VIT-AP UNIVERSITY

Computer Vision

(Course Code: 4047)

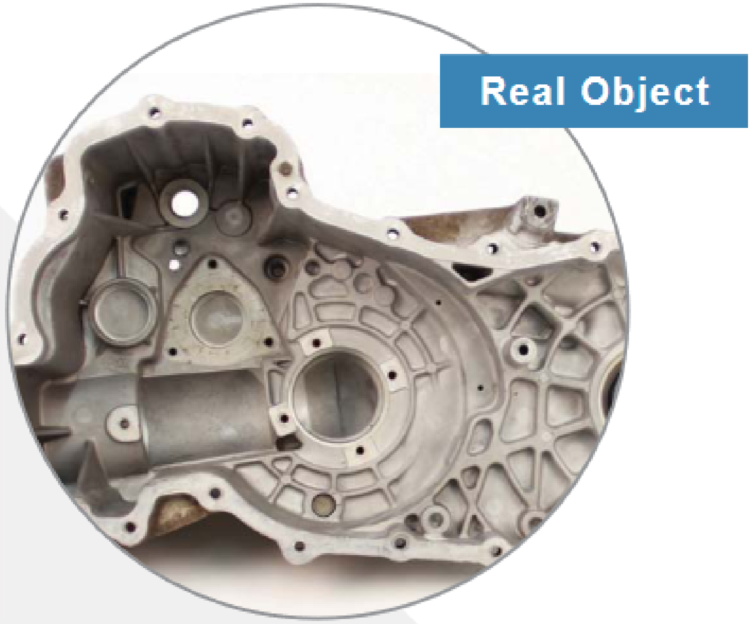
Module-6: 3D Reconstruction

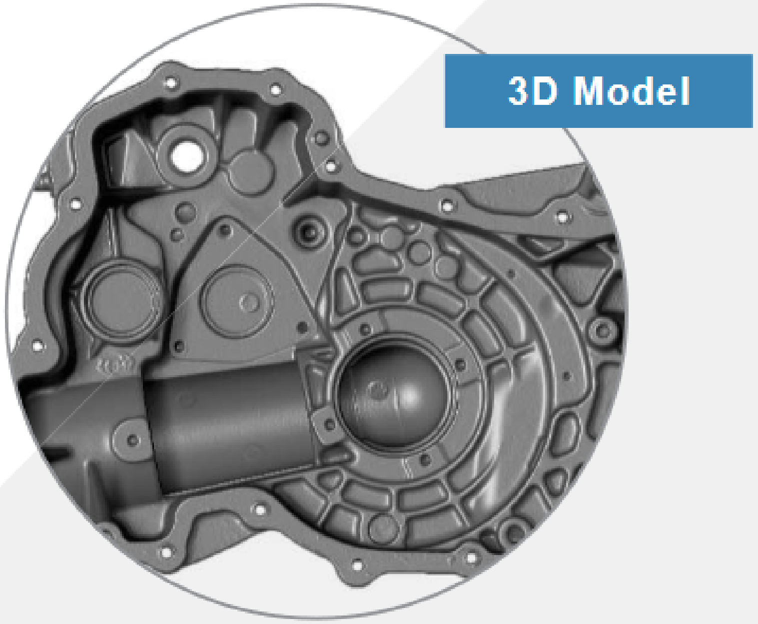
Lecture-1: 3D Point Representations

Gundimeda Venugopal, Professor of Practice, SCOPE

3D Scanning and Reconstruction: bringing physical and digital worlds

3D scanners are tri-dimensional measurement devices used to capture real-world objects or environments so that they can be remodeled or analyzed in the digital world. The latest generation of 3D scanners do not require contact with the physical object being captured.

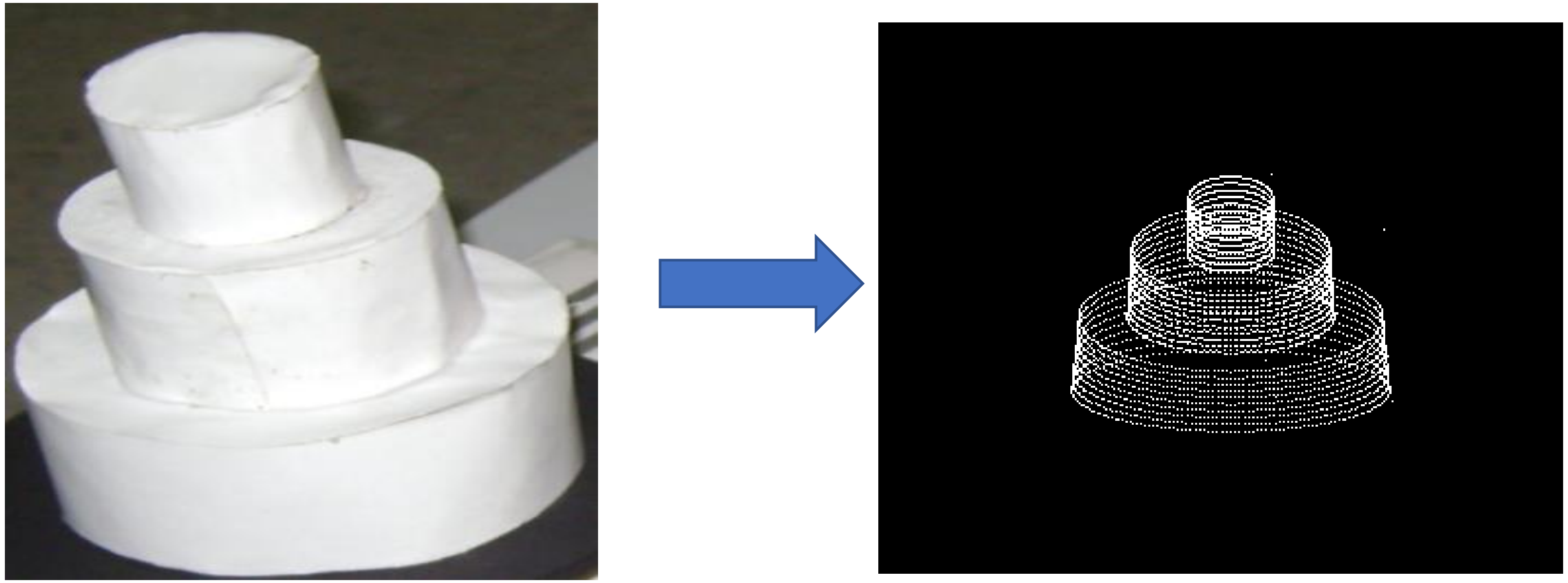




3D scanners can be used to get complete or partial 3D measurements of any physical object. The majority of these devices generate points or measures of extremely high density when compared to traditional "point-by-point" measurement devices.

