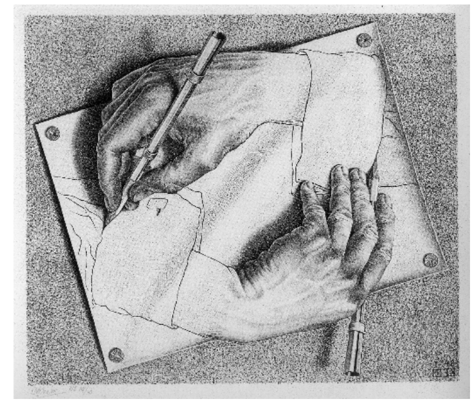
…what about **computers**?

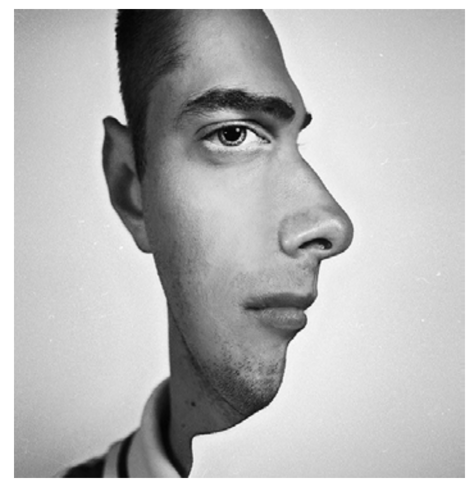
eyes → camera

brain → CPU

~~other~~ → ~~sensors, …~~

Computer Vision



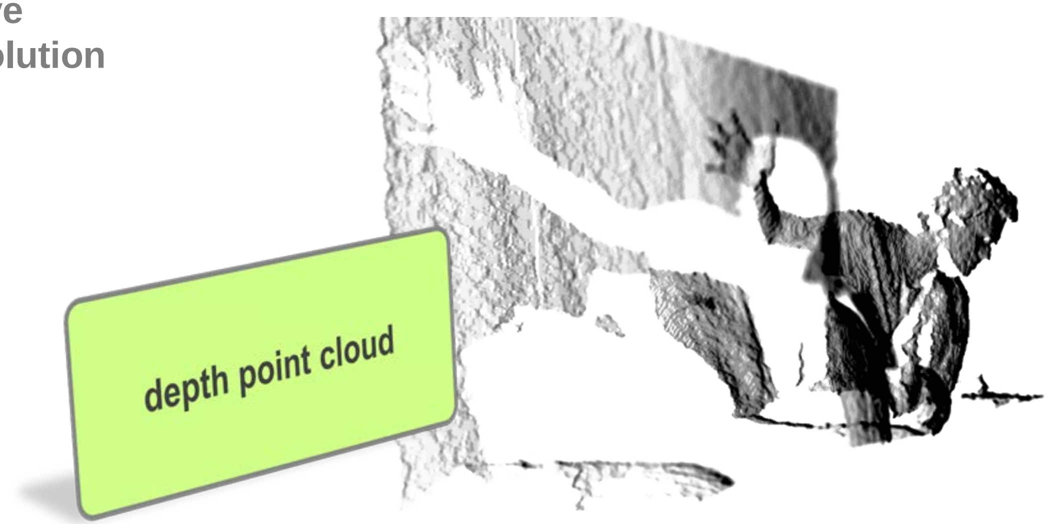


Why? Human-Machine interaction, deaf people, animation, surgery, leisure, etc.

Human Body Analysis using Depth Data

Big
Expensive
Low-resolution

Small
~150€
high-resolution

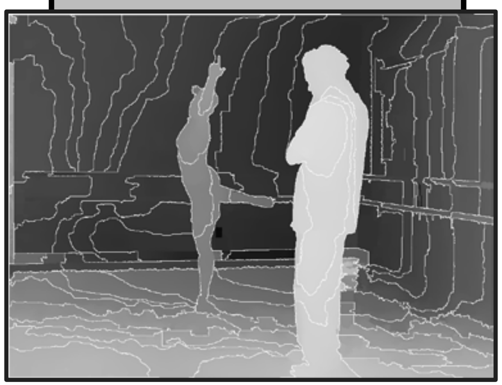


Depth map / Point Cloud processing

Objective: To obtain information from depth camera frames

Depth Data

- Poor texture
- Sharp edges
- Low resolution



?

