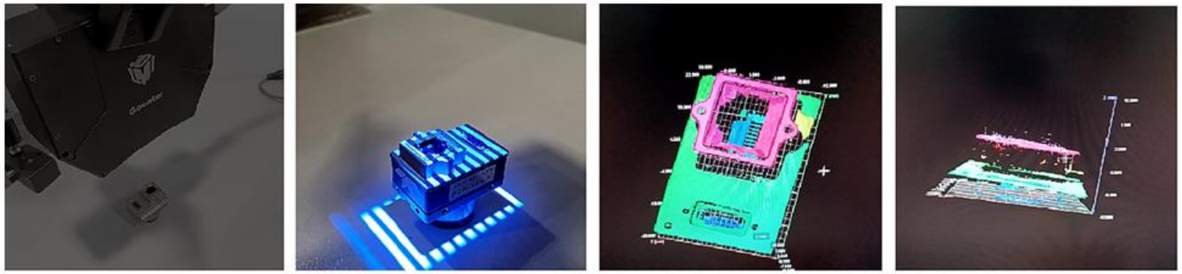
3D Vision

Paulo Dias

Infaimon training 2022:

A visão artificial tem avançado consideravelmente nos últimos anos, ao nível dos objetivos, das possibilidades e das ferramentas. Os **sistemas de visão 3D** deram um salto na qualidade e abriram um leque de possibilidades para melhorar os procedimentos e as aplicações.

Vejamos um exemplo de Inspeção de pinos com recurso à técnica 3D:



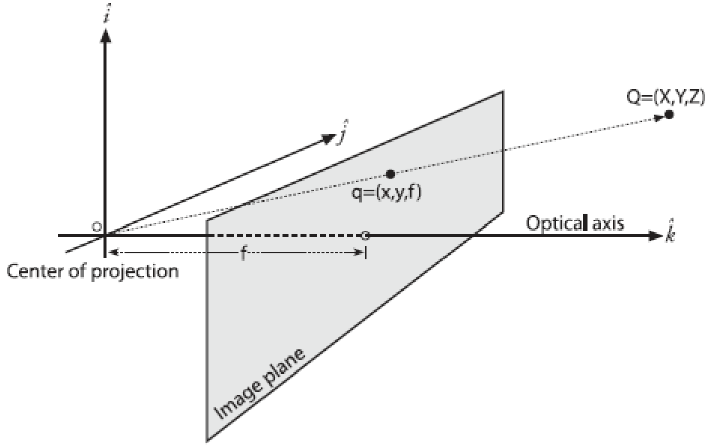
A visão 3D permite realizar inspeções independentemente das condições de luz e da cor do objeto: uma opção flexível e multitecnológica que pode ser usada numa série de produtos numa fábrica, por exemplo. Com recurso a tecnologias de reconstrução 3D, é possível **resolver problemas muito comuns na visão artificial 2D,** como a análise de superfícies de baixo contraste ou com muito brilho.

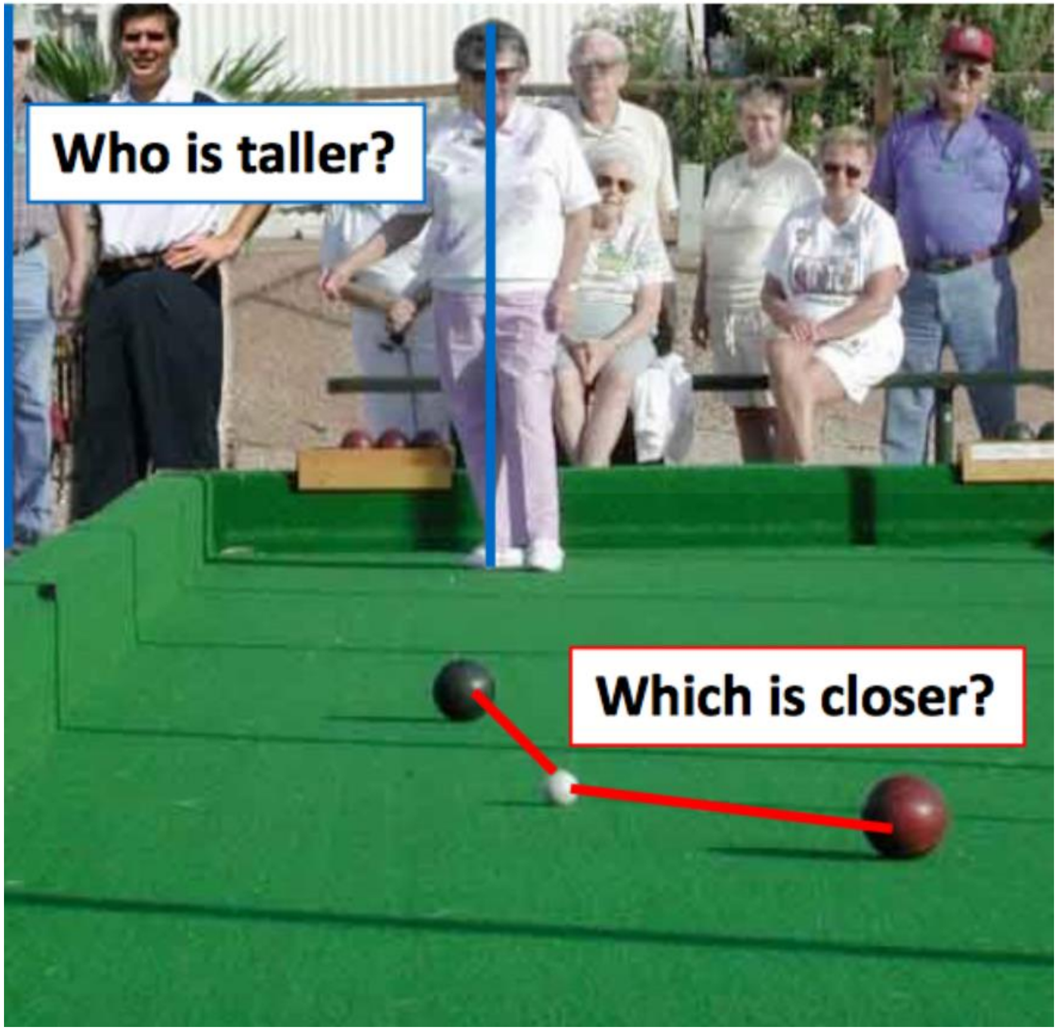
Summary

- Methods for 3D data acquisition
  - Introduction
  - Passive
    - shape from X (stereo, motion, shading, focus)
  - Active range sensing
    - Structured Light Systems
    - Laser Range Finder – Time of Flight
- 3D vision applications
- Manipulation of range/depth images
    - Edges
    - Triangulation
    - Registration
    - Texture

Introduction

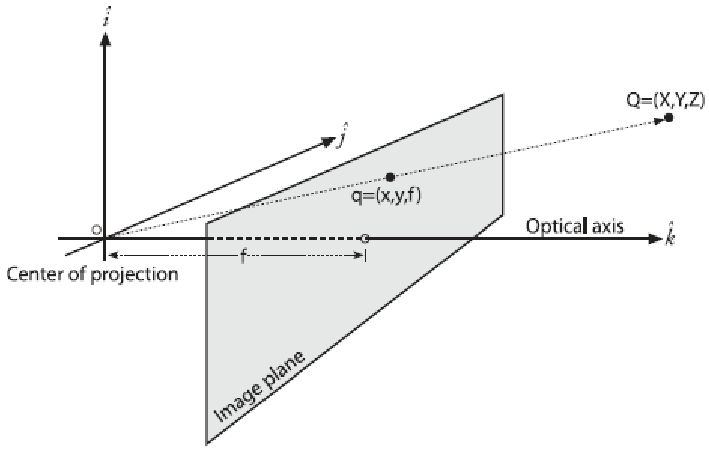
- Depth estimation is difficult when real world scenario are projected from 3D to 2D in camera images.



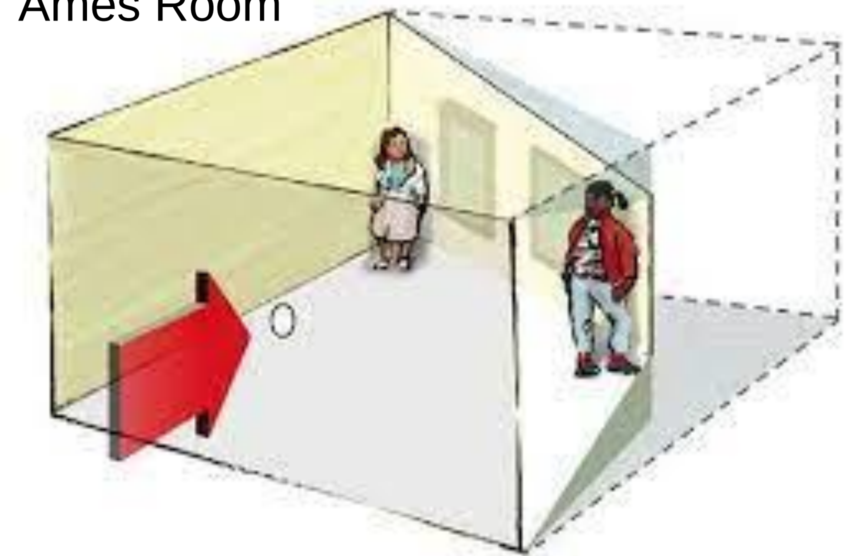


Introduction

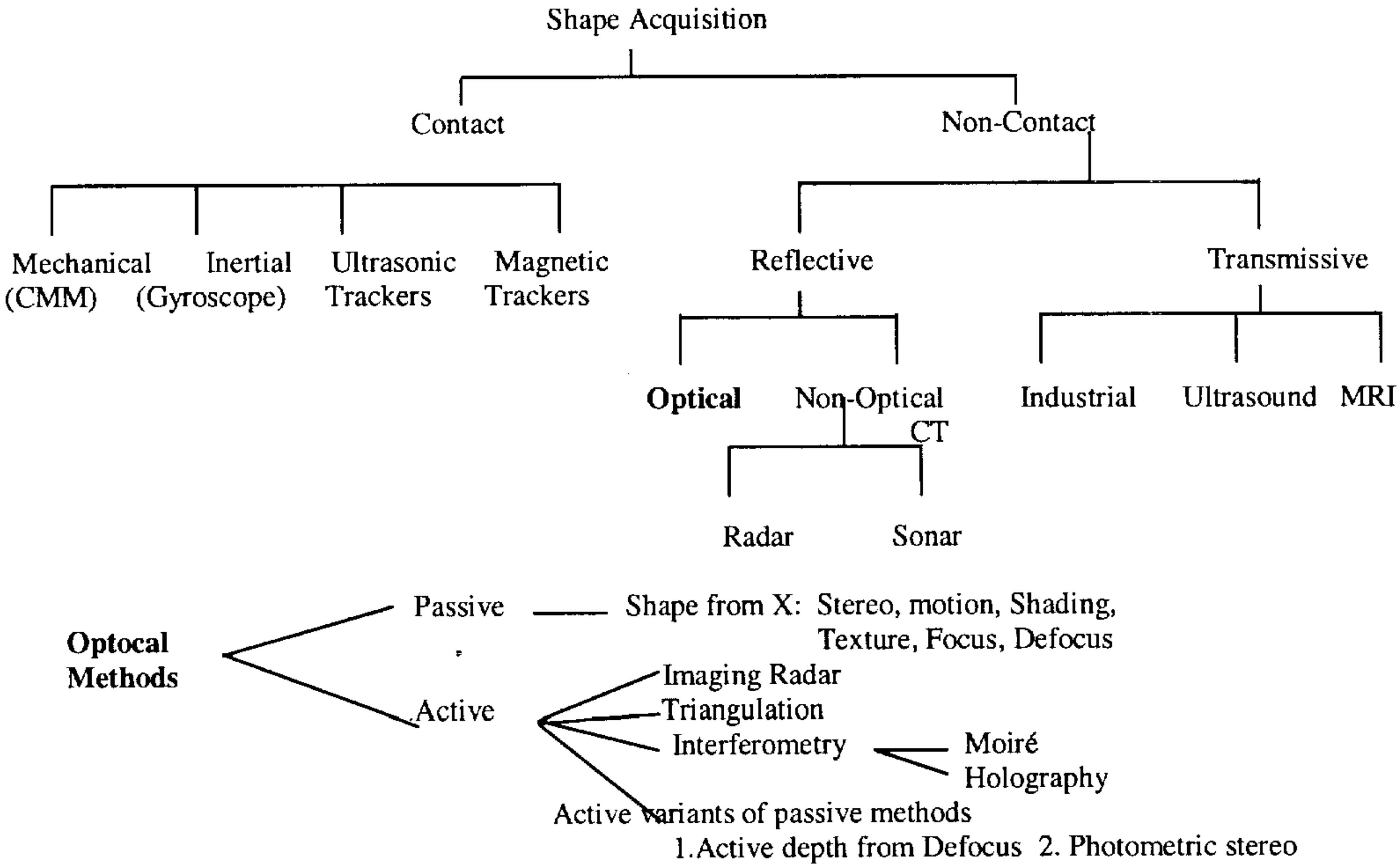
- Depth estimation is difficult when real world scenario are projected from 3D to 2D in camera images.