3D Scanning

- **3D Scanner**: is a device that analyzes a real-world object or environment to collect data on its shape and possibly its appearance. The collected data can then be used to construct three dimensional models useful for a wide variety of applications.



- A 3D scanner works by capturing data from a physical object's surface to describe its shape in an accurate, digital, three-dimensional format.
- 2D scanner is a device that optically scans images, printed text, handwriting, or an object, and converts it to a digital image.
- 3D scanner used to analysis 3D objects, also it concerns with the outer shape of the object.

How 3D Scanning works

A 3D scanner works by capturing data from a physical object's surface to describe its shape in an accurate digital 3D format.

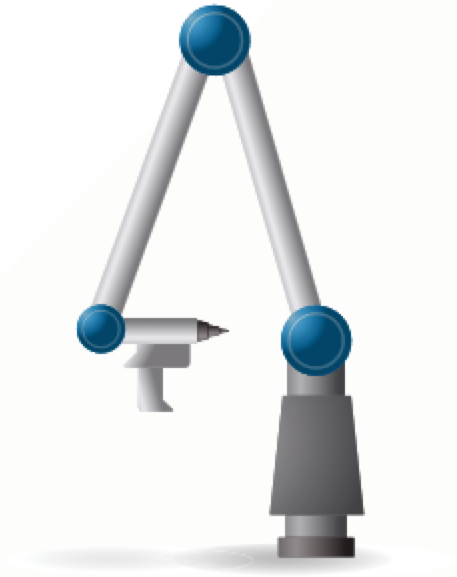
There are TWO major categories of scanners based on the way they capture data:

- **WHITE-LIGHT AND STRUCTURED-LIGHT SYSTEMS**

How 3D Scanning works

Take single snapshots or scans

- **SCAN ARMS AND PORTABLE HANDHELD SCANNERS**

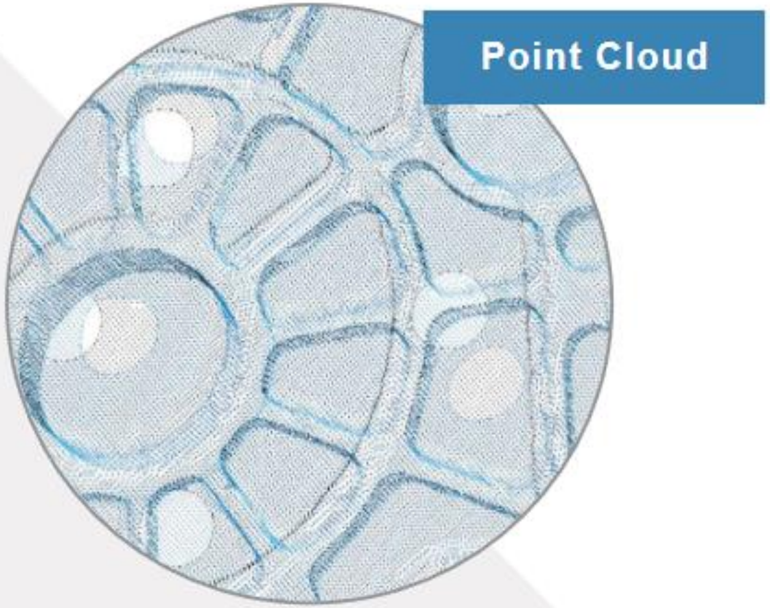


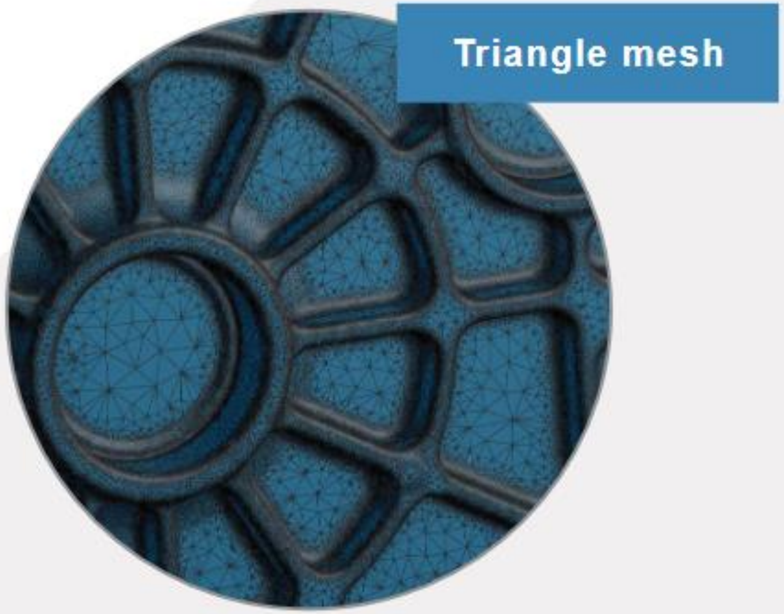
Capture multiple images continuously

How 3D Scanning works

Scanning results are represented using free-form, unstructured three-dimensional data, usually in the form of a point cloud or a triangle mesh. Certain types of scanners also acquire color information for applications where this is important.