2D descriptors

- SIFT
- SURF
- Freak
- …

3D descriptors

- 3DSC [1]
- VFH [2]
- SHOT [3]
- …

Other proposals

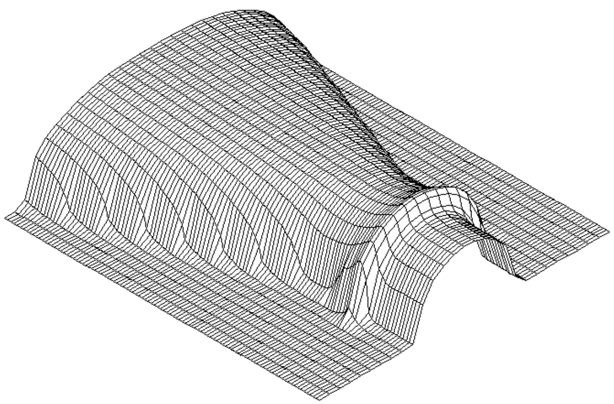
- **ORD [4]**
- R-NBLS

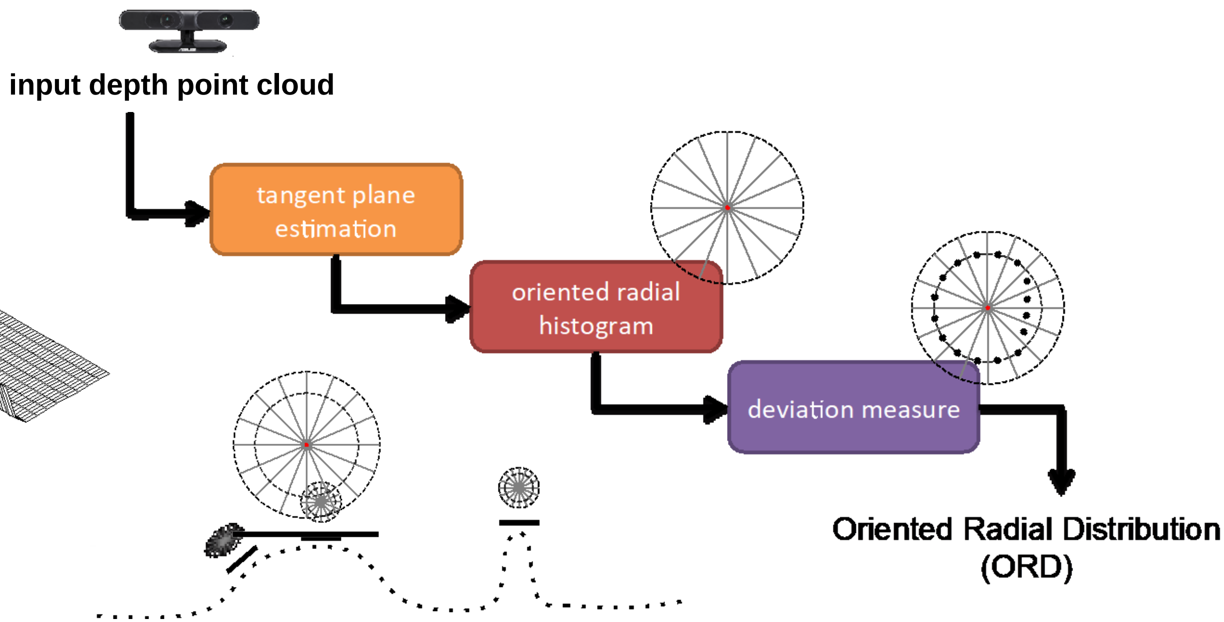
[1] Andrea Frome et al. Recognizing objects in range data using regional point descriptors. ECCV 2004
[2] R.B. Rusu, et al. Fast 3d recognition and pose using the viewpoint feature histogram. IROS 2010
[3] F. Tombari, et al. Unique signatures of histograms for local surface description. ECCV 2010
[4] X. Suau, et al, "Oriented Radial Distribution on Depth Data: Application to the Detection of End-Effectors," ICASSP 2012

Oriented Radial Distribution

Objective: Detect prominent and flat zones of a depth point cloud

depth point cloud
(simplified example)





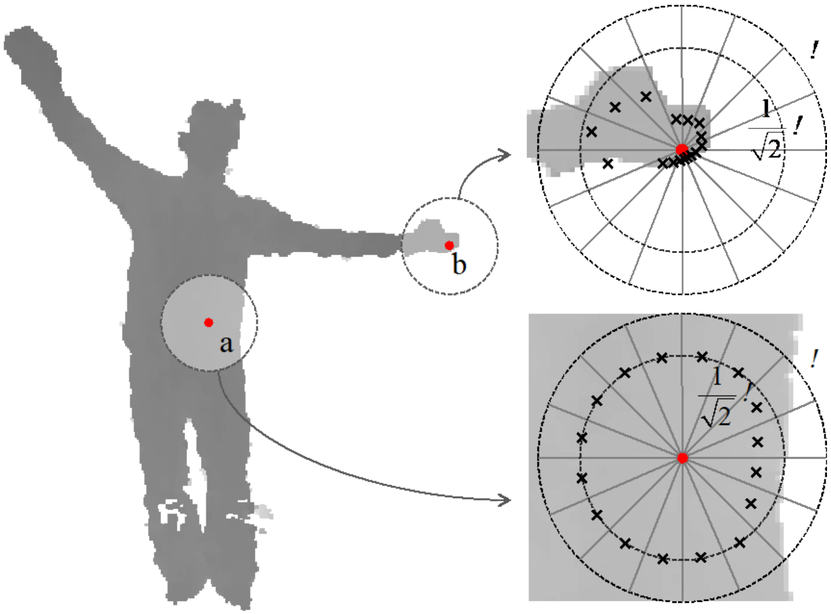
[4] X. Suau, et al, "Oriented Radial Distribution on Depth Data: Application to the Detection of End-Effectors," ICASSP 2012

Oriented Radial Distribution

Objective: Detect prominent and flat zones of a depth point cloud

ORD Characteristics

- **Oriented** to the surface normals
- **Local** computation (neighborhood of a point)
- **Multiscale** (disk radius)
- Output: **histogram** or **scalar**



Deviation measure: $$\Theta(\mathbf{z}, \Omega, \xi) = \frac{1}{\frac{1}{\sqrt{2}}\rho K_f} \sum_{j=0}^{K_f} \left(\bar{\delta}_j - \frac{1}{\sqrt{2}}\rho \right)$$

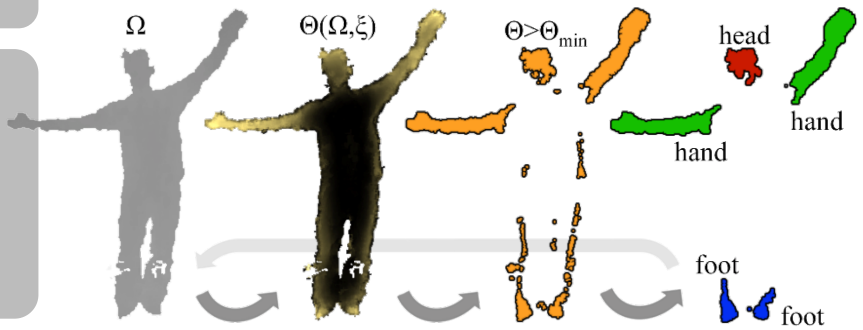
a : low deviation
→ **FLAT**

b : high deviation
→ **PROMINENT**

Oriented Radial Distribution

Classification of prominent zones

We propose to use probabilistic descriptors
- **Y :** Position, relative height of zone
- **S :** Size, area of the zone
- **A :** Shape, PCA aspect ratio of zone