Ames Room



Overview



[Mada2003]

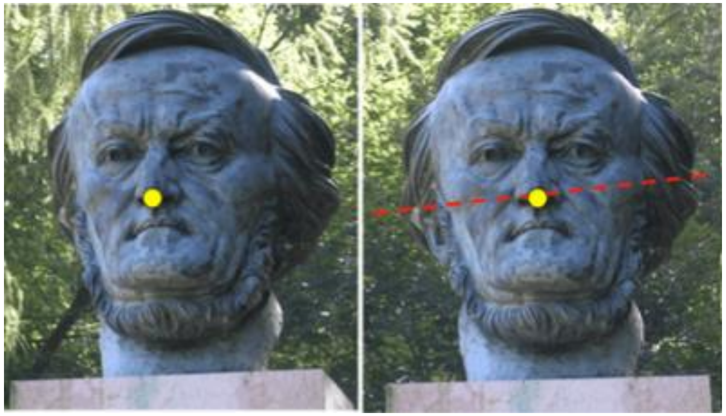
Summary

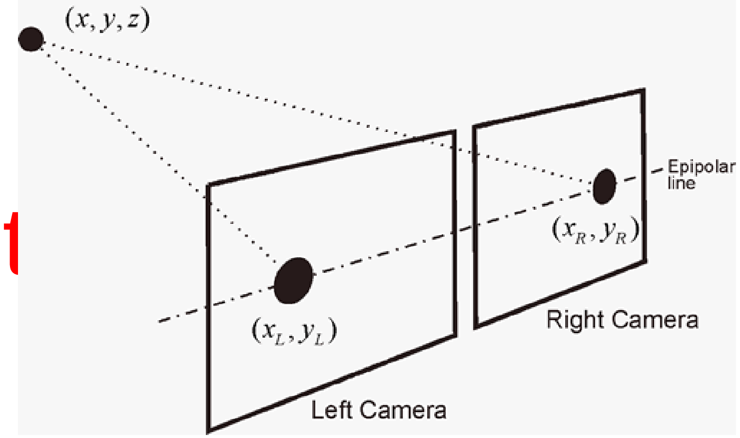
- Methods for 3D data acquisition
  - Introduction
  - Passive
    - shape from X (stereo, motion, shading, focus)
  - Active range sensing
    - Structured Light Systems
    - Laser Range Finder – Time of Flight
- 3D vision applications
- Manipulation of range/depth images
    - Edges
    - Triangulation
    - Registration
    - Texture

Passive - Shape from stereo

- See last lecture

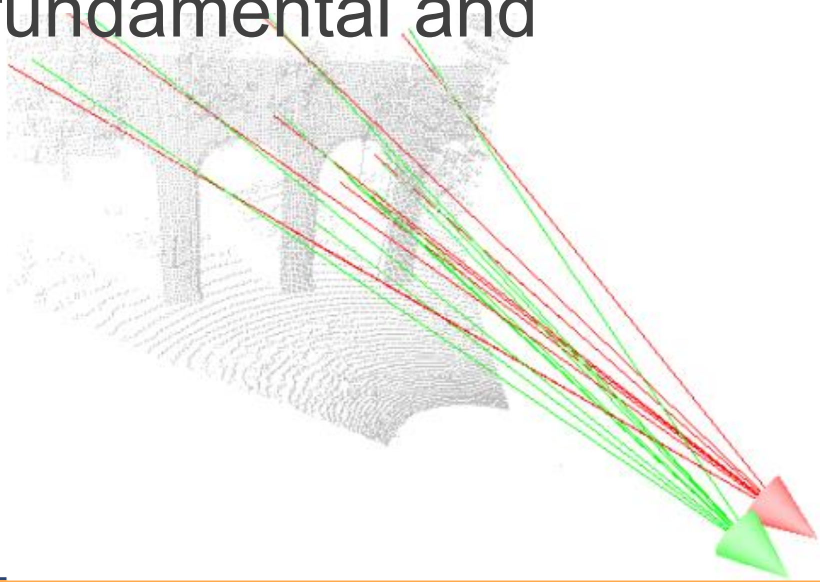
- Pros
  - Cheap (use cameras)
  - Fast acquisition
- Cons
  - Highly dependant on correspondences qualit
  - Still challenging



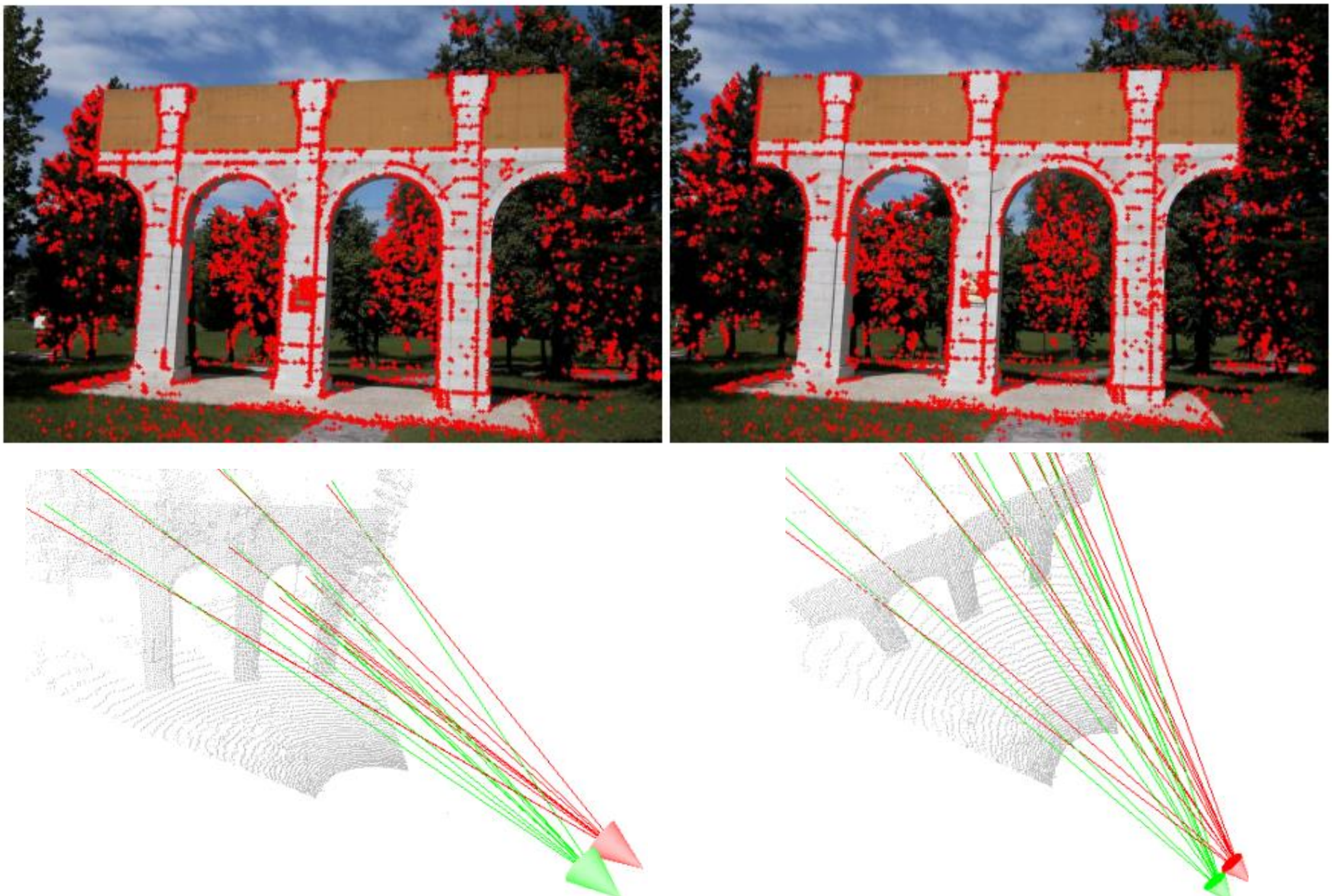


Passive - Shape from motion

- Shape from motion
  - Similar to stereovision in many ways
  - Successive images might be considered as stereo pairs
  - With texture, possible to find correspondences (matching techniques, optical flow…) and find fundamental and essential matrix.

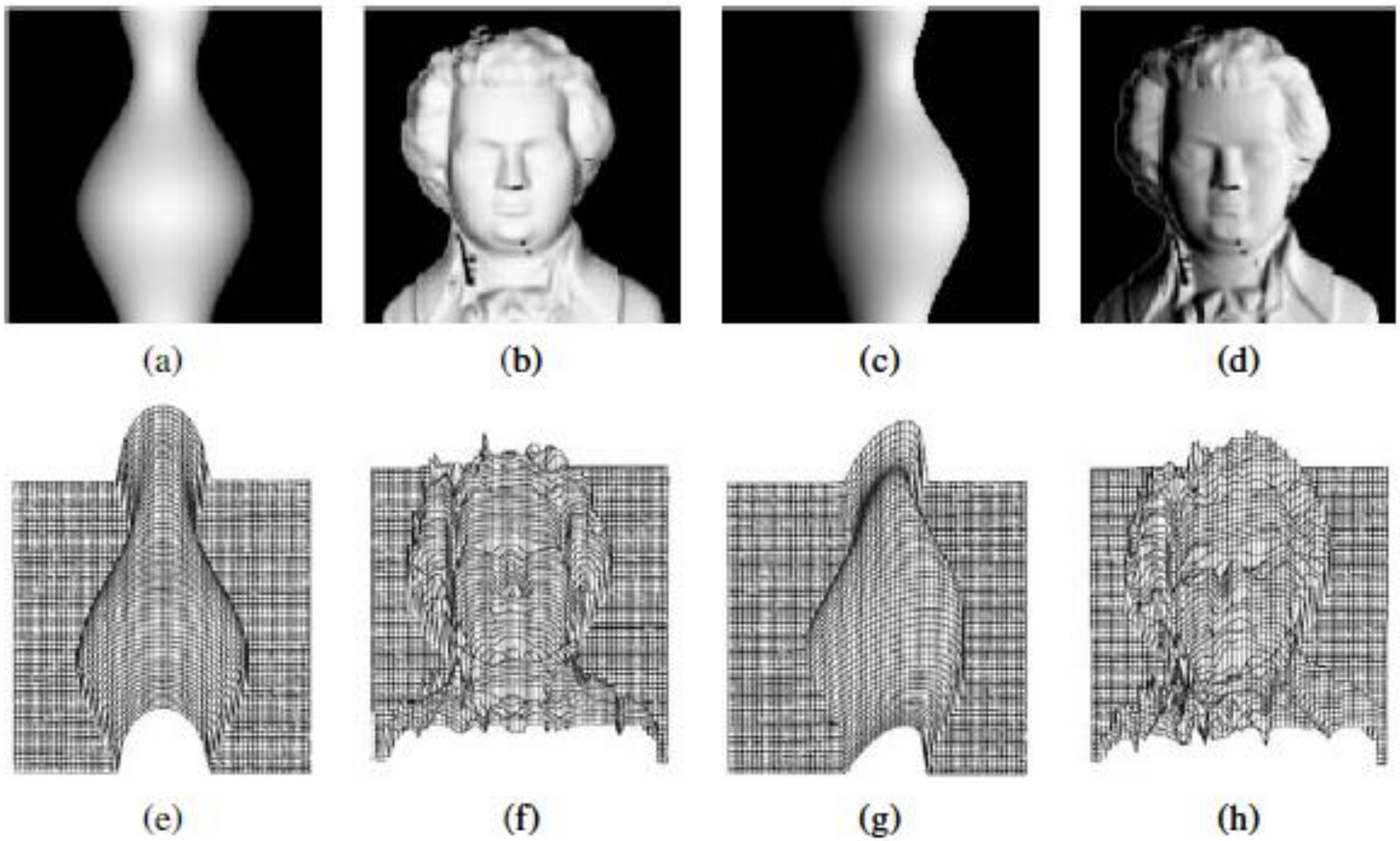


Passive - Shape from motion



- Shape from Shading
  - Given a continuous surface, and known illumination, intensity variation in the surfaces depends of its orientation.
  - Most surfaces are not uniform and lighting difficult to control - normally combined with other methods.

Passive - Shape from shading



Depth Map and 3D Imaging Applications: Algorithms and Technologies
IGI Global Editors: Aamir Saeed Malik, Tae-Sun Choi, Humaira Nisar
Three-Dimensional Scene Reconstruction: A Review of Approaches

Passive - Shape from focus

- Shape from focus
  - Objects away from focal plan are out of focus.
  - With several images with different focus, possible to extract depth information.

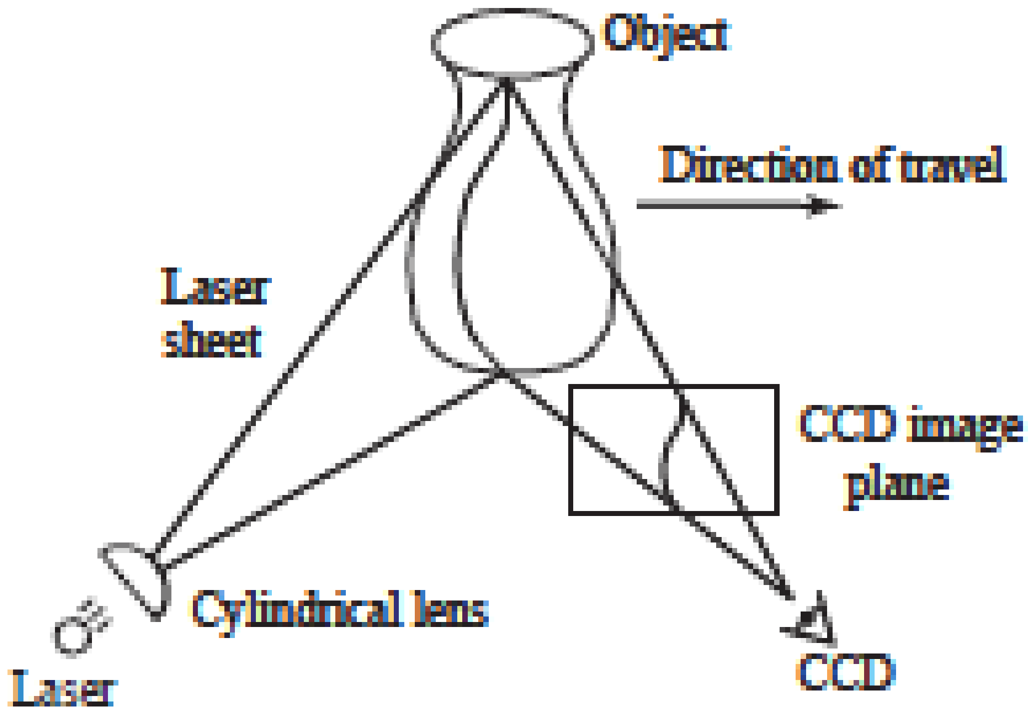
Favari and Soatto: A Geometric Approach to Shape from Defocus

Summary

- Methods for 3D data acquisition
  - Introduction
  - Passive
    - shape from X (stereo, motion, shading, focus)
  - Active range sensing
    - Structured Light Systems
    - Laser Range Finder – Time of Flight
- 3D vision applications
- Manipulation of range/depth images
    - Edges
    - Triangulation
    - Registration
    - Texture

Active – Structured Light Techniques

- Projection of a known pattern
- Acquisition with camera, 3D from pattern deformation in scene.