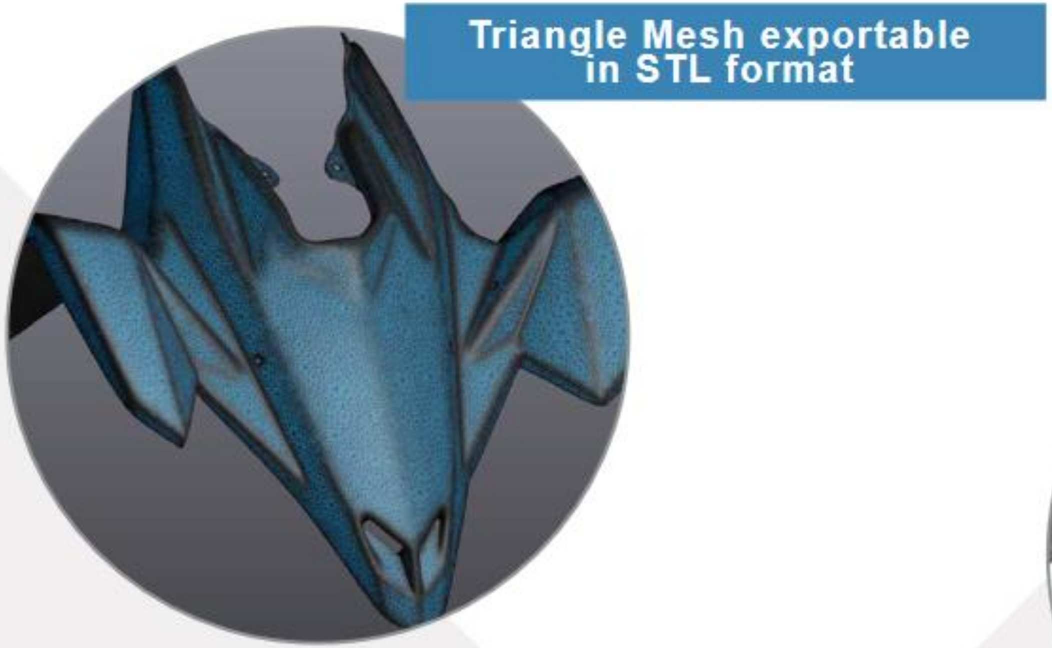
Images/scans are brought into a common reference system, where data is merged into a complete model. This process – called alignment or registration – can be performed during the scan itself or as a post-processing step.

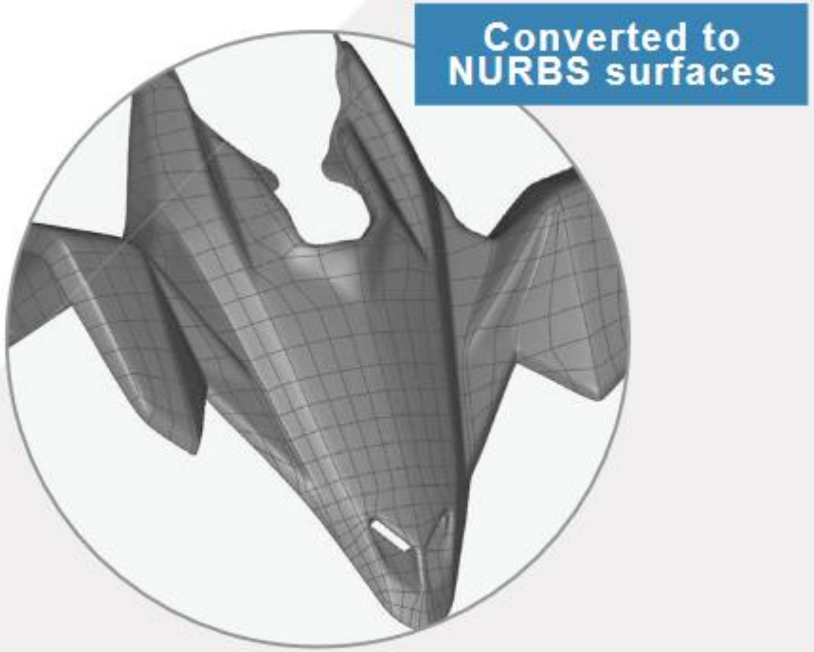




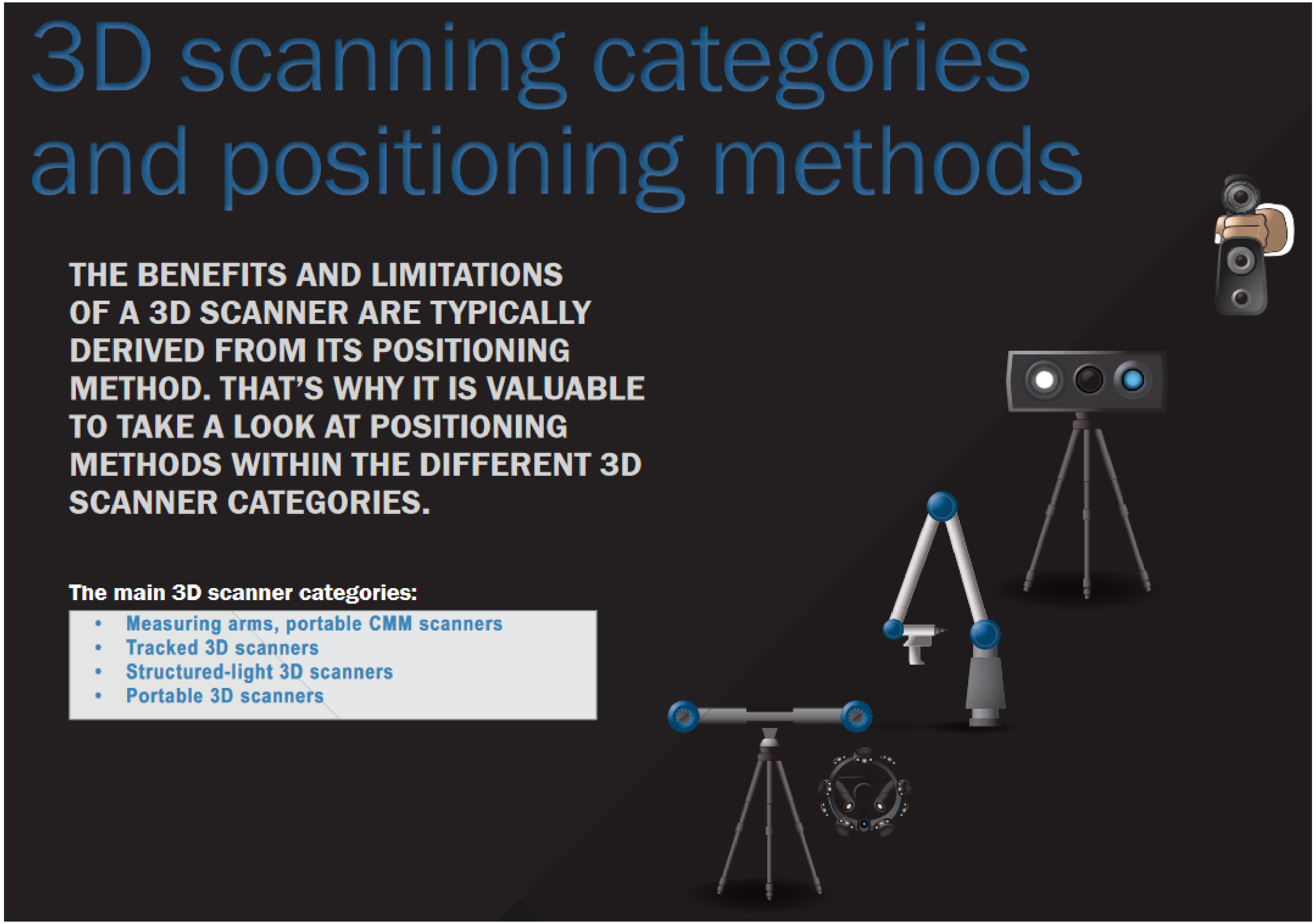
How 3D Scanning works

Computer software can be used to clean up the scan data, filling holes, correcting errors and improving data quality. The resulting triangle mesh is typically exported as an STL (STereoLithography or Standard Tessellation Language) file or converted to Non-Uniform Rational B-Spline (NURBS) surfaces for CAD modeling.