Statistical moments of the descriptors

λ_k	$\mu_{\lambda_k}^{head}$	$\sigma_{\lambda_k}^{head}$	$\mu_{\lambda_k}^{hand}$	$\sigma_{\lambda_k}^{hand}$	$\mu_{\lambda_k}^{foot}$	$\sigma_{\lambda_k}^{foot}$
Y	62.18	7.48	29.43	29.06	−71.31	10.89
S	58.58	10.00	64.24	24.89	46.68	10.75
A	0.58	0.17	0.11	0.13	0.41	0.19

A blob **B** is classified
γ_i**={head,hand,foot,**nothing**}**
depending on its
combined probability

$$P(\mathcal{B}=\gamma_i) = P\Big((Y_{\mathcal{B}}=\gamma_i)\wedge(S_{\mathcal{B}}=\gamma_i)\wedge(A_{\mathcal{B}}=\gamma_i)\Big)$$
$$= f_Y^{\gamma_i}(\mathcal{B})\cdot f_S^{\gamma_i}(\mathcal{B})\cdot f_A^{\gamma_i}(\mathcal{B})$$

$$\text{with PDF:}\quad f_{\lambda_k}^{\gamma_i}(\mathcal{B}) = \frac{1}{\sigma_{\lambda_k}^{\gamma_i}\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{\lambda_k(\mathcal{B})-\mu_{\lambda_k}^{\gamma_i}}{\sigma_{\lambda_k}^{\gamma_i}}\right)^2}$$

Other Point Feature Representations

- **Signature of Histograms of OrienTations (SHOT)**
- **VFH signatures**
- **Point Feature Histograms (PFH)**
- **Fast Point Feature Histograms (FPFH)**
  - See tutorials in PCL: http://pointclouds.org/documentation/tutorials

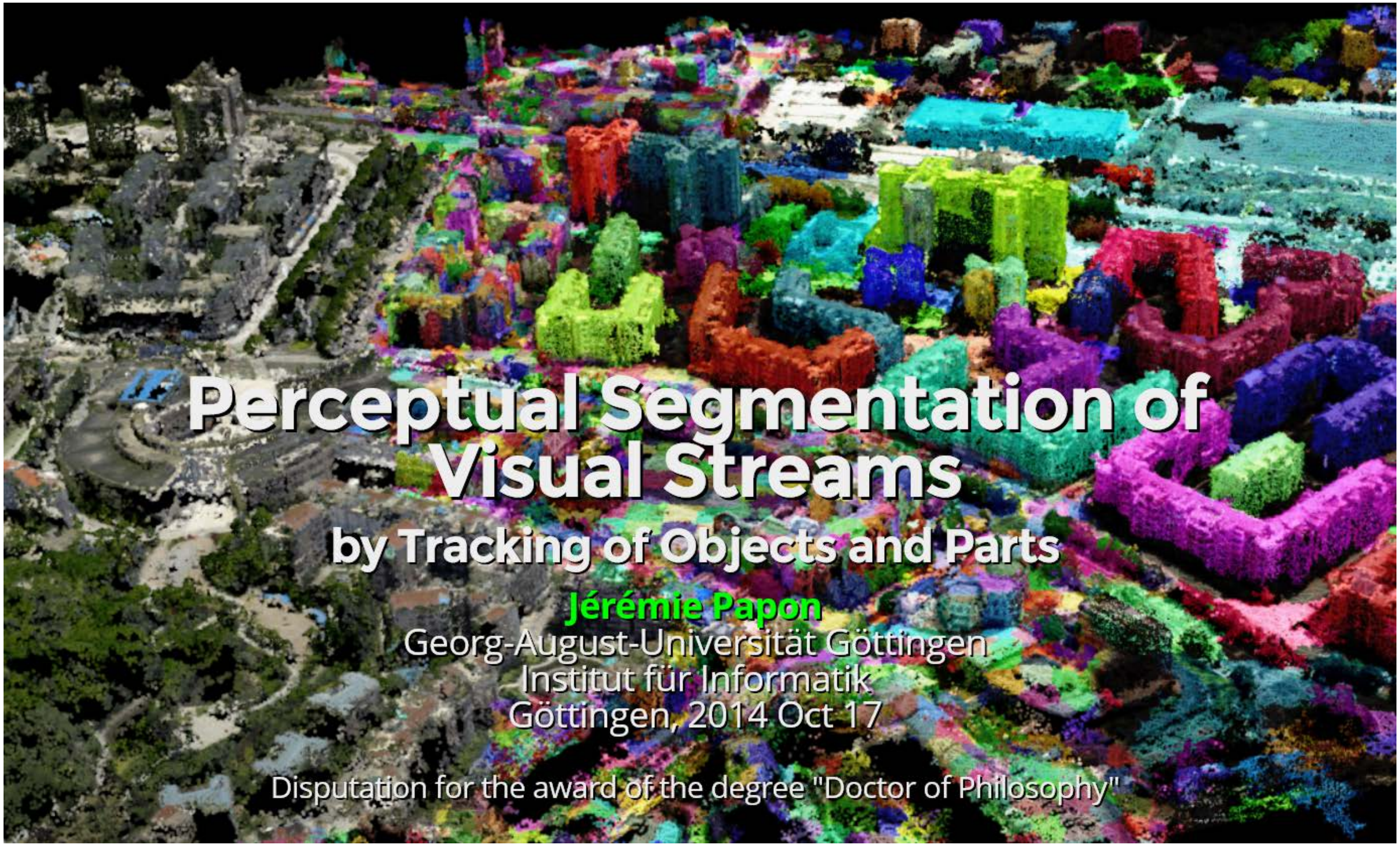
Point feature representations

A good point feature representation distinguishes itself from a bad one, by being able to capture the same local surface characteristics in the presence of:

- rigid transformations - 3D rotations and translations in the data should not influence the resultant feature vector F estimation
- varying sampling density - a local surface patch sampled more or less densely should have the same feature vector signature
- noise - the point feature representation must retain the same or very similar values in the presence of mild noise in the data.