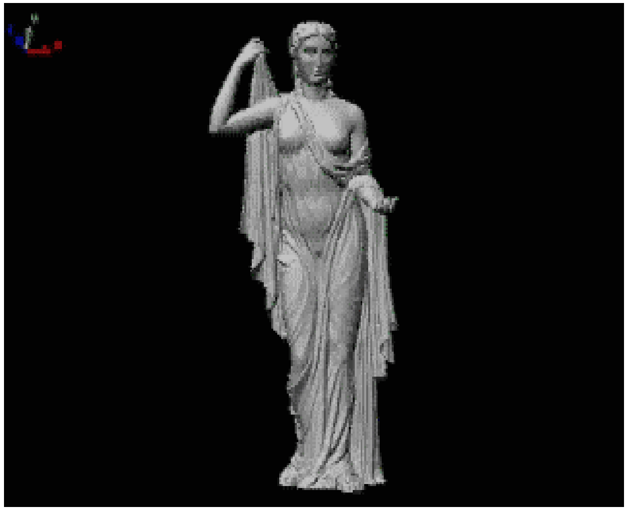


Active – Structured Light Techniques

- Several commercial for small distances

Shape Grabber

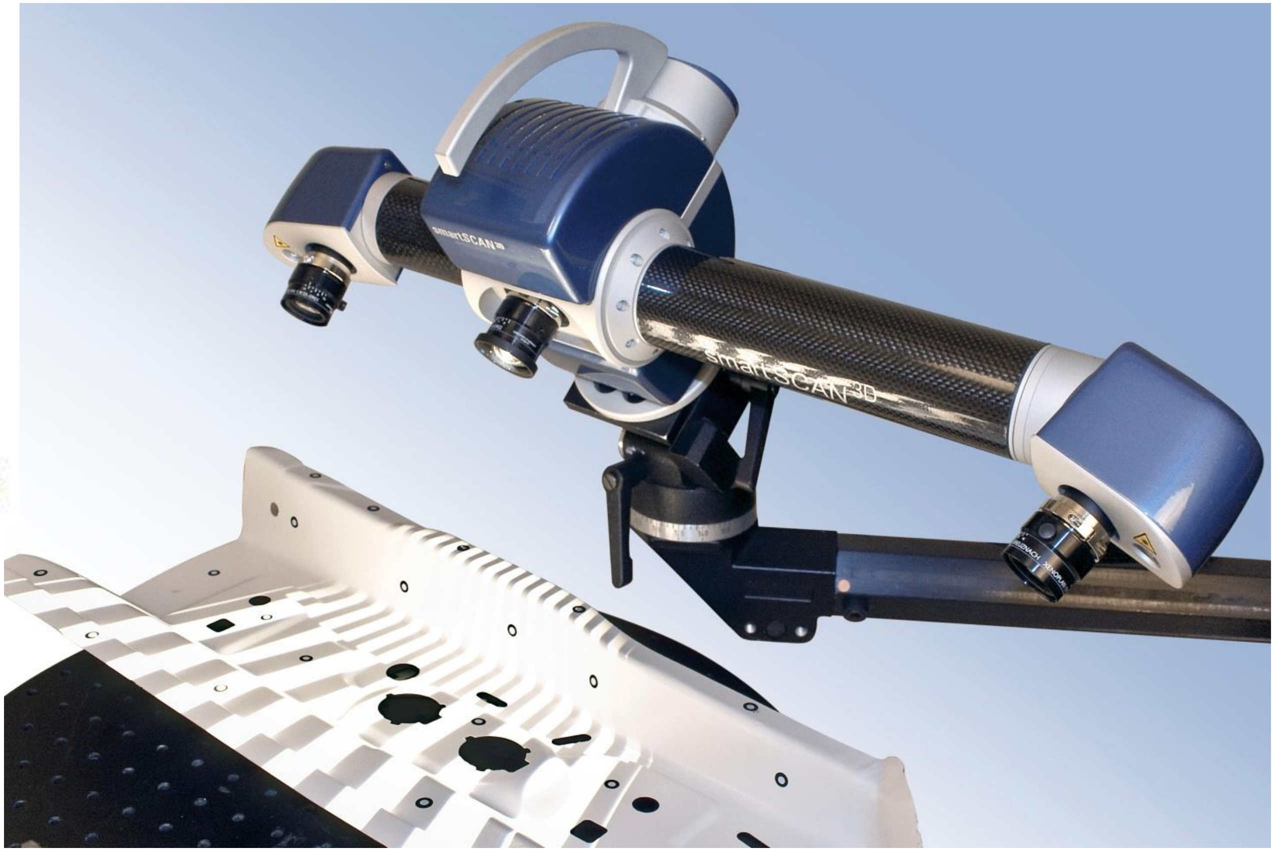
Minolta Vivid



Marble Statue of Aphrodite scanned with the VIVID 910 using the rotary stage option

Active – Structured Light Techniques

- Commercial solutions



SmartScan Breuckman

Skull with 1.5 Million points – Error below 30 µm

Active – Structured Light Techniques

- Pros
  - Very accurate
- Cons
  - Takes time (often need to scan through an area)
  - Sensitive to environment brightness, usually only implemented in dark or indoor areas.
  - Short range

Active - Laser Range Finder – Long Range

- For larger areas (buildings, rooms) use of Laser Range Finders - LRF.

Riegl

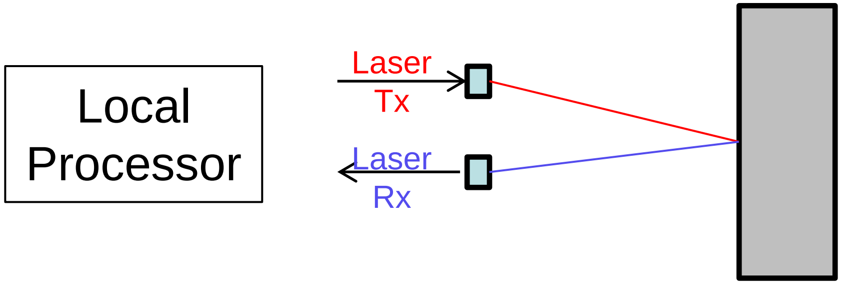
Zoller & Frölich

Cyra

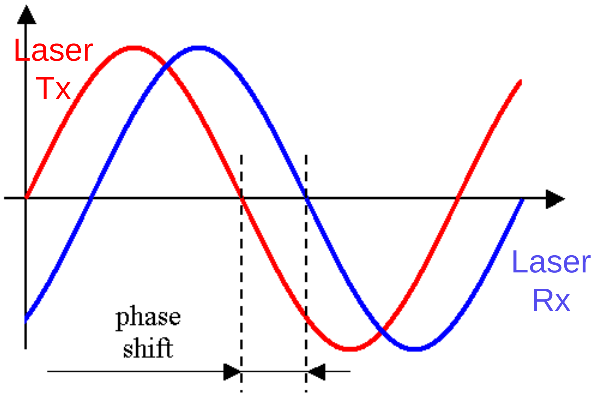
Leica BLK 360º
Available @ IEETA

Active - Laser Range Finder – Long Range

- Working principle:
  - Light Pulse Time of Flight.



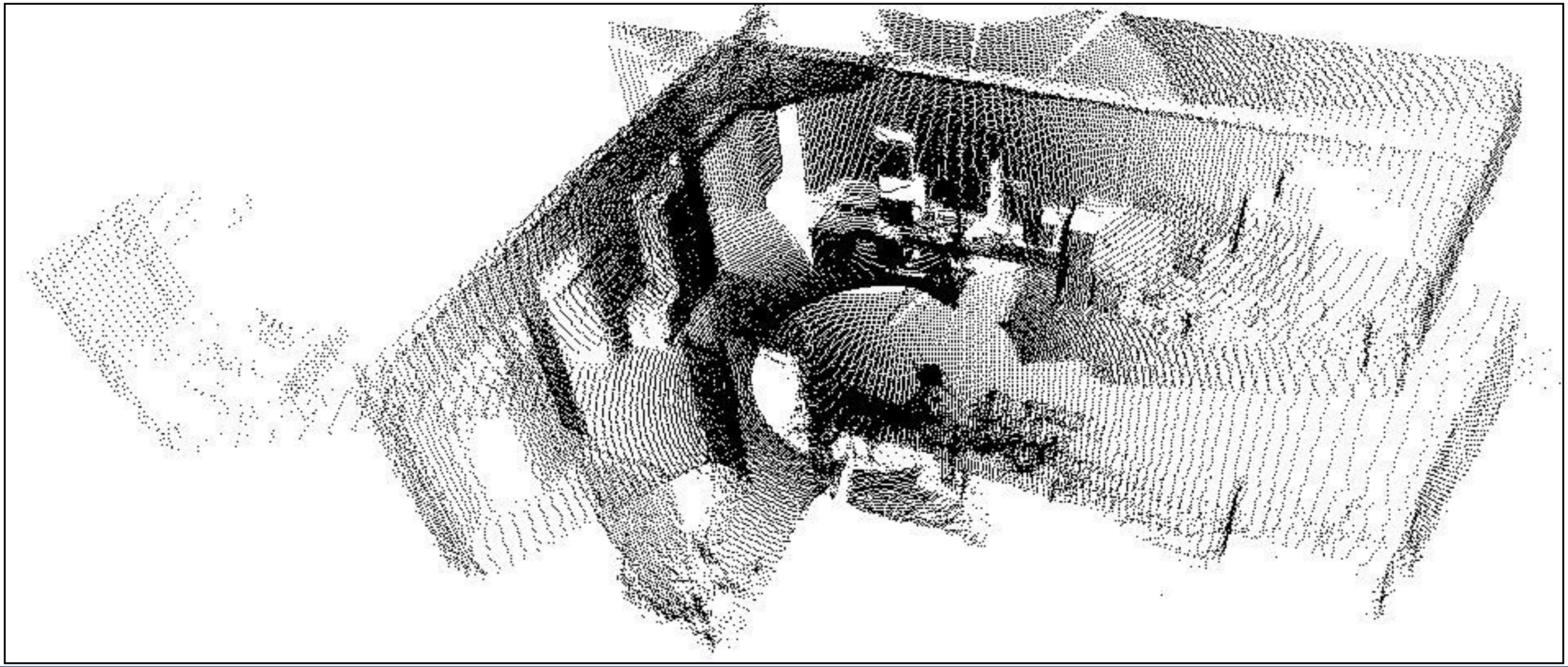
  - Phase Shift: Amplitude of frequency modulation – Comparison of phases.



Range Image example

Reflectance
1000 x 175 (Riegl LMS Z210)

3D Cloud of points



Range Image example

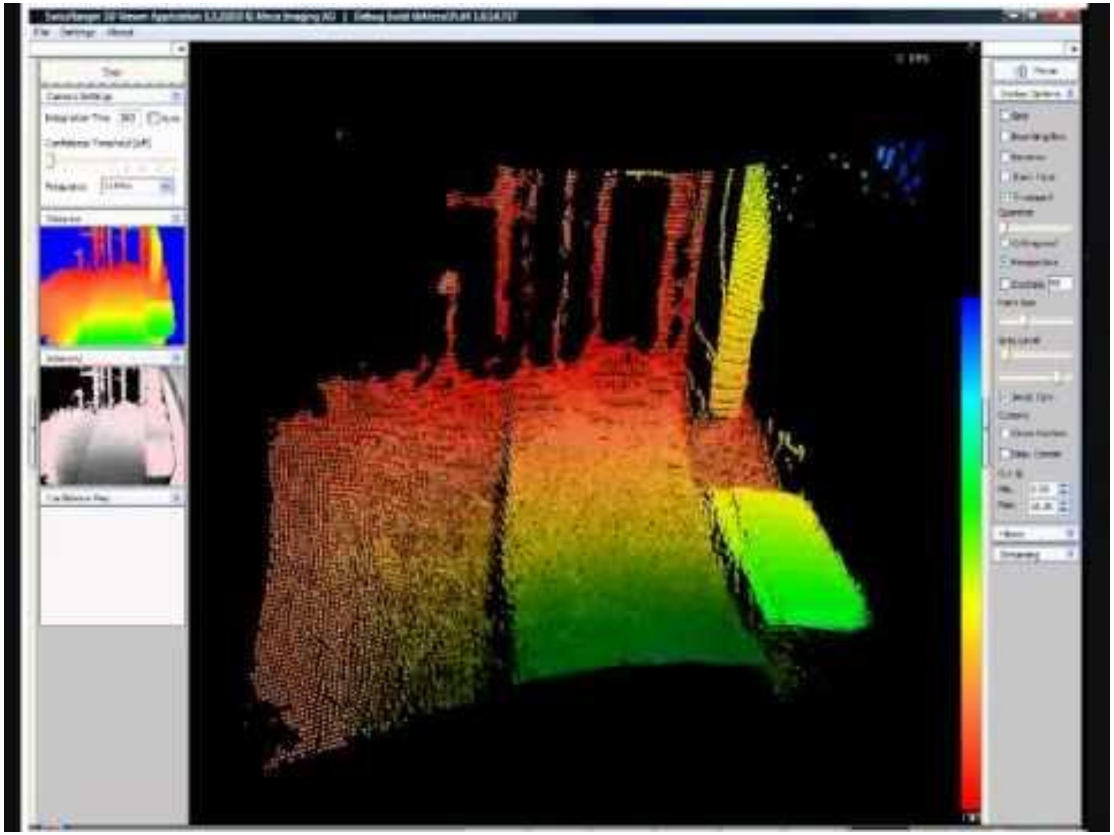
- deti lobby – LEICA BLK 360º
  - 6 scans
  - 70,009,139 points
  - Digital Photographs
  - Infra red sensor

Active – Laser Range Finder

- Pros
  - Independent from external lighting
  - No need of texture in scene
  - Provide directly 3D measurements
- Cons
  - Expensive sensors
  - Large sensors = acquisition more difficult
  - Limited spatial resolution
  - No colour texture map, or black and white reflectance