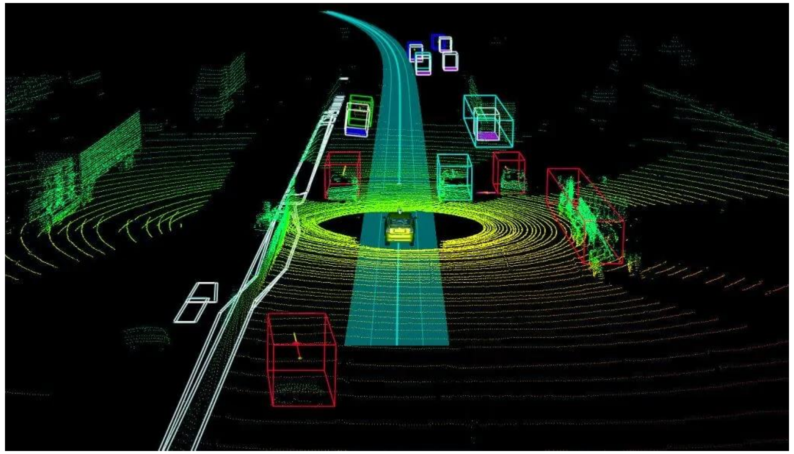
3D Scanning Categories and positioning methods



What is 3D Reconstruction

- **3D Reconstruction** is the task of creating a 3D model or representation of an object or scene from 2D images or other data sources.
- The goal of 3D reconstruction is to create a virtual representation of an object or scene that can be used for a variety of purposes, such as visualization, animation, simulation, and analysis.
- 3D reconstruction can be used in fields such as computer vision, robotics and virtual reality.

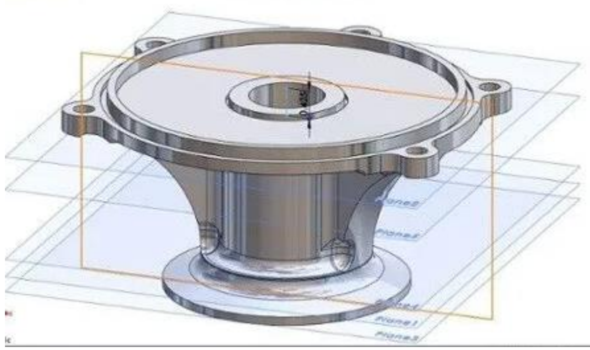
- Healthcare: Dentistry, Orthopaedics, Opthalmology
- Gaming: 3D objects, Models,
- Industries: Design and printing of 3D parts. CAD/CAM
- Autonomous Vehicles: Traffic Models
- Architecture and Construction - Building Information Modeling
- … Lot of other application



3D point cloud data is most useful for autonomous vehicles that need to understand their environment and perform simultaneous localization and mapping in real-time in order to keep moving in the right direction avoiding obstacles.