R. B. Rusu, "Semantic 3D Object Maps for Everyday Manipulation in Human Living Environments," PhD TUM 2009

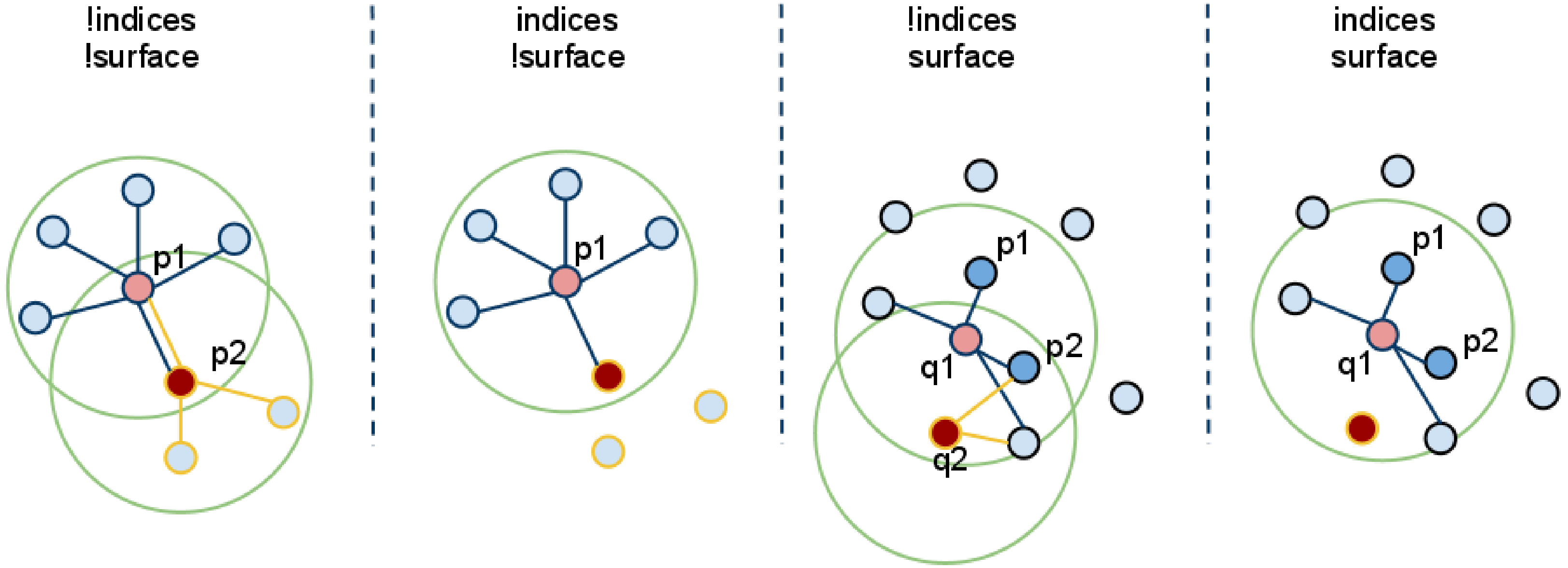
Unorganized Point Cloud Processing



http://jpapon.github.io

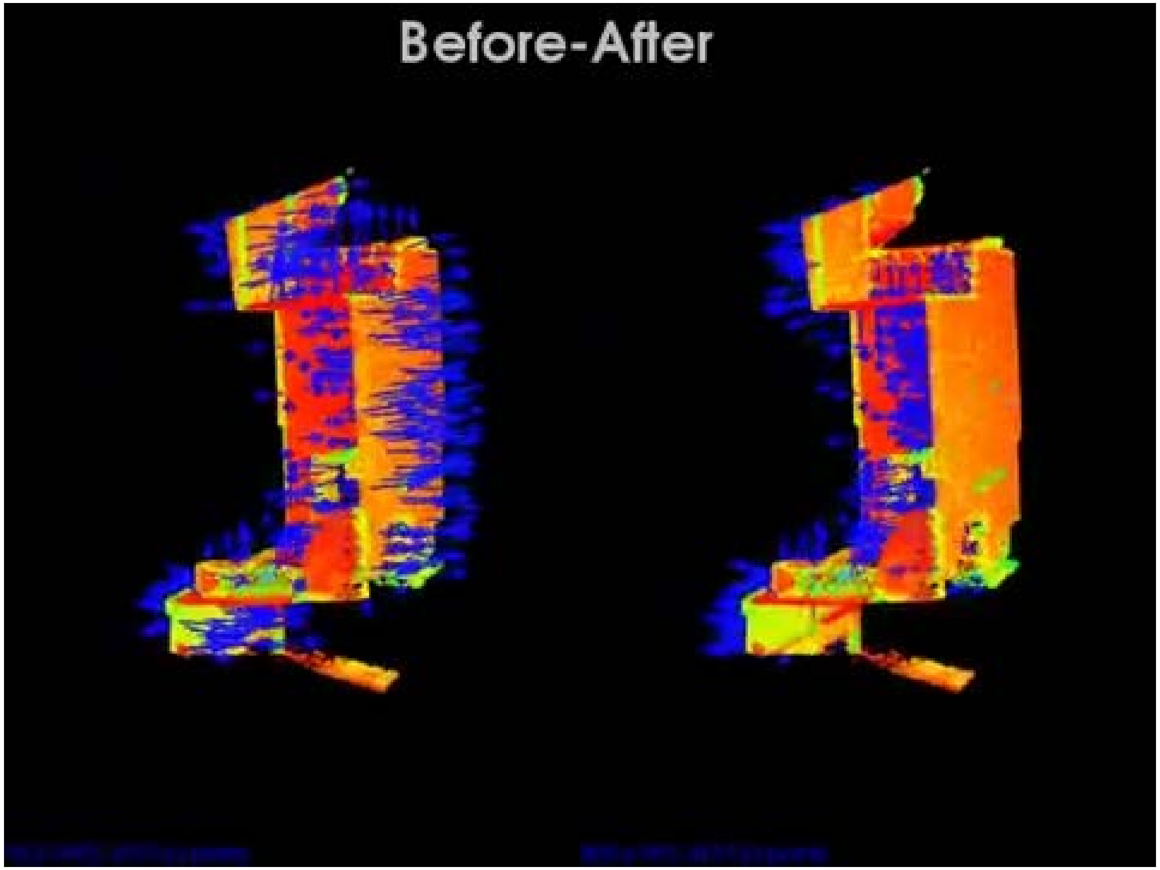
Unorganized Point Cloud Processing

Importance of the local neighborhood



Estimating Surface Normals

- Consistent normal orientation (PCL)



Martin Matilla, Alignment of 3D Point Clouds and RPS Detection

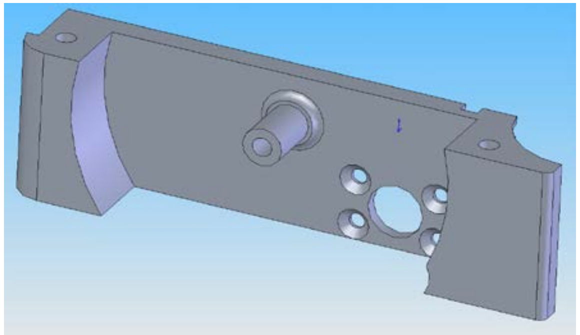
POINT CLOUD PROCESSING IN AUTOMOTIVE INDUSTRY

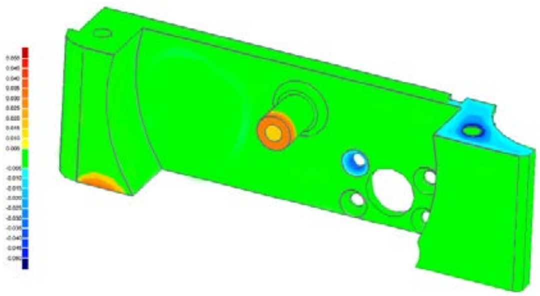
Industrial (automotive) Production

Computer Aided Design (CAD)

Computer Numerical Control (CNC)