Active – 3D ToF Cameras

- Phase shift principle of emitted and received infrared light to measure depth



Swiss Ranger SR4000
3D ToF Camera

Resolution: 176x144
Range: 5–8m
54 fps

Comparing methods

- Overall performance
  - **Structured Light**:
    - Best depth accuracy performance
    - Shortest range
    - Require Dark environment
  - **Time of Flight (ToF):**
    - Up to hundred meters depending on emitting power
  - **Camera Array:**
    - Largest depth error
    - Range depend on baseline (distance between cameras) typically around 10m
    - Need bright environment

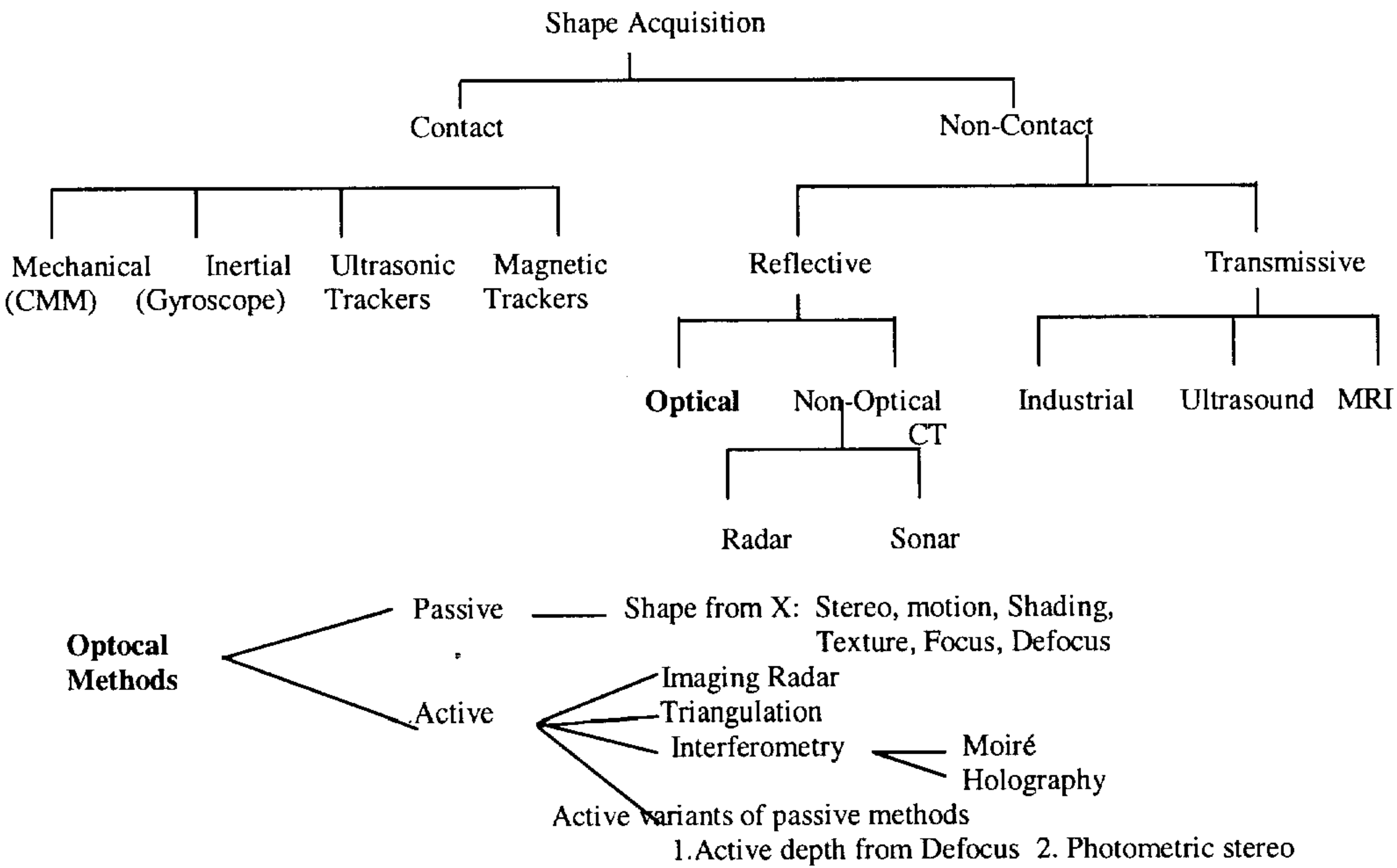
Comparing methods

- Cost
  - **Structured Light**:
    - Highest cost
  - **Time of Flight (ToF):**
    - Moderate cost – might decrease significantly
  - **Camera Array:**
    - Lowest cost
    - Development mainly on software side

Active vs Passive in a nutshell

	Range (ToF)	Intensity (Camera arrays)
Cost	Expensive sensors but decreasing	Low cost any digital camera
Acquisition	Often difficult with large sensors	Easy, with a digital camera
Depth error	Intermediate depending on sensor	Typically largest depth error, despite high resolution degrades with non-ideal point matching
Texture map	No colour texture map, or black and white reflectance	Possibility to provide a realistic colour texture map
Lighting	Independent from external lighting	Highly dependent on lighting conditions
Texture relevance	No need of texture in scene	Texture is crucial for good results
3D processing	Provide directly 3D measurements	Need processing to extract 3D depth from images

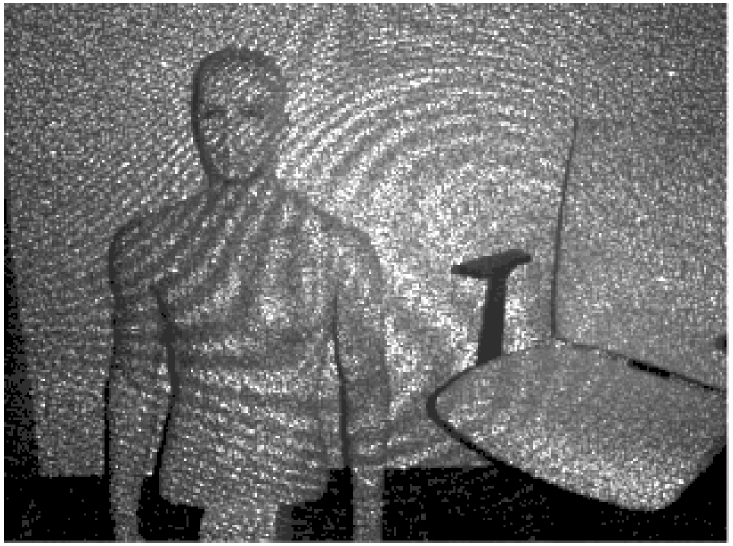
Kinect?

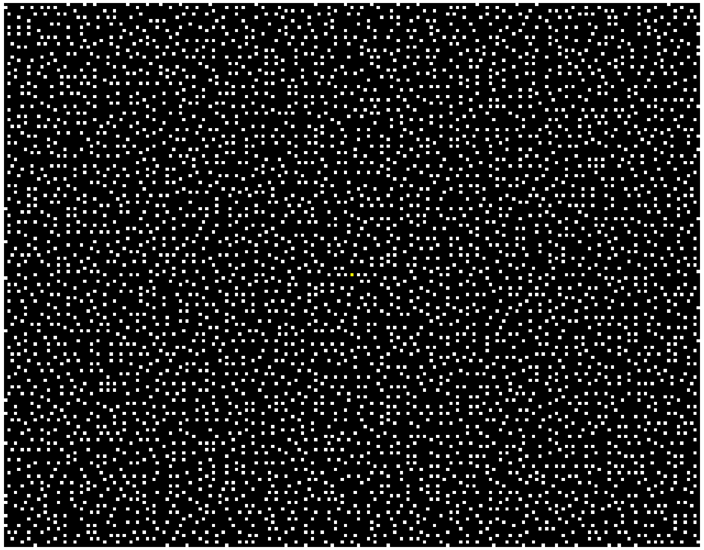


[Mada2003]

E o Kinect?

- Active – Infrared pattern





Summary

- Methods for 3D data acquisition
  - Introduction
  - Passive
    - shape from X (stereo, motion, shading, focus)
  - Active range sensing
    - Structured Light Systems
    - Laser Range Finder – Time of Flight
- 3D vision applications
- Manipulation of range/depth images
    - Edges
    - Triangulation
    - Registration
    - Texture

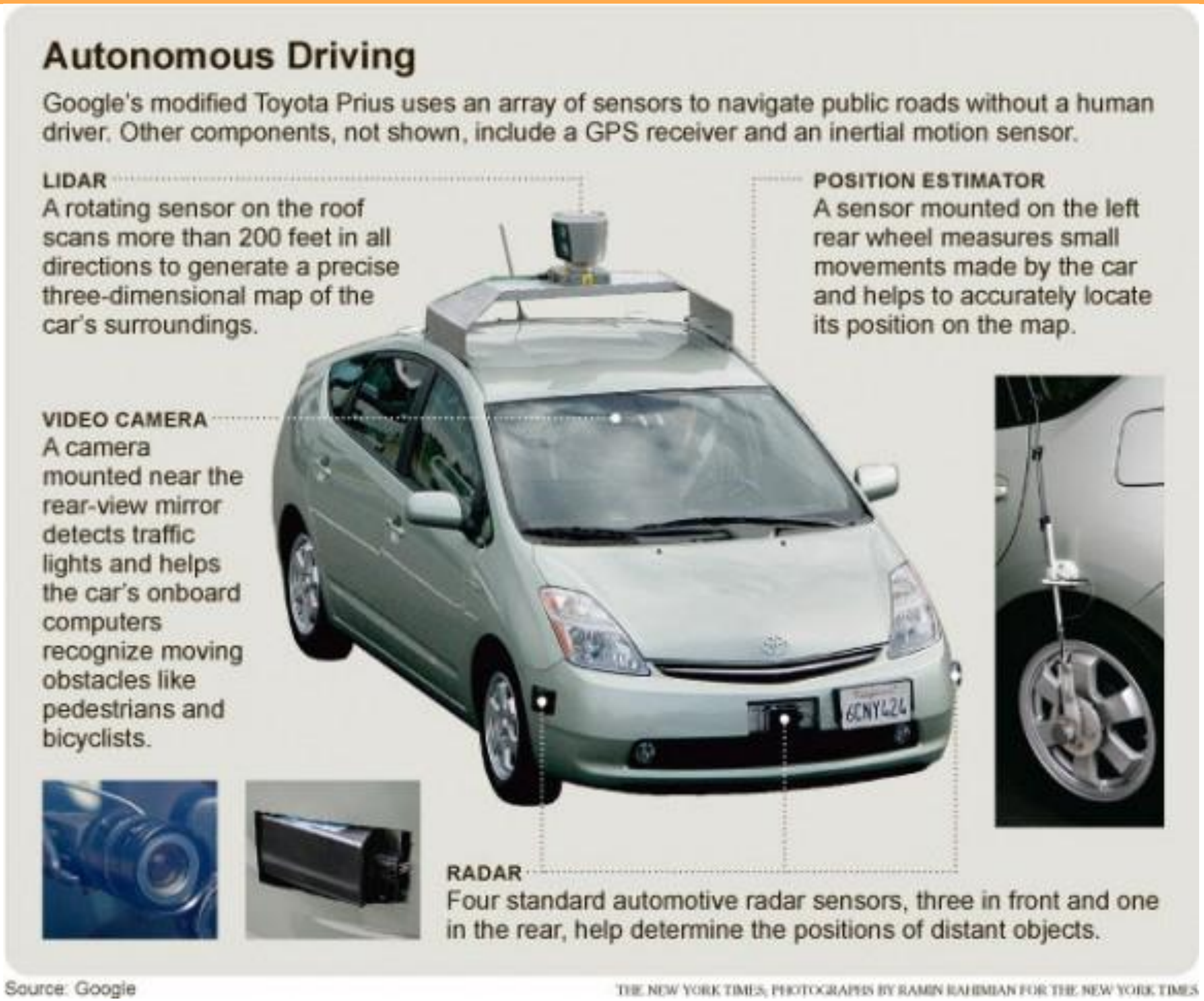
3D Vision applications

Robotics:

Navigation, localization, mapping,

avoiding collision, …

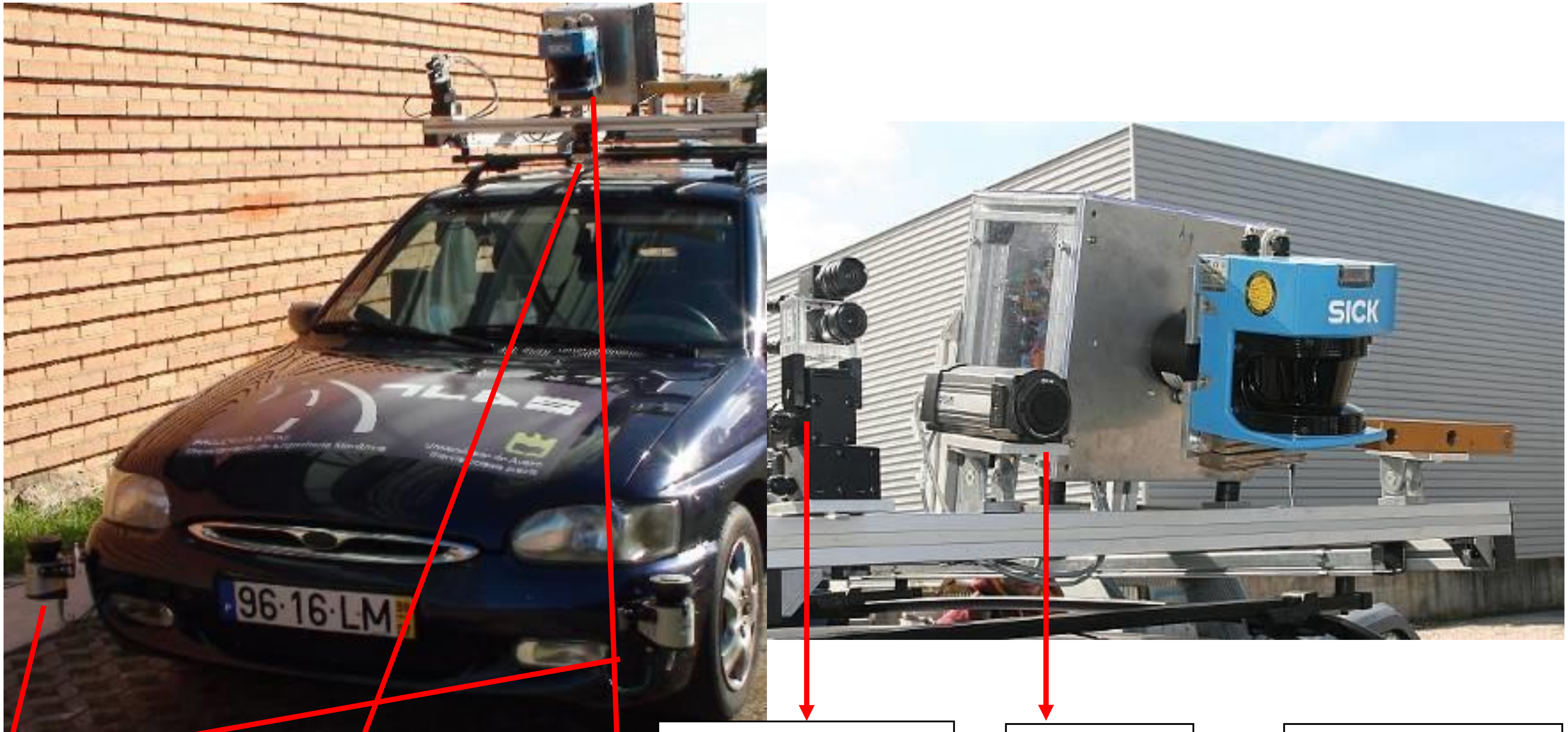
Projects: google car



Darpa Grand Challenge: Stanford

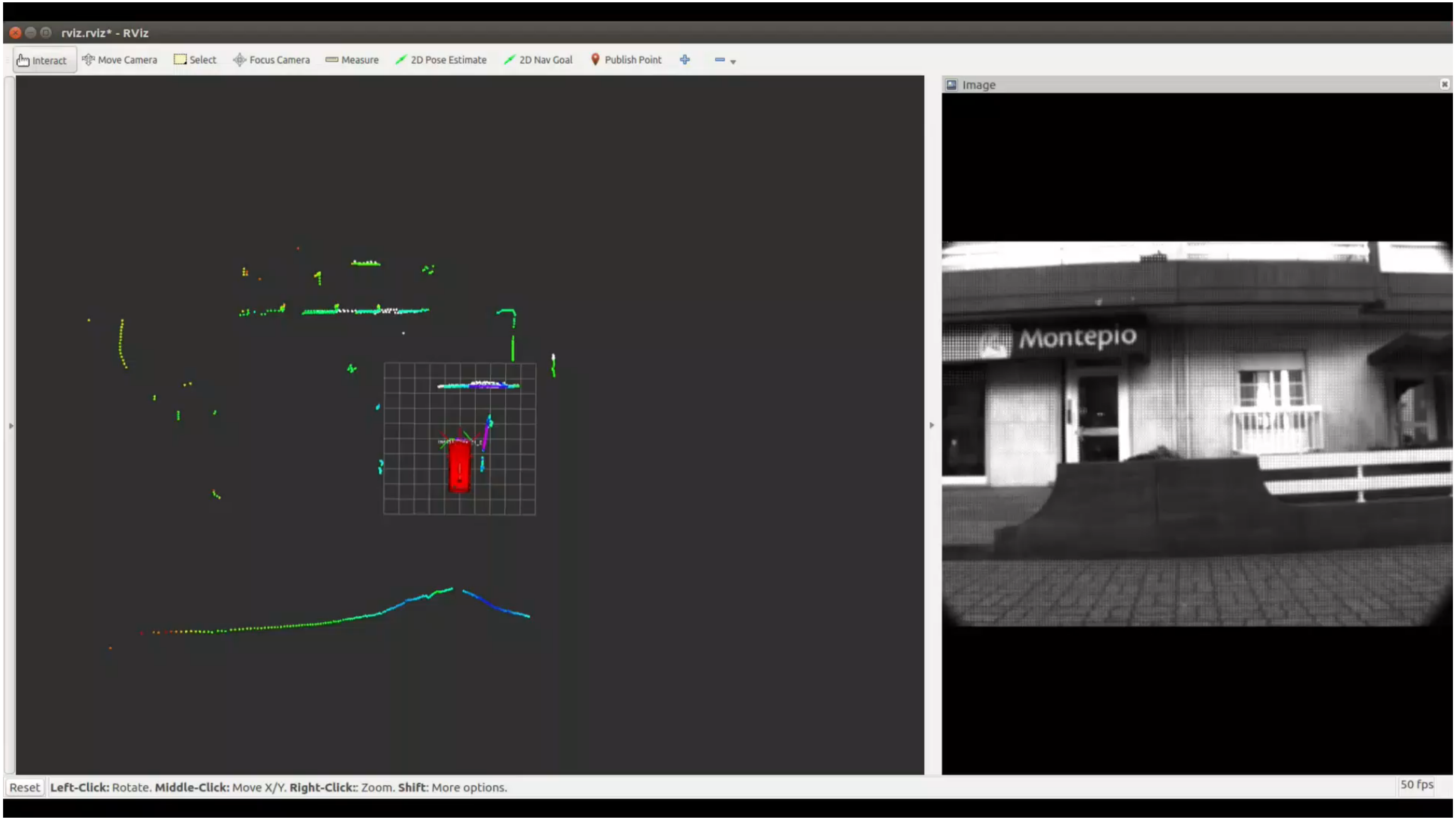
Atlas Car: Universidade de Aveiro

Atlas Car: Universidade de Aveiro



Visão / Condução Autónoma – AtlasCar v2

Visão / Condução Autónoma – AtlasCar v2

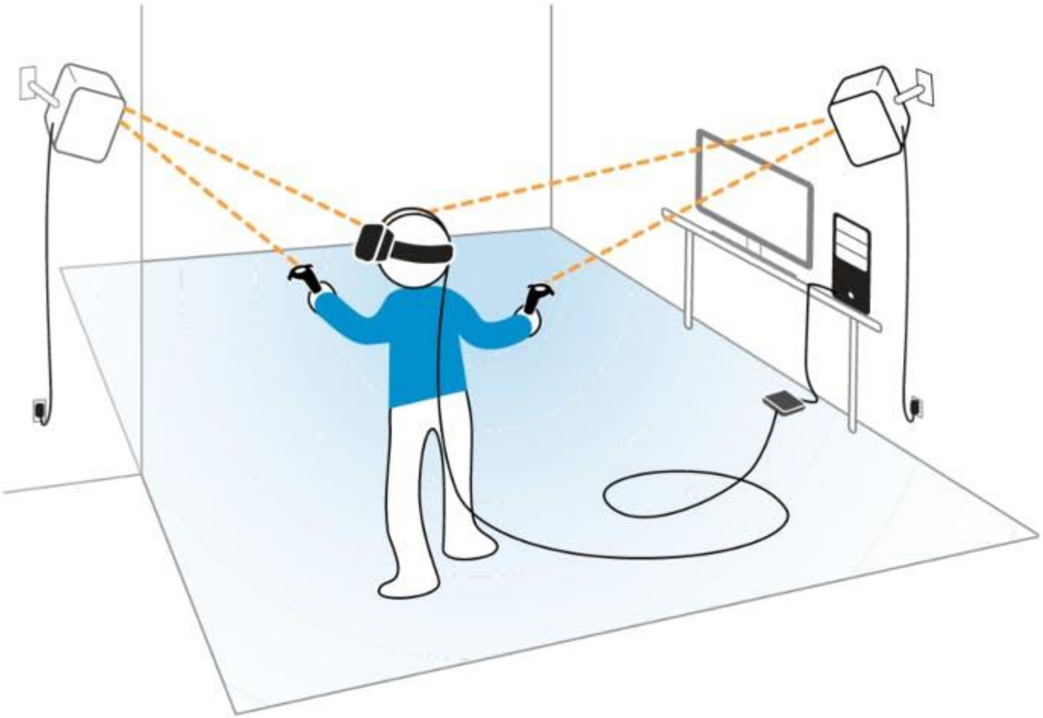


3D Vision applications

AR / VR: sensing real 3D environments and reconstructing them in the virtual world

VR/AR Interaction

- Devices must respond accurately to the 3D movement -> need high-performance depth sensor



HTC Vive lighthouse

Summary

- Methods for 3D data acquisition
  - Introduction
  - Passive
    - shape from X (stereo, motion, shading, focus)
  - Active range sensing
    - Structured Light Systems
    - Laser Range Finder – Time of Flight
- 3D vision applications
- Manipulation of range/depth images
    - Edges
    - Triangulation
    - Registration
    - Texture

Range Image

- Range image is a rectangular array of numbers that quantifies the distance from the sensor to the surfaces within the field of view.
- Also referred as depth image and easily transform to cloud of points.

Range image characteristics

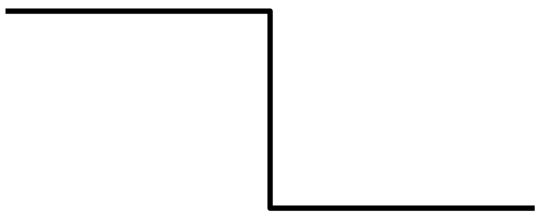
- **Edges in intensity images**

  edges related to intensity changes (due to geometry or aspect - for example colour or shadow)

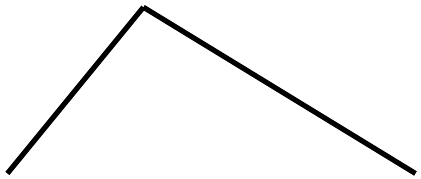
Range image characteristics

- **Edges in range images**,

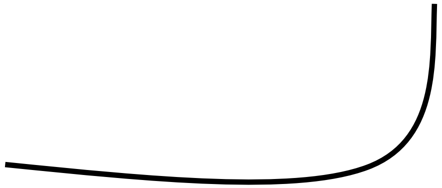
  3 different type of edges:



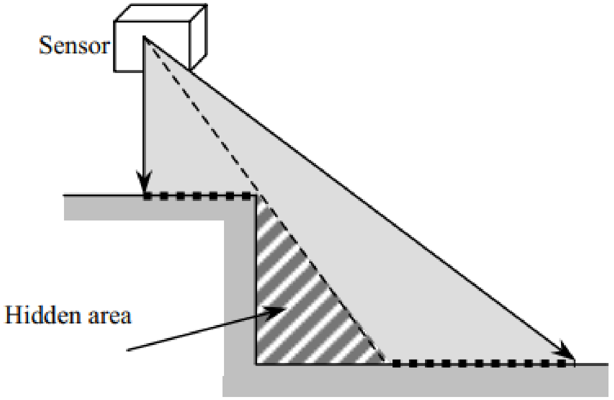
Jump or step edge



Roof or crease edge



smooth edge



Range Image Processing

Adapted Sick (Aveiro)

IEETA

Mechanical Lab.

Riegl (Italy)