(a) Image

(b) Silhouette

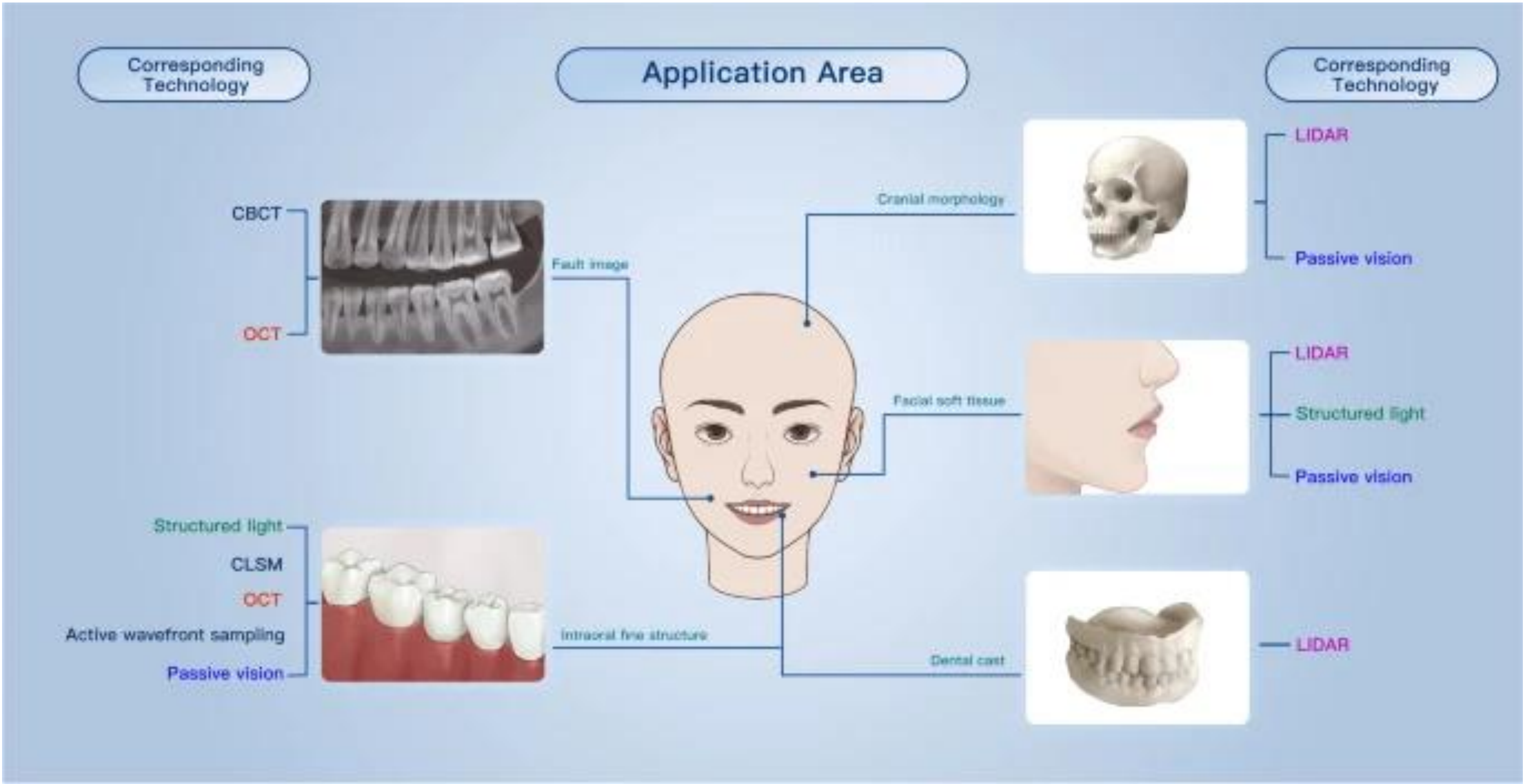


3D CAD Model

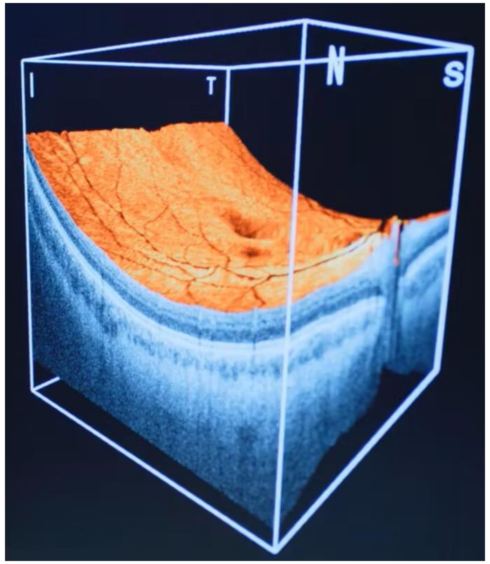
3D Scanning and Reconstruction applications

- **1. Architecture and Construction: Building Information Modeling (BIM)**: 3D reconstruction helps create accurate models of buildings and infrastructure, facilitating better planning, design, and management. **Site Analysis**: It can be used to assess construction sites and visualize how new structures will fit into existing environments.
- **2. Cultural Heritage and Archaeology:**
- **Digital Preservation**: 3D reconstruction is employed to digitally preserve historical sites and artifacts, allowing for virtual tours and research without physical degradation.
- **Excavation Documentation**: Archaeologists use 3D models to document excavation sites, which aids in analysis and presentation.
- **3. Medical Imaging**
- **Surgical Planning**: 3D reconstruction from medical scans (like CT or MRI) helps in visualizing complex anatomies, assisting surgeons in planning procedures.
- **Prosthetics and Implants**: Custom prosthetics and implants can be designed using 3D models of patients' anatomy.
- **4. Robotics and Autonomous Vehicles**
- **Navigation and Mapping**: Robots and self-driving cars use 3D reconstruction for environment mapping, obstacle detection, and navigation.
- **Simultaneous Localization and Mapping (SLAM)**: This technique combines 3D reconstruction with real-time localization for autonomous systems.
- **5. Virtual Reality (VR) and Augmented Reality (AR)**
- **Immersive Experiences**: 3D reconstruction enhances VR and AR applications by creating realistic environments and objects that users can interact with.
- **Training and Simulation**: These technologies are leveraged in training simulations for various fields, including aviation, medicine, and military.
- **6. Gaming and Entertainment**
- **Game Development**: 3D reconstruction is used to create realistic game environments and characters, enhancing player immersion.
- **Film Production**: In visual effects, 3D reconstruction helps integrate CGI with live-action footage seamlessly.
- **7. Forensics and Crime Scene Investigation**
- **Crime Scene Reconstruction**: 3D models of crime scenes can be created to analyze evidence, visualize scenarios, and present findings in court.
- **Accident Reconstruction**: This application helps in analyzing vehicle accidents to determine causes and contributing factors.
- **8. Manufacturing and Quality Control**
- **Reverse Engineering**: 3D scanning and reconstruction are used to create models from existing parts for replication or modification.
- **Quality Inspection**: 3D models help in inspecting parts for compliance with specifications and tolerances.
- **9. Education and Training**
- **Interactive Learning**: 3D models can be used in educational settings to provide interactive and engaging learning experiences in subjects like biology, history, and engineering.
- **Skill Development**: Fields like surgery or engineering can benefit from 3D simulations for hands-on training.