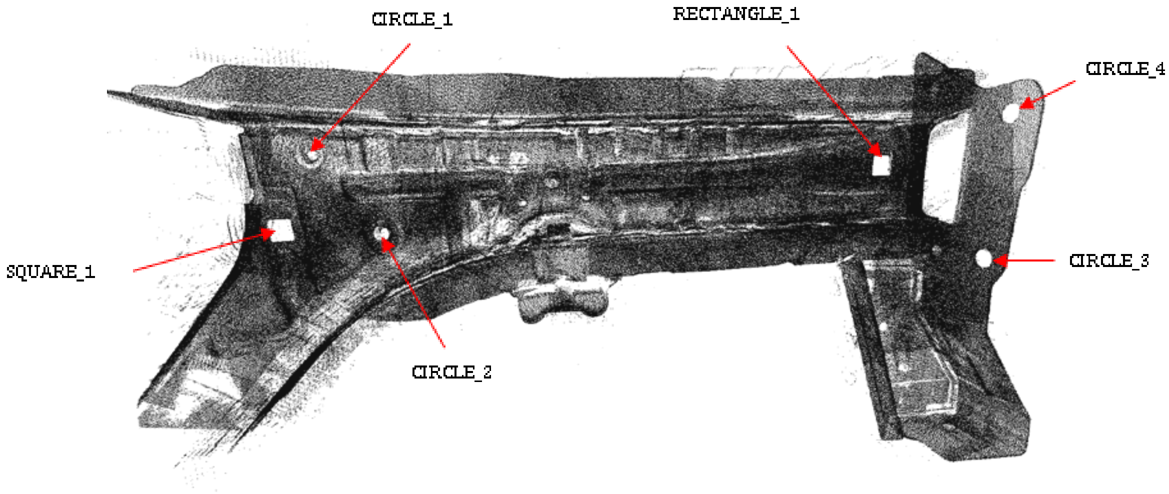
Difference CAD - CNC





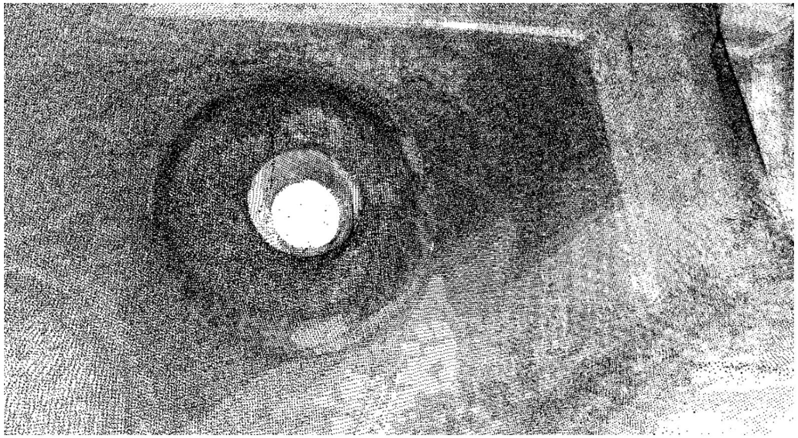
Development

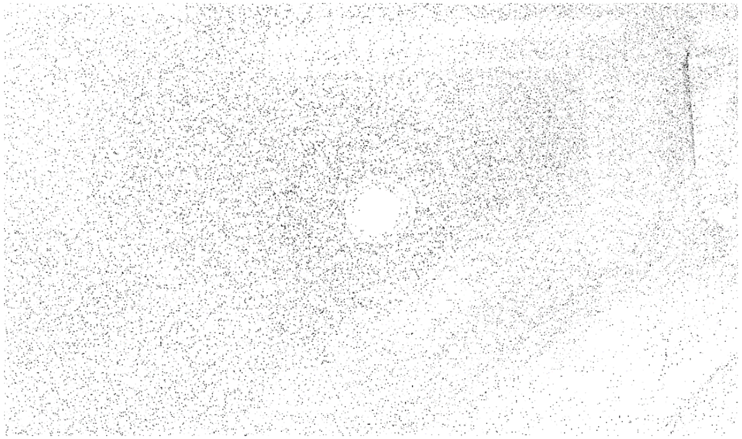
Data



IDENTIFIER	X (MM)	Y (MM)	Z (MM)	DIAMETER(MM)
RECTANGLE_1	-442.0	467.6	348.5	20 - 25
SQUARE_1	100.0	405.0	308.0	20
CIRCLE_1	61.1	415.4	371.0	24.0
CIRCLE_2	9.9	430.5	302.5	14.0
CIRCLE_3	-491.012	536.68	266.12	16.0
CIRCLE_4	-488.873	545.532	381.389	14.5