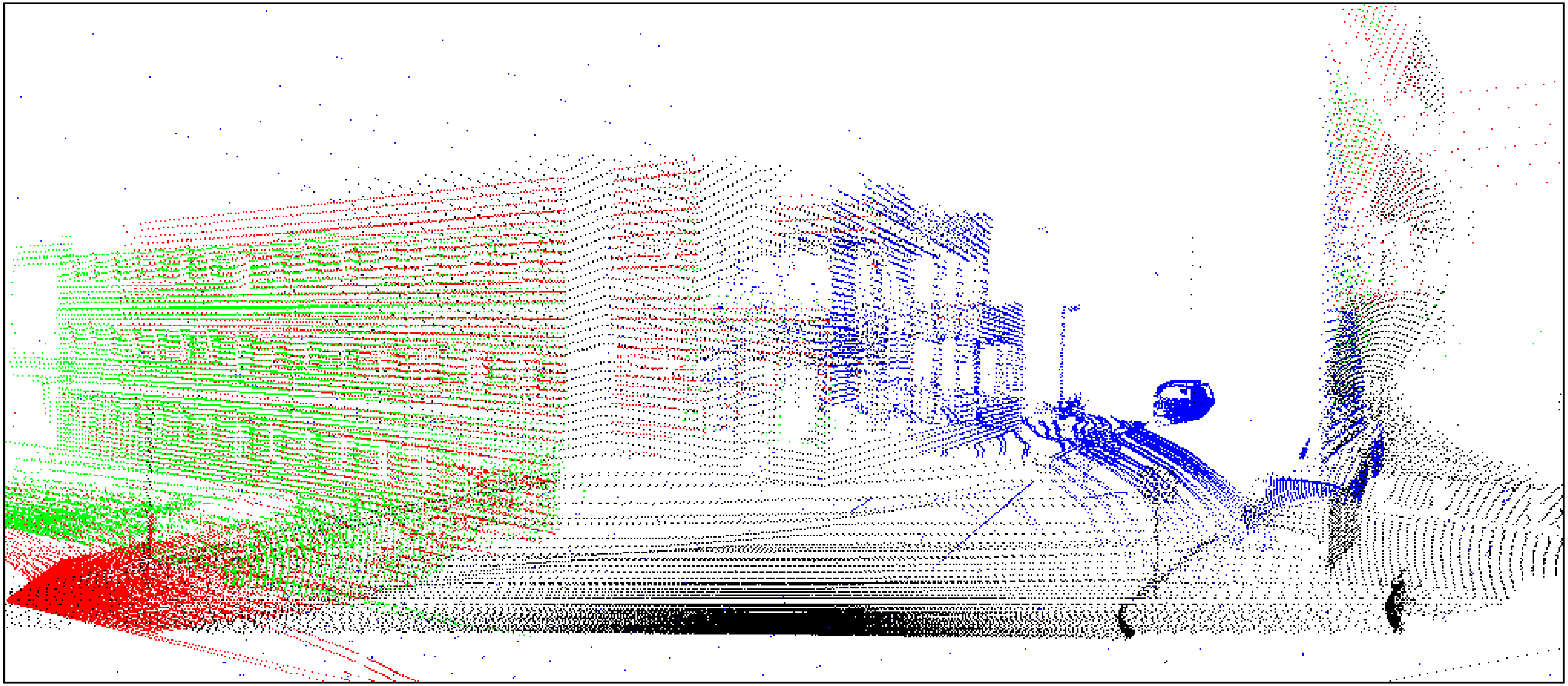
Farmhouse Laveno

Assemblée nationale

Sala dello scrutionio

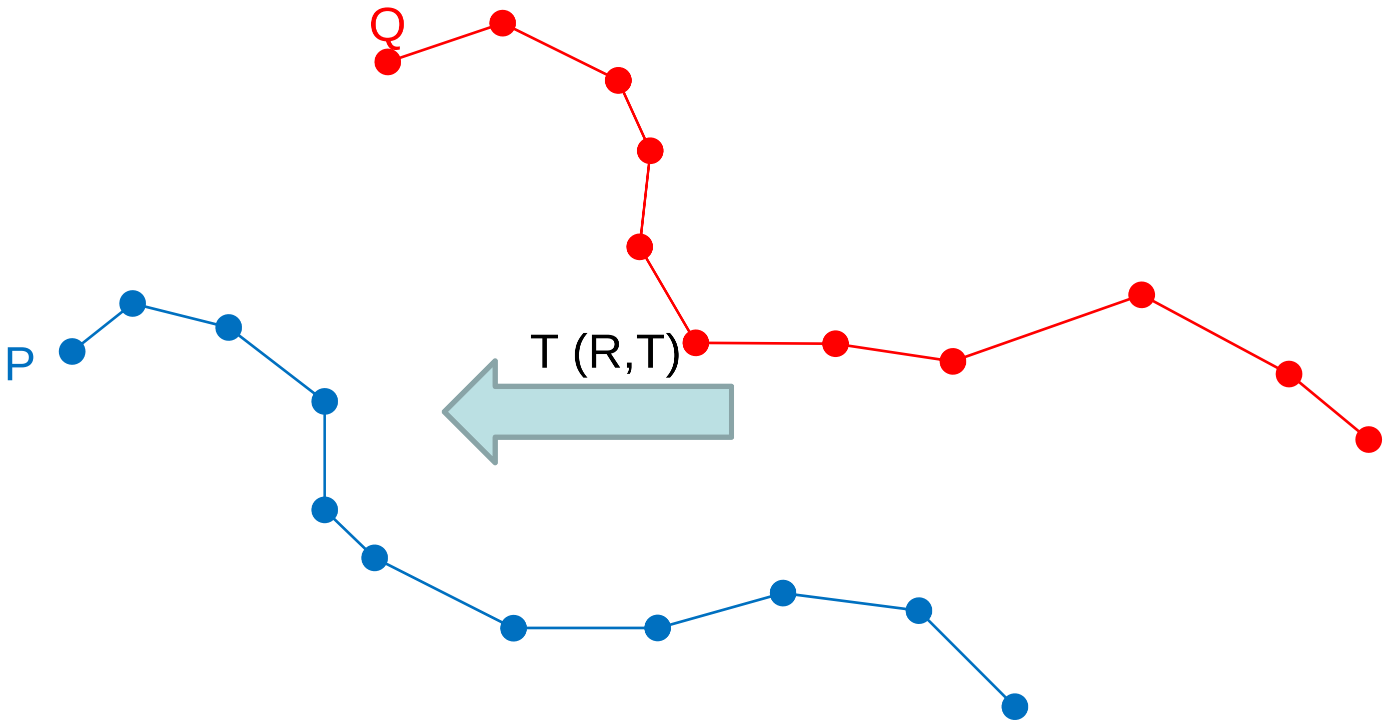
47

Registration



Registration

- Registration estimates Rigid Body Transform that minimize the distance between 2 scans

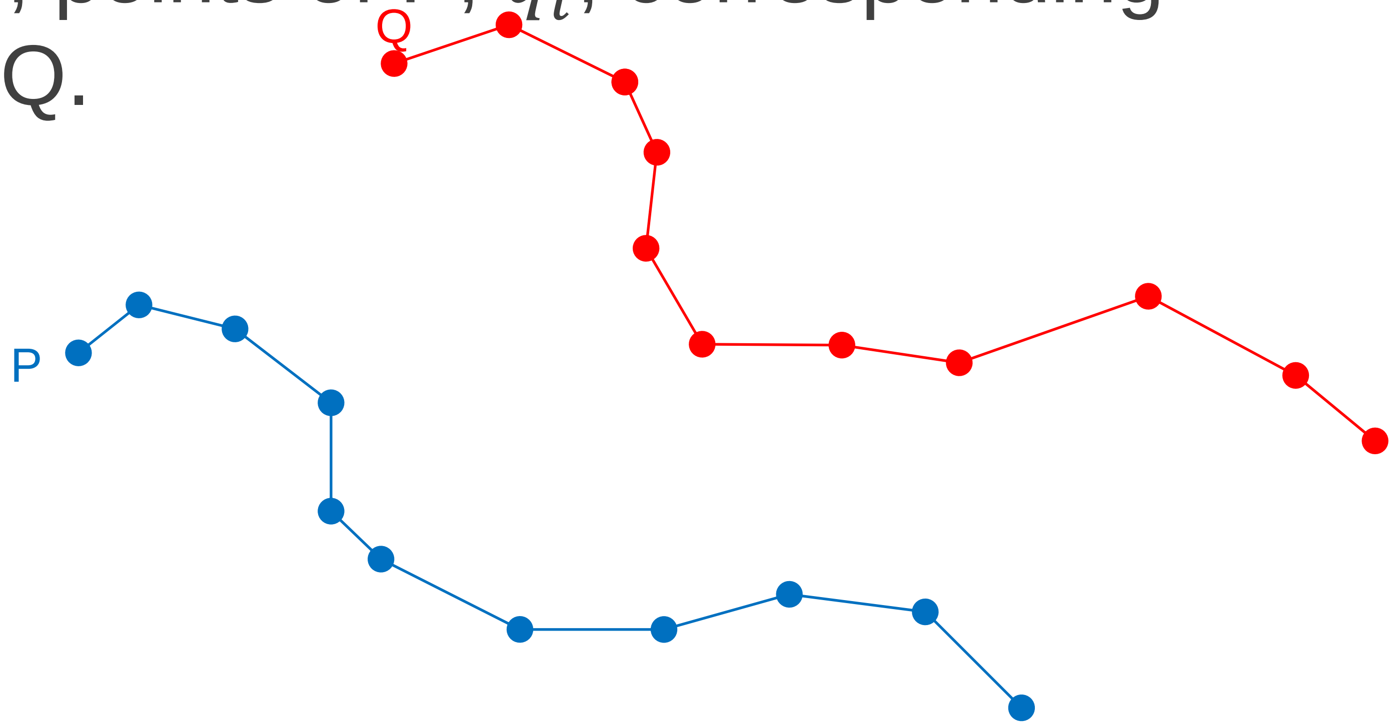


Registration

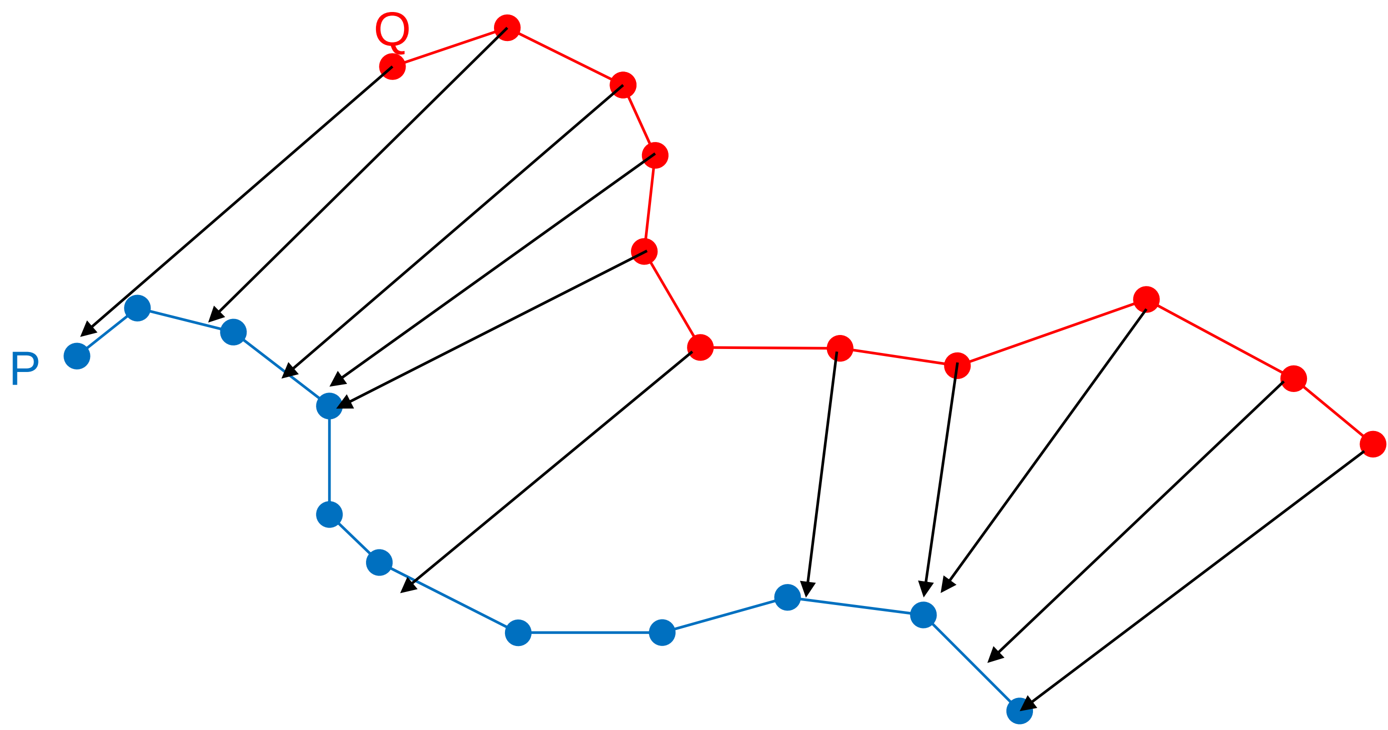
- An approximation of distance between 2 scans:

$$Error = \sum_{i}^{N_p} (Tq_i - p_i)^2$$

- Where p_i, points of P, q_i, corresponding points in Q.

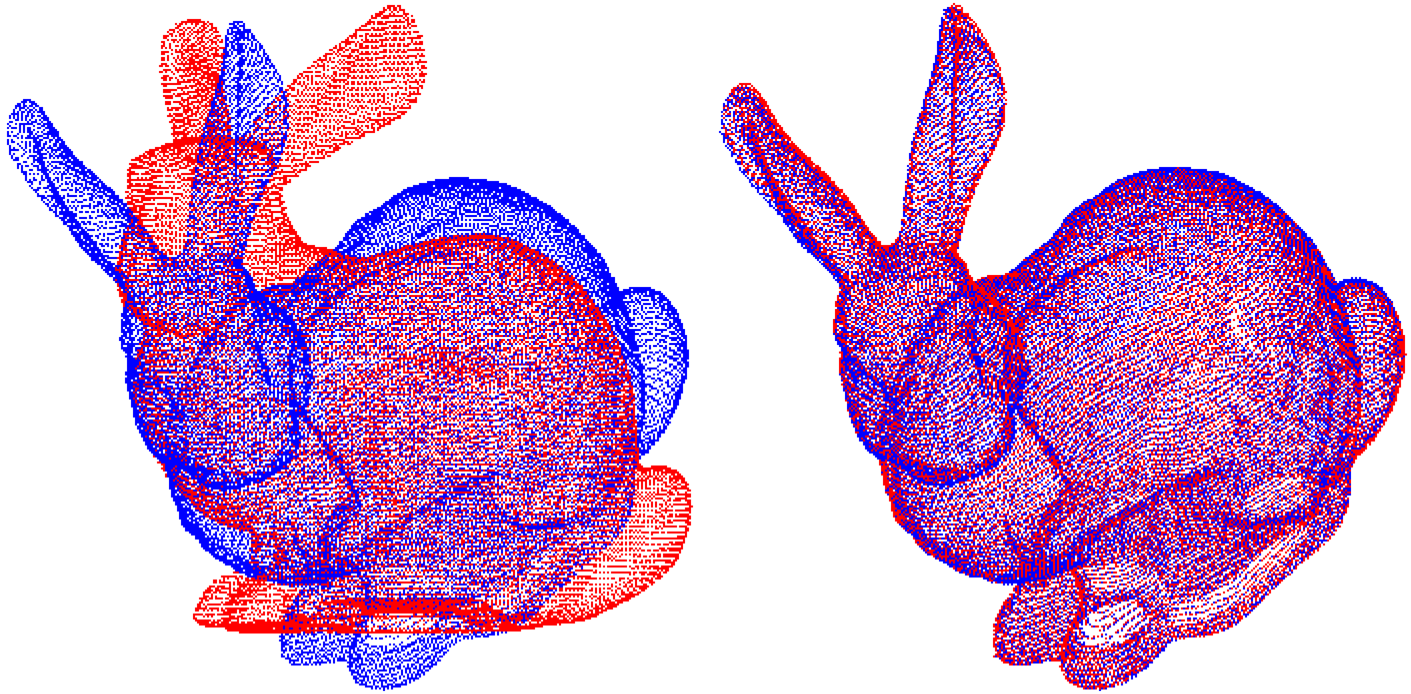


- Iterative Closest Point algortithm [Besl92]:
  - Find closest point
  - Compute transform that minimizes error
  - Repeat until ending condition.