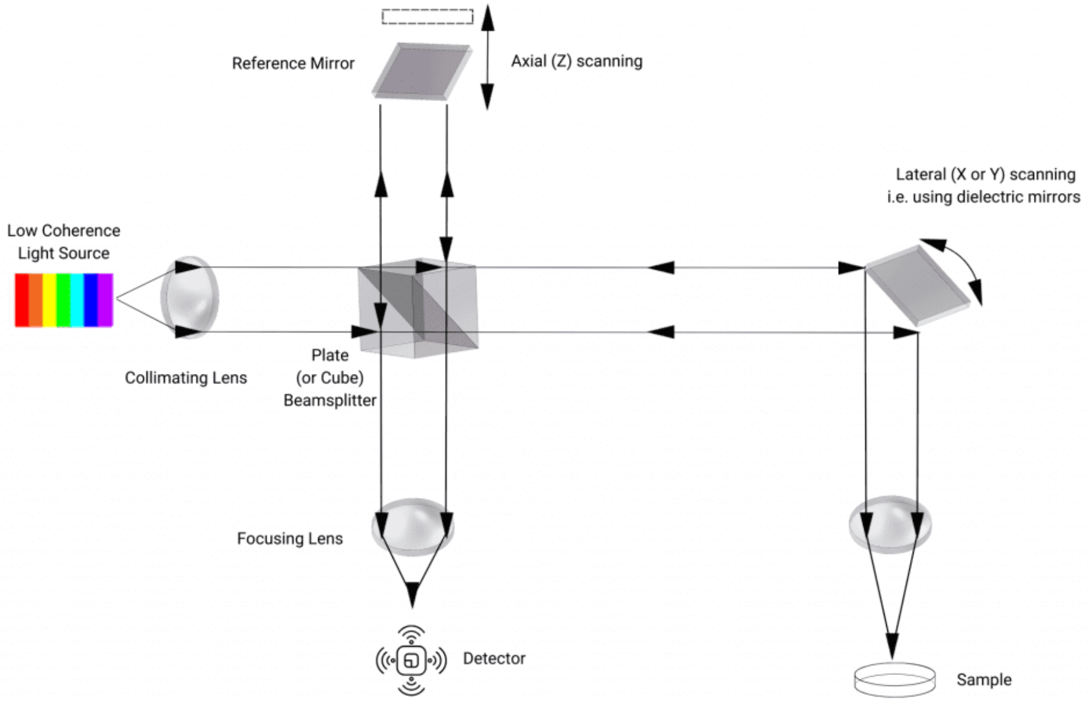
3D Reconstruction in Dentistry



Optical Coherence Tomography (OCT): Eye Retina Scan

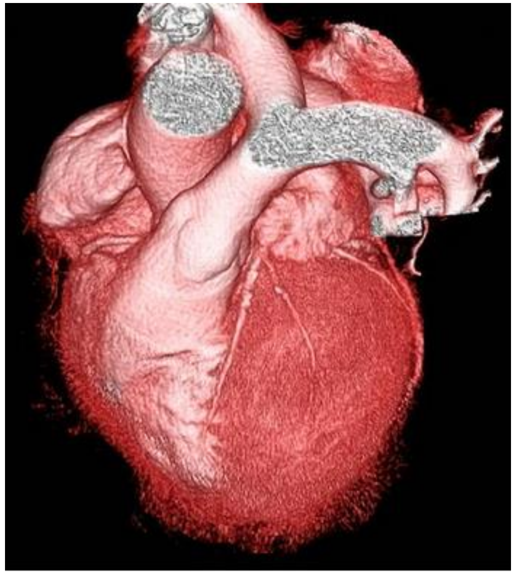


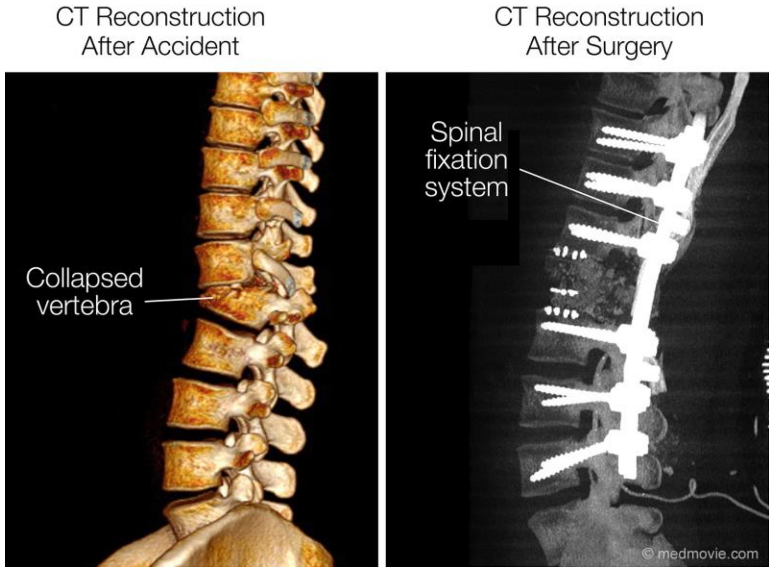
- OCT 3D Scan for Eye: The OCT 3D images provides a cross sectional view of the inner anatomy of your eyes, specifically the retinas, to provide a deeper analysis of your eye health going into the eye test.
- How does the OCT scan work?
  - The OCT 3D scan works like an ultrasound. However, rather than using sound waves, it uses light waves to see the detailed layers of the retina. The retina is at the back of the eye.
  - It converts light into images, which allows you to see.
  - The OCT machine takes over 1,000 images of your eye, from the retina to the optic nerve. This then creates a detailed 3D image, which lets us see your eye's structure.
  - To generate 3D images, OCT employs Fourier Transform techniques. The interference pattern obtained from the combined beams is processed using Fourier Transform algorithms to extract depth information. By scanning the sample arm along one axis, a series of A-scans (depth profiles) are acquired. These A-scans, when combined, form a two-dimensional (2D) cross-sectional image known as a B-scan. By sequentially acquiring multiple B-scans, a 3D volumetric image can be reconstructed.
- Key Applications of OCT
  - Ophthalmology
  - Neuroscience
  - Guided Surgery
  - Industrial Processing and Non-Destructive Testing (detecting defects or irregularities in materials).. Additionally, OCT is employed in the evaluation of coatings, paints, and thin films, providing detailed structural information and enhancing the efficiency of industrial processes.