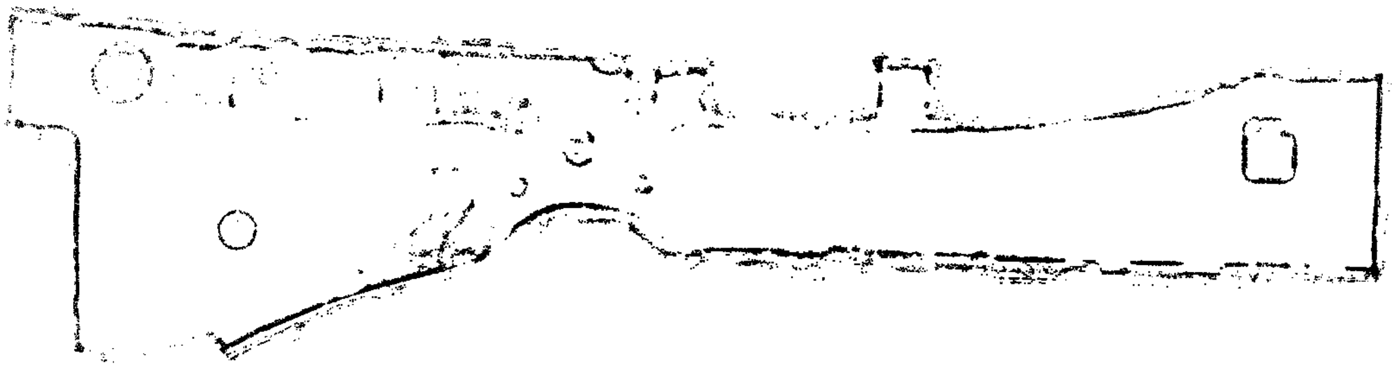
Decimation



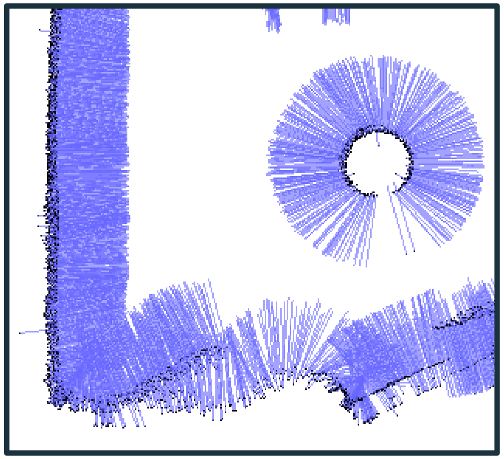


Plane Segmentation

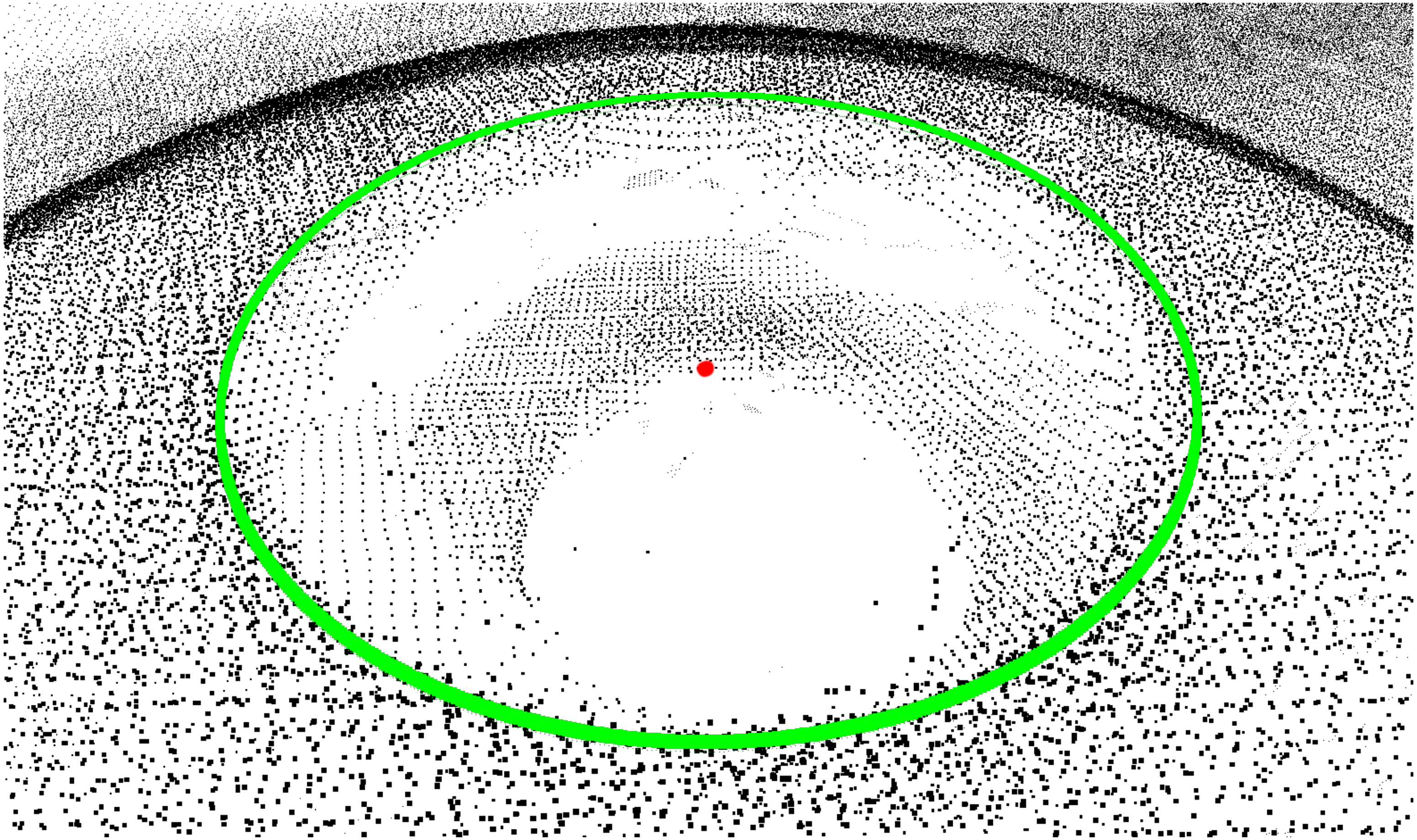
Edge Detection



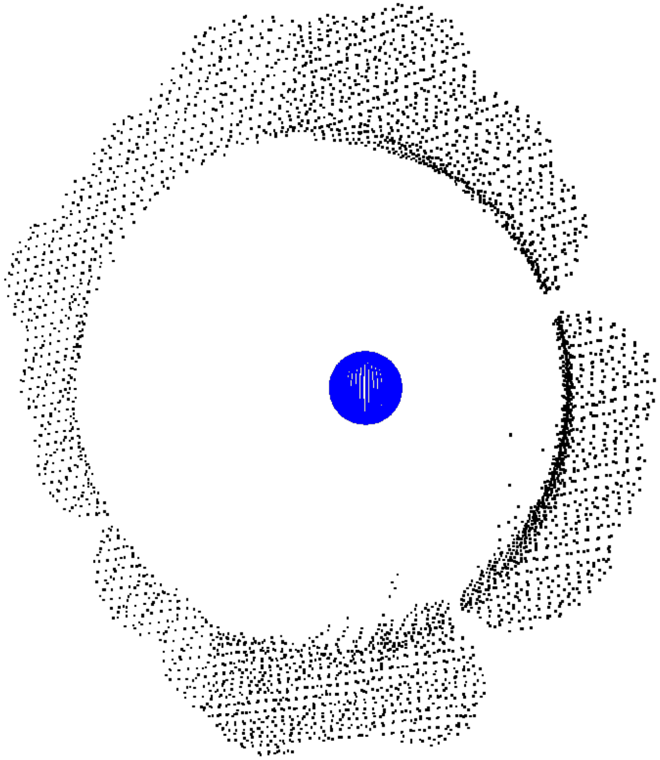
Edge Clustering



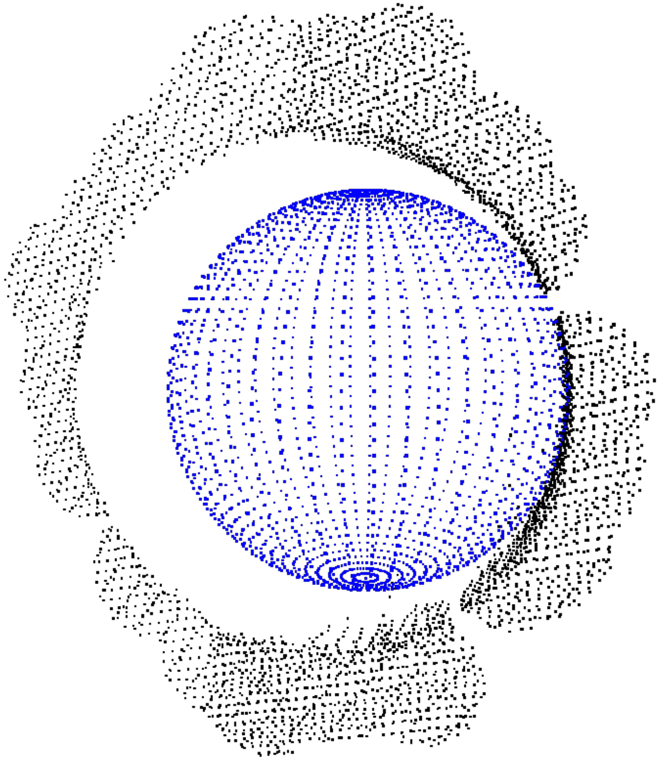
Coarse Detection



Fine Extraction

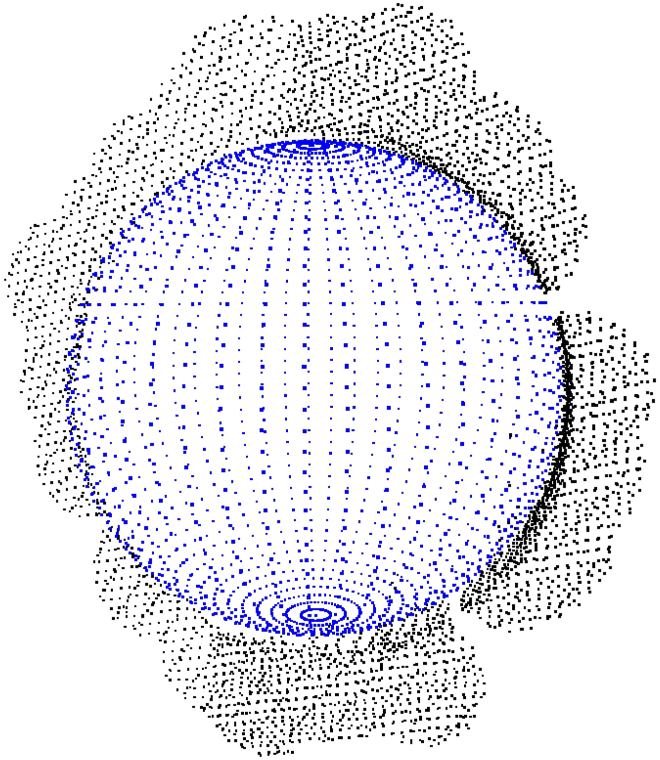


Fine Extraction

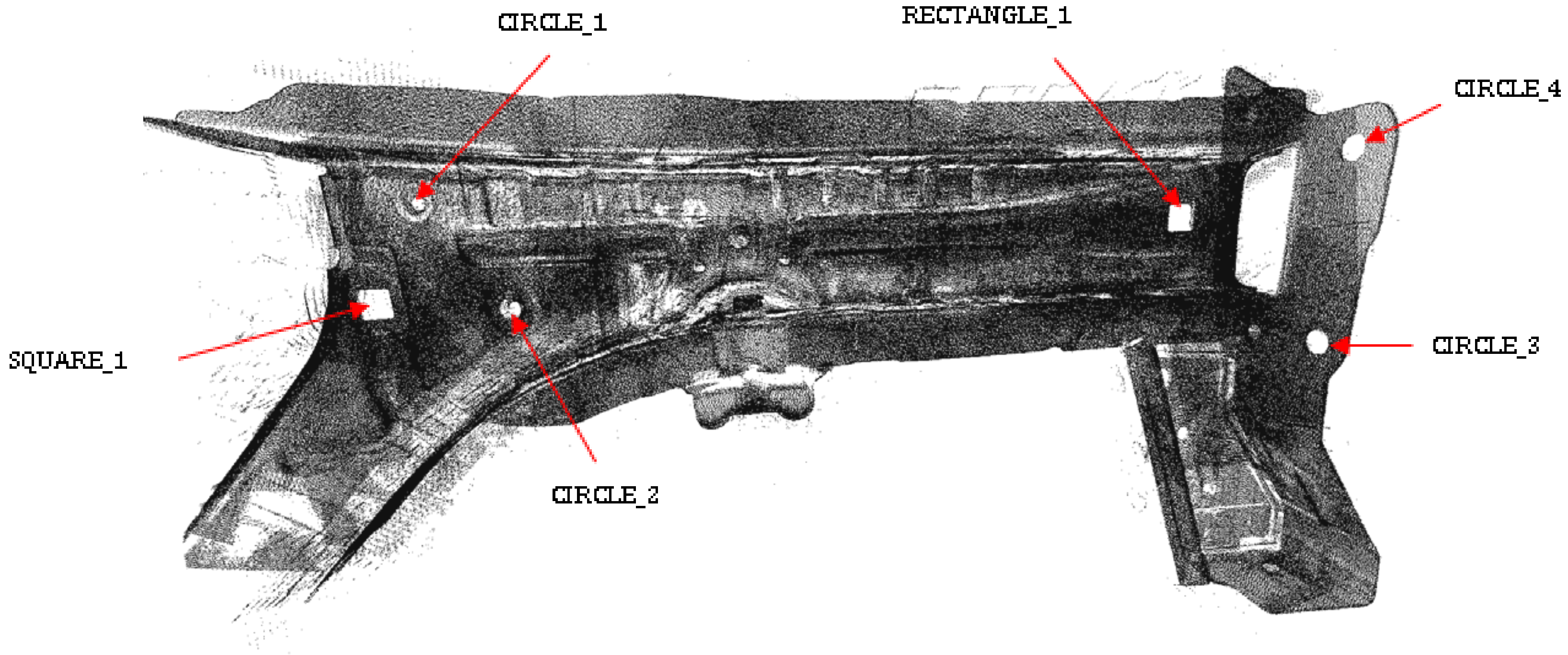


Development

Fine Extraction



Results



Deep Learning with sets and Point Clouds

PointNet: **Neural network that directly consumes point clouds**

- respects the permutation invariance of points
- each point is processed identically and independently, represented by just its three coordinates (x, y, z)
- additional dimensions added: normals and local/global features
- provides a unified architecture for applications: object classification, part segmentation, scene semantic parsing

stanford.edu/~rqi/pointnet

[Qi 2016] C.R. Qi et al, **PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation**, *arXiv:1612.00593, 2016*

[Qi 2017] C.R. Qi et al, **PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space**, arXiv:1706.02413 2017

Motto and fundamental questions

Motto:

Computer Vision = "Teaching computers to see"

***Antonio Torralba** (MIT)*

Talk@CVC 20th Anniversary, Barcelona July 9th, 2015

Exciting time for CV: new architectures, DBs, productivization, future

*Fundamental problems: **reconstruct 3D world**, recognize…*

Questions around this:

*Q1: Is "**projective vision**" a natural way to capture the 3D world?*

*Q2: Do we need **photometry** to get **geometry**?*

*Q3: Does **3D vision** mean the same than **3D geometry**?*

*Q4: Does 2D/3D matter for "**Teaching computers to see**"?*