Registration - ICP

- Iterative Closest Point algortithm [Besl92]:
  - Stanford Bunny example



Stanford Bunny example: https://www.youtube.com/watch?v=uzOCS_gdZuM

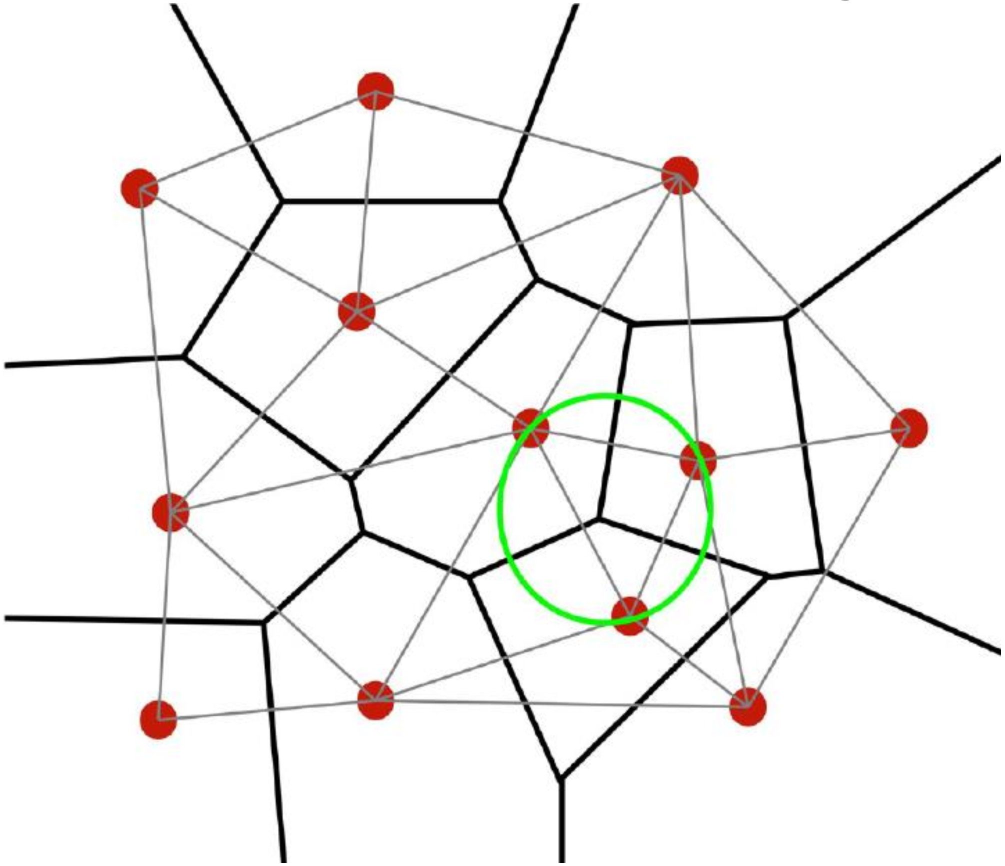
Registration – ICP problems

- Surfaces are matching only in small area may result in many outliers
- Algorithm might fall in local minima
- Typically requires an initial guess - 3 corresponding points, additional information such as GPS,…)

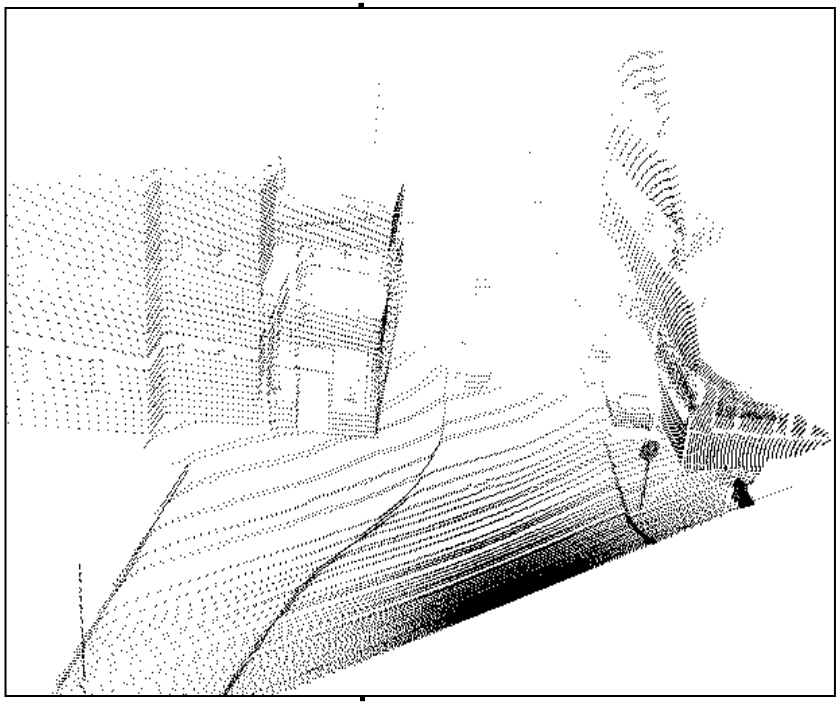
From points to surfaces

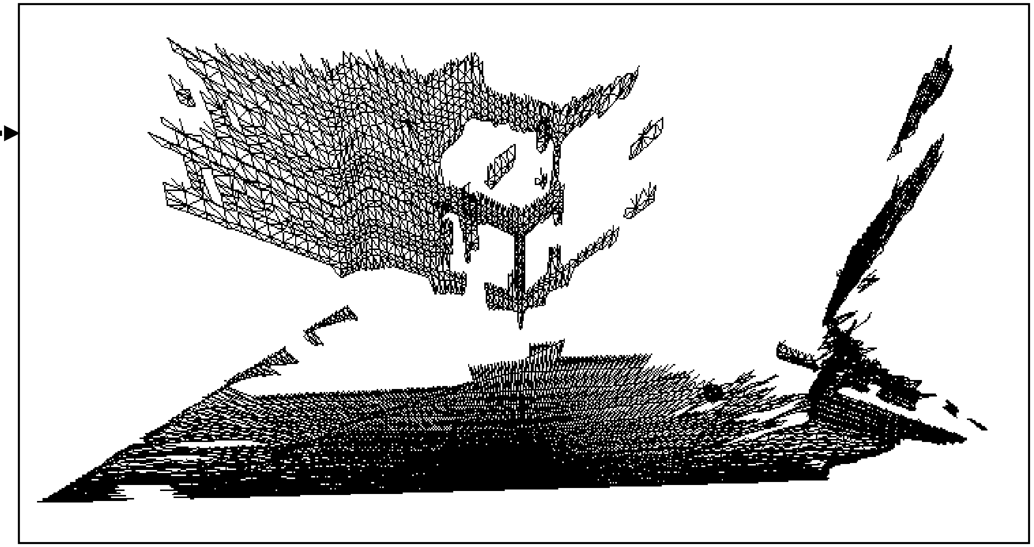
- From cloud points to surfaces:
  - Non parametric curves (triangles,…)
  - Parametric curves (cylinders, quadrics, …)

- Delaunay triangulation: for a set of 2D points P ensure that none points of the set is inside the circumcircle of any triangle.

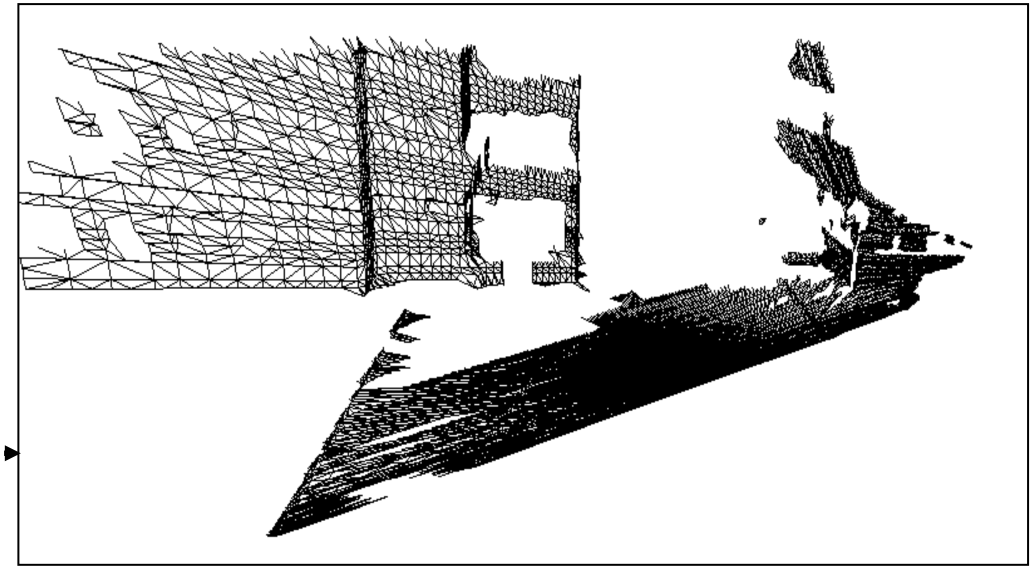


Triangulation





Triangulated model - IEETA



Other triangulation algorithms

- Marching cubes
- Marching triangles
- Ball-pivoting
- Poisson Surface Reconstruction
- Moving least-squares (MLS)

  – Possible to test some with open source Meshlab from Visual Computing Lab (http://meshlab.sourceforge.net/)

Zippering – Union of diferente meshes

- Remove overlapping portion of meshes
- Clip mesh together
- Remove triangles introduced in clipping

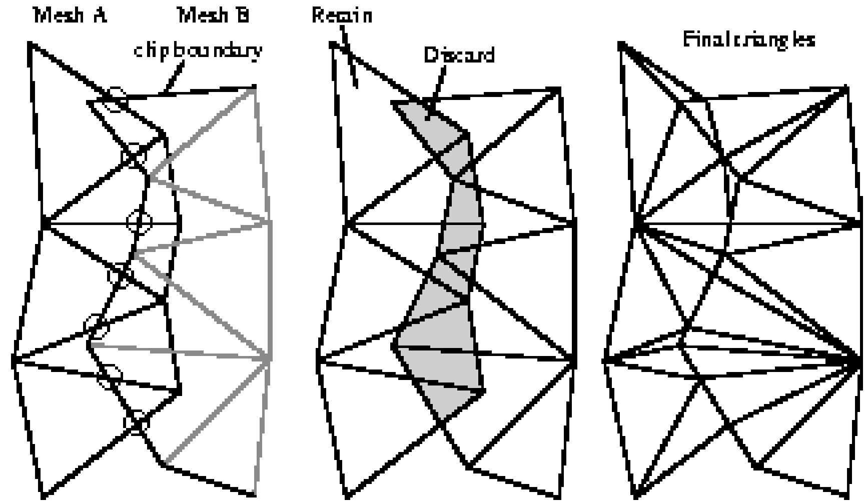


Figure 5: Mesh A is clipped against the boundary of mesh B. Circles (left) show intersection between edges of A and B's boundary. Portions of triangles from A are discarded (middle) and then both meshes incorporate the points of intersection (right).

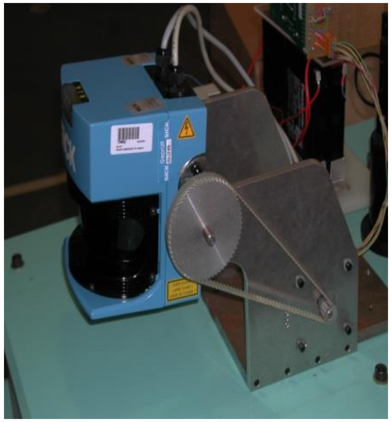
- Some 3D reconstruction techniques provide automatically texture:
  - Shape from X.
  - Structured Light Techniques
- Other do not (initial Laser Range Finder)

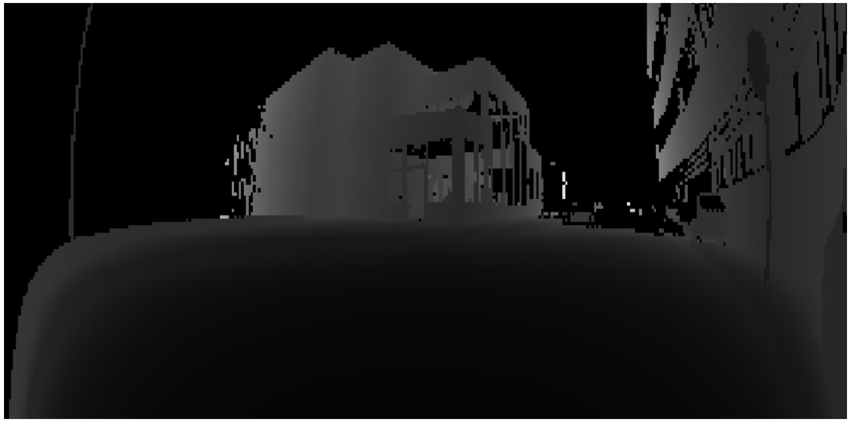
Texture mapping

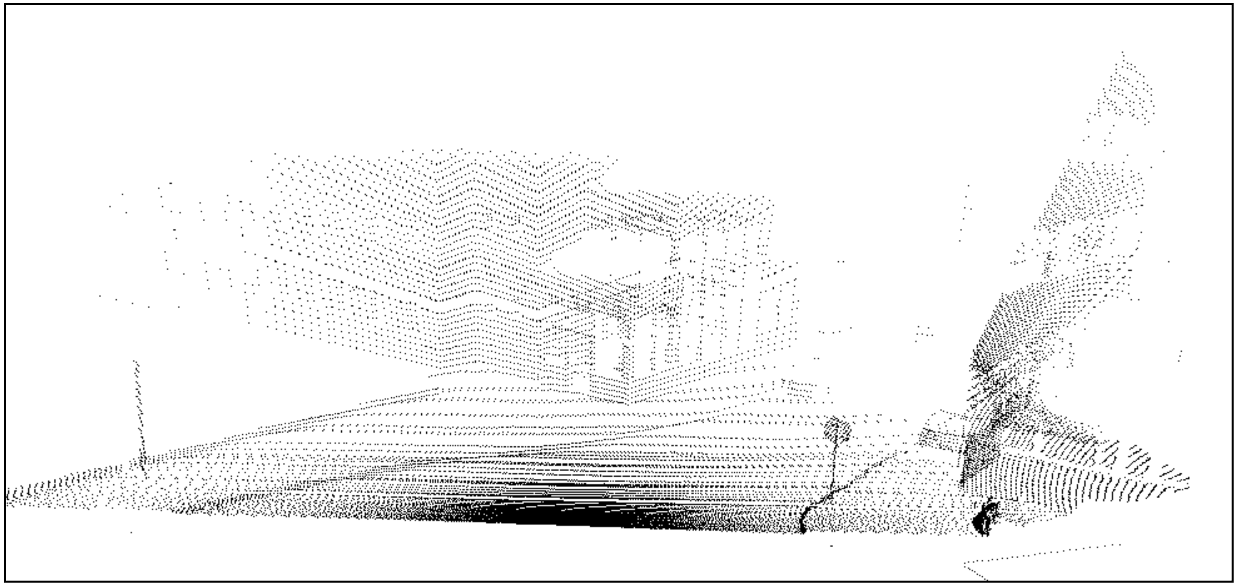
- Additional acquisition of images
- Camera calibration (might be fixed to the 3D sensor)

Textura

Range image







Cloud of points

Digital photographs

3072 x 2048
(Canon EOS 300D)

Texture Mapping – Camera calibration

Tsai camera model

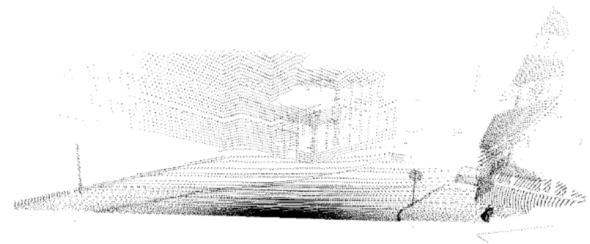
11 parameters:

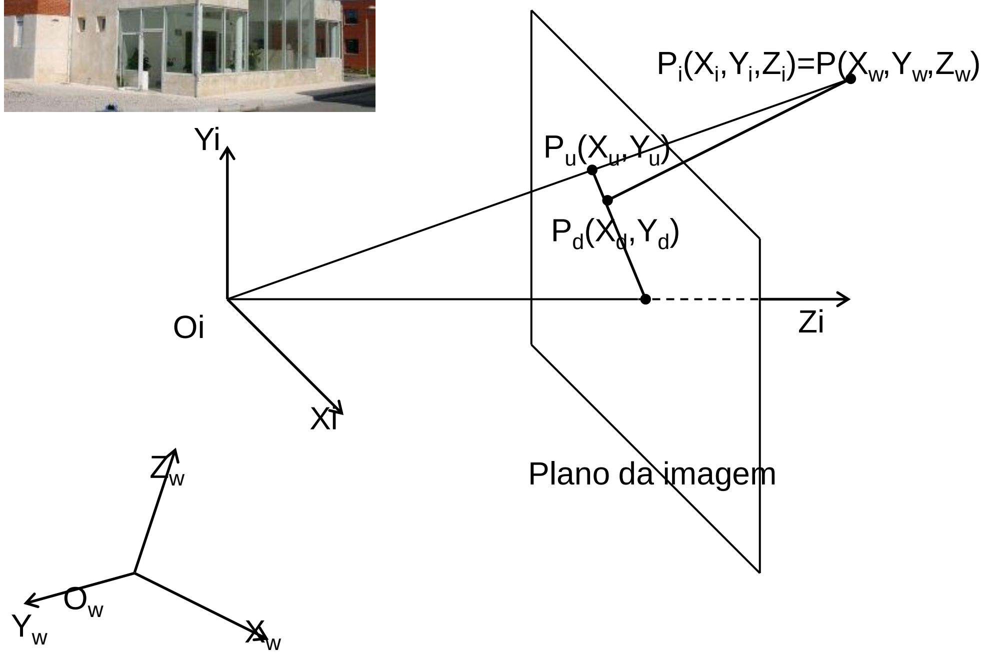
5 internals

- *f* – focal length,
- *k* – radial distortion,
- *Cx*, *Cy* – image centre,
- *Sx* – scale facto,

6 Externals

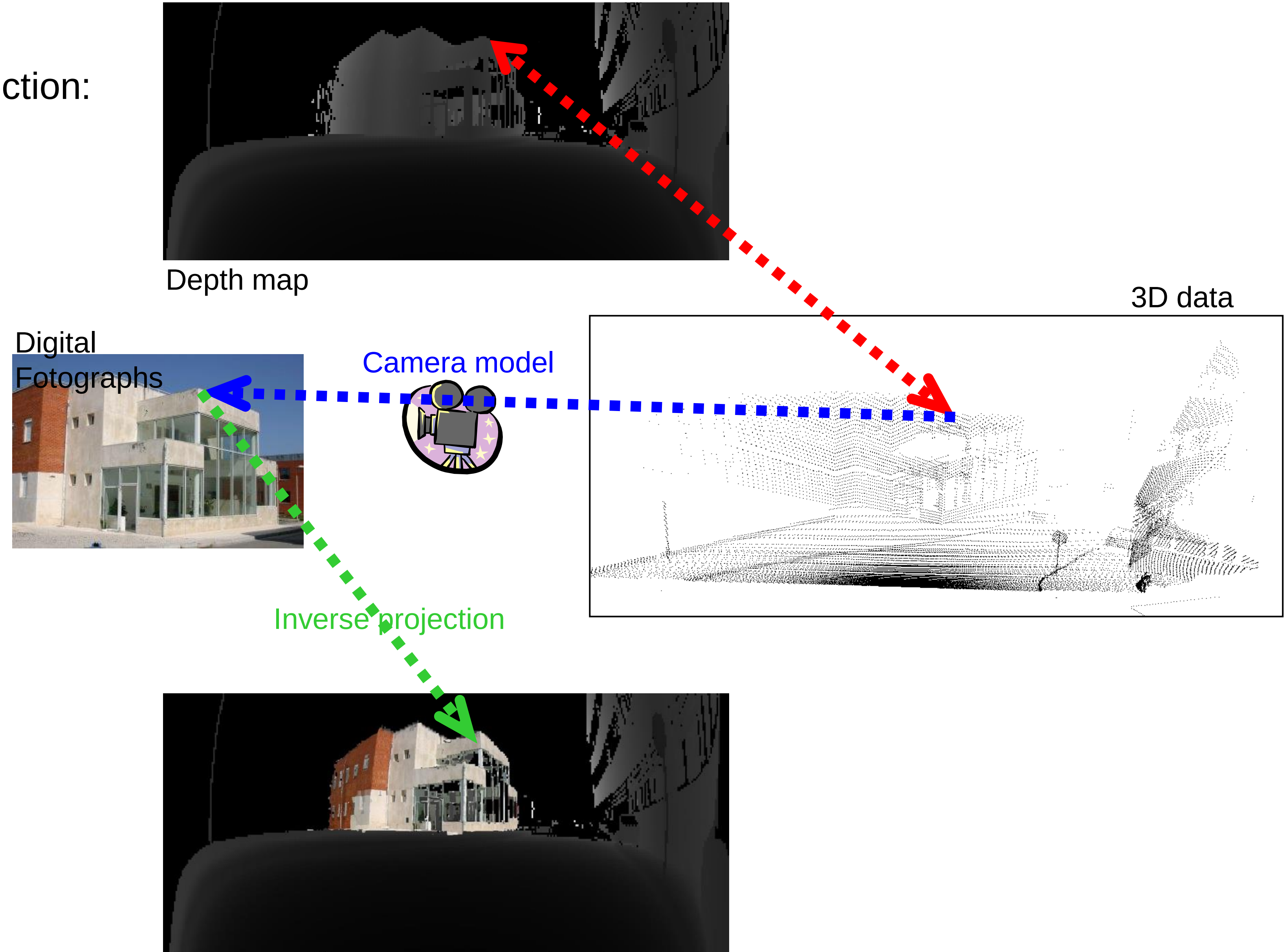
- *Rx*, *Ry*, *Rz* – rotation,
- *Tx*, *Ty*, *Tz* – translation.



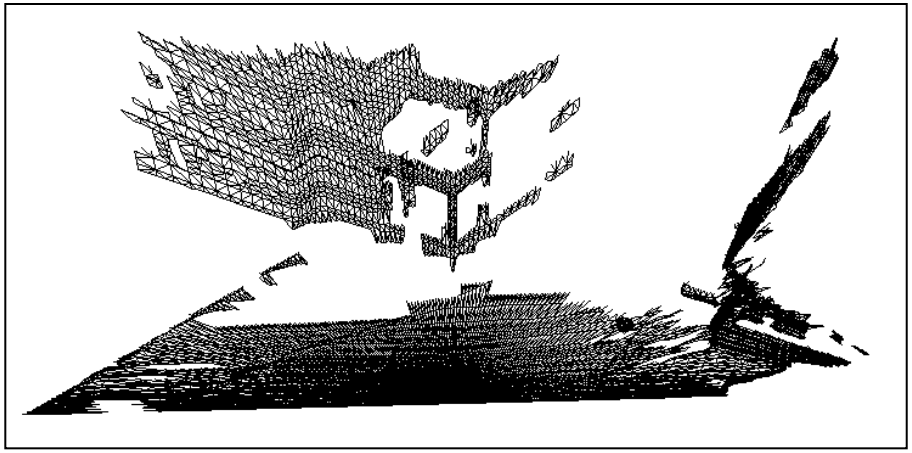


Camera Calibration

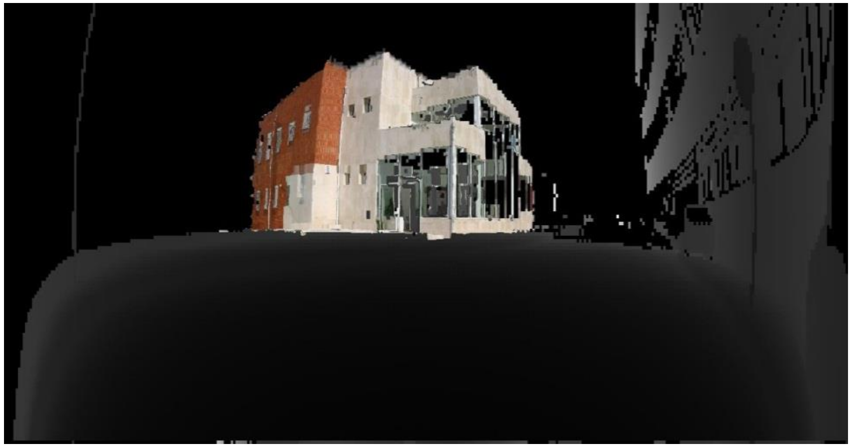
Re-projection:



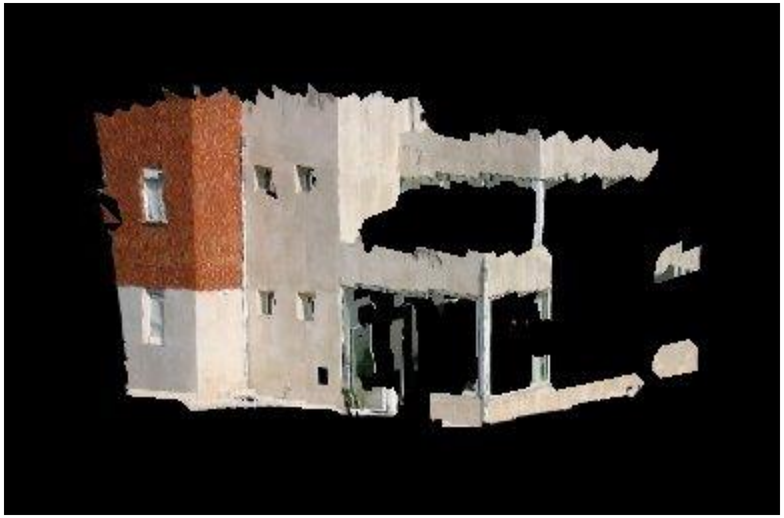
Texture

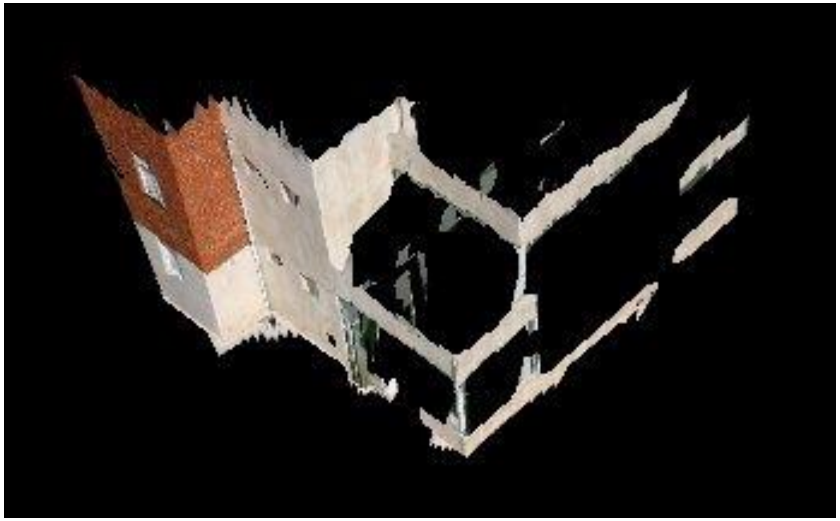


+



=





Texture mapping

- Other sensor combines calibrated 3D sensor and cameras

Some references

- Mada, S. K., Smith, M. L., Smith, L. N., and Midha, P. S. (2003). Overview of passive and active vision techniques for hand-held 3D data acquistion. In Shearer, A., Murtagh, F. D., Mahon, J., and Whelan, P. F., editors, *Proceedings of the SPIE: Optical Metrology, Imaging, and Machine Vision*, volume 4877, pages 16–27.
- Gary Bradski and Adrian Kaehler. *Learning OpenCV: Computer Vision with the OpenCV Library*. O'Reilly, Cambridge, MA, 2008.
- Szeliski, R. (2010).. Computer Vision: Algorithms and Applications, Springer
- P. Besl and N. McKay. A method for Registration of 3-D Shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 14(2):239 - 256, February 1992.
- Reg G. Willson, *Modeling and Calibration of Automated Zoom Lenses.*Ph.D. thesis, Department of Electrical and Computer Engineering, Carnegie Mellon University, January 1994
- Z. Zhang. *A flexible new technique for camera calibration.*IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 11, pp. 1330-1334, 2000
- Richard Hartley and Andrew Zisserman (2003). Multiple View Geometry in Computer Vision. Cambridge University Press. pp. 155–157. ISBN 0-521-54051-8.
- Turk, Greg and Mark Levoy "Zippered Polygon Meshes from Range Images" SIGGRAPH 1994
- https://blog.cometlabs.io/depth-sensors-are-the-key-to-unlocking-next-level-computer-vision-applications-3499533d3246