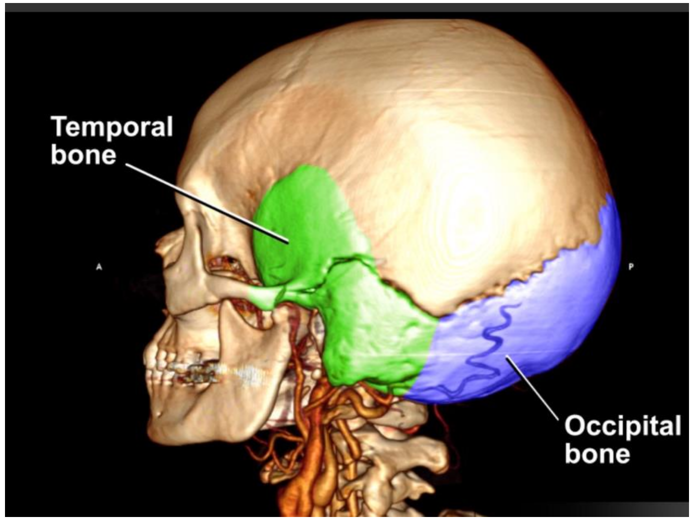


CT Scan and MRI

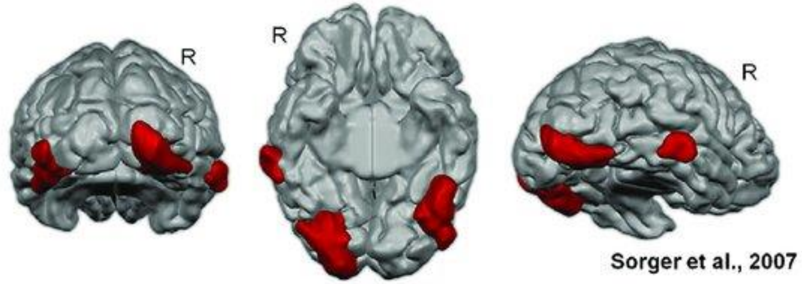
CT Scan

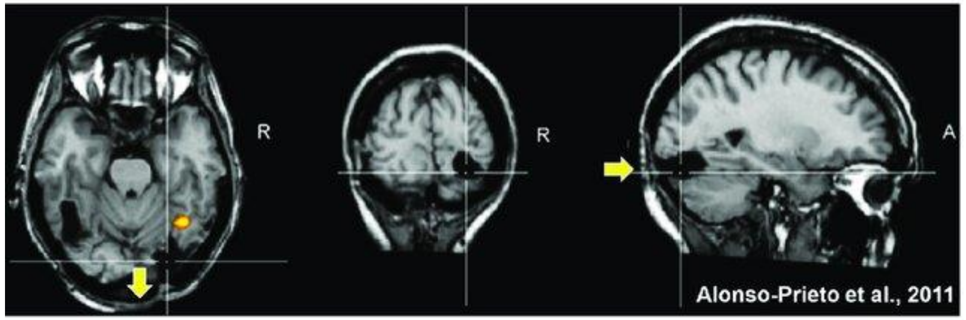


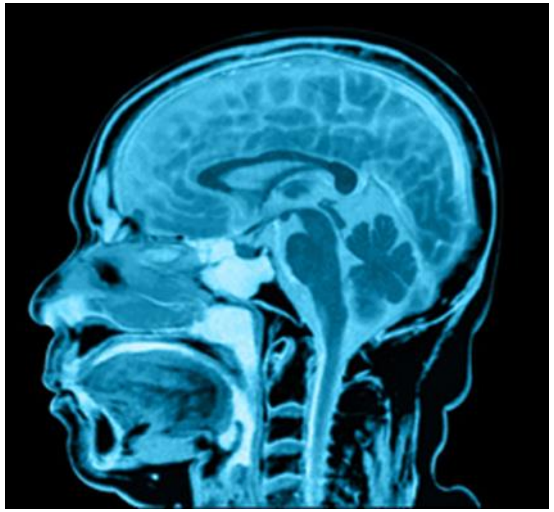




MRI

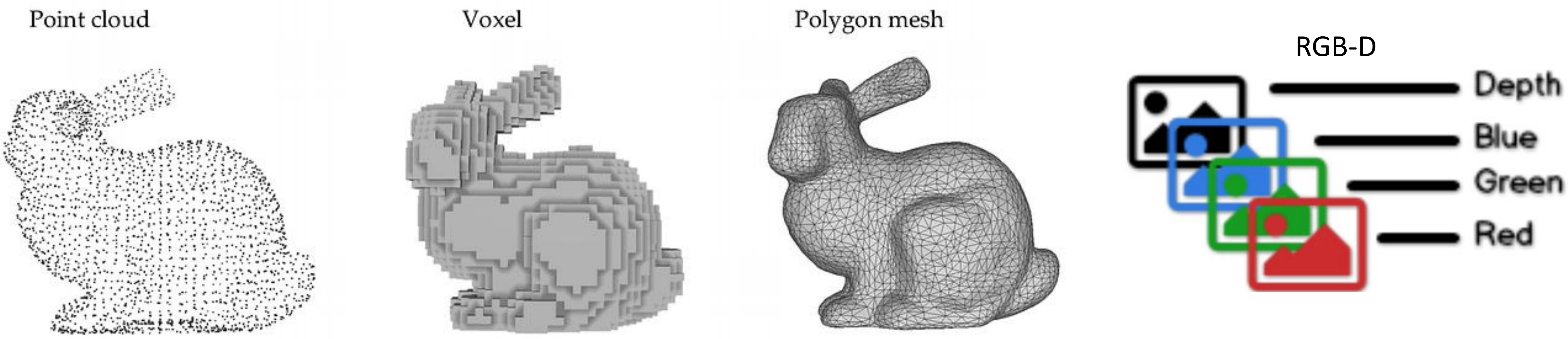






3D Representation

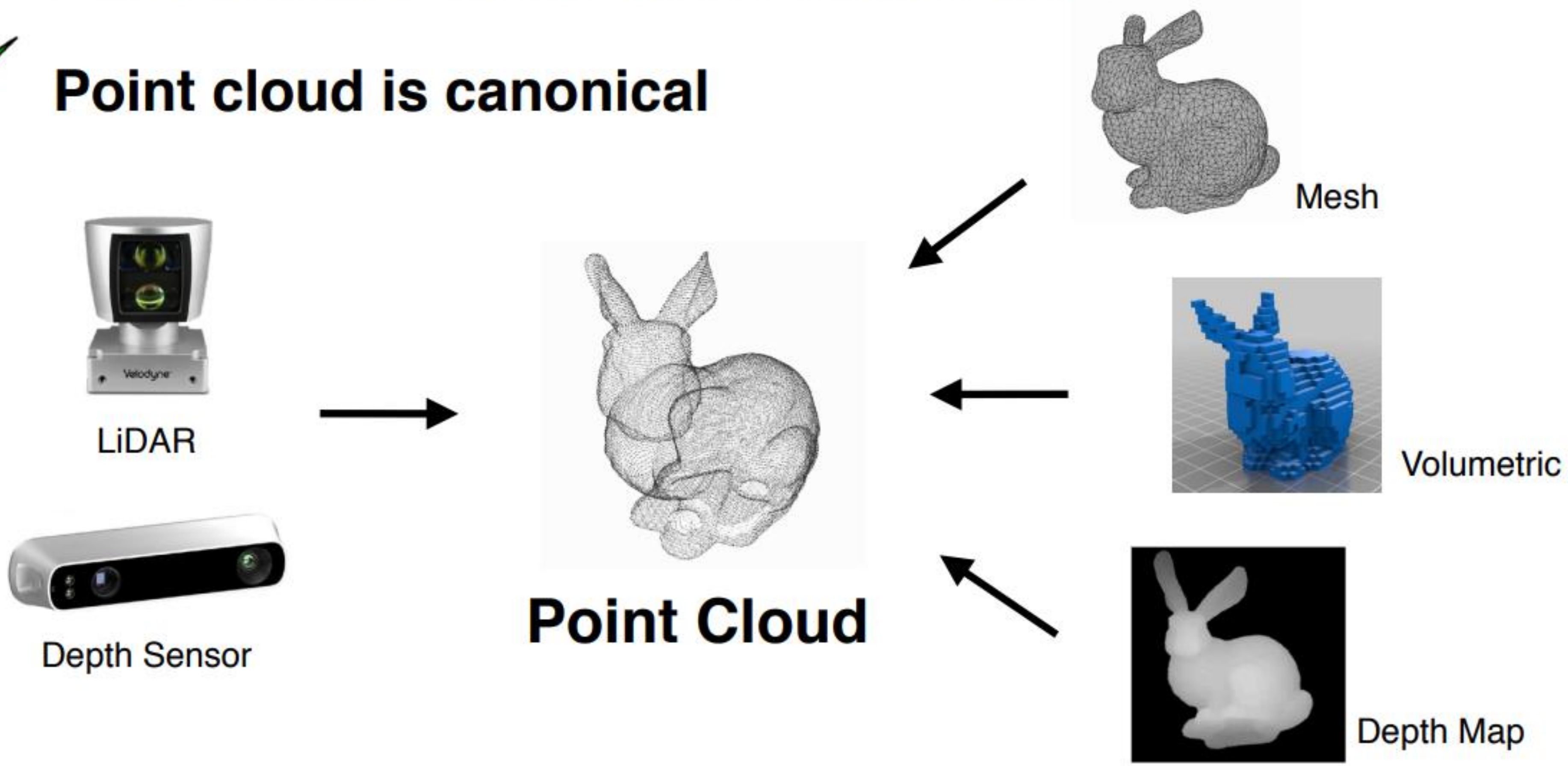
Commonly used 3D data representations: point clouds, voxel representations, and meshes:



Point cloud, voxel, Polygon mesh and RGB-D representation of 3D models

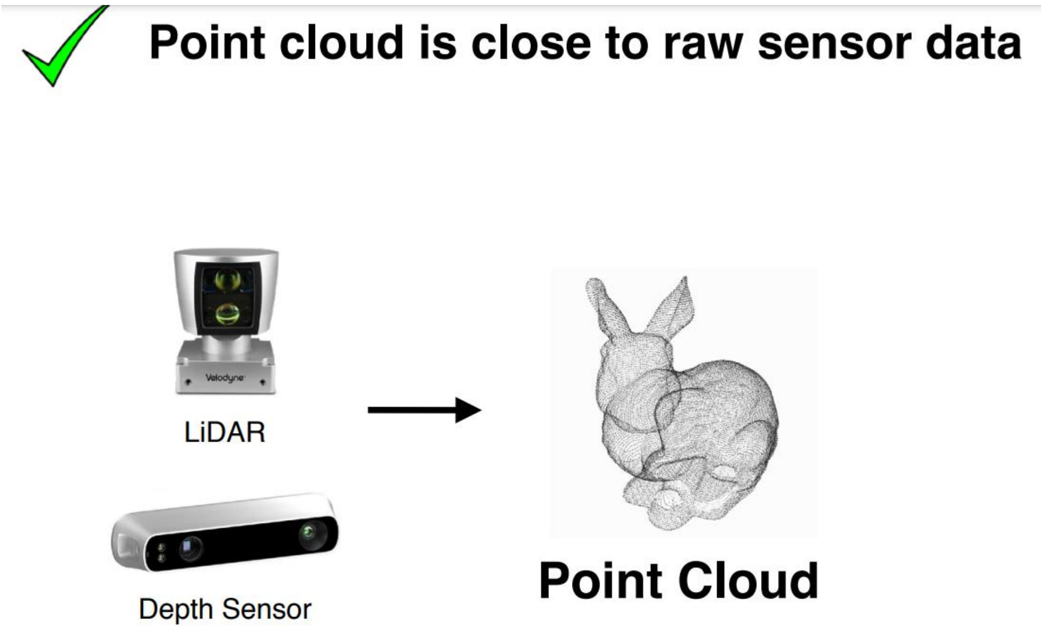
3D Representation: Point Cloud

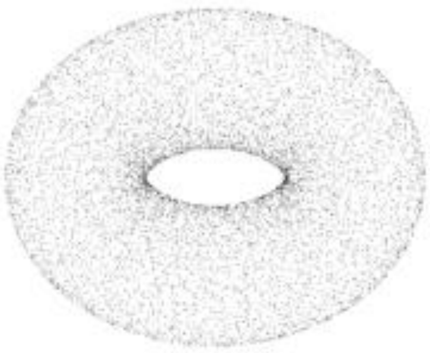
- **Point cloud is close to raw sensor data**
- **Point cloud is canonical**



3D Representation: Point Cloud

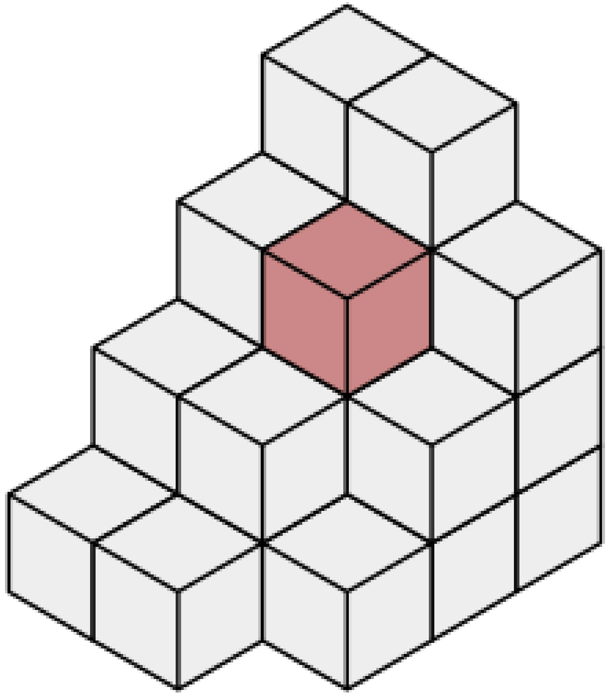
- A 3D point cloud is a straightforward way to represent objects in three dimensions. It consists of individual 3D points, with each point represented by a tuple (x, y, z).
- Depth cameras often provide raw measurements in the form of 3D point clouds.   e.g., LIDAR
- Point clouds are unordered and irregular data types.
- Unlike images, where pixels have clear neighboring pixels, point clouds lack regular structures, and applying convolutions directly is usually not feasible. (for Deep Learning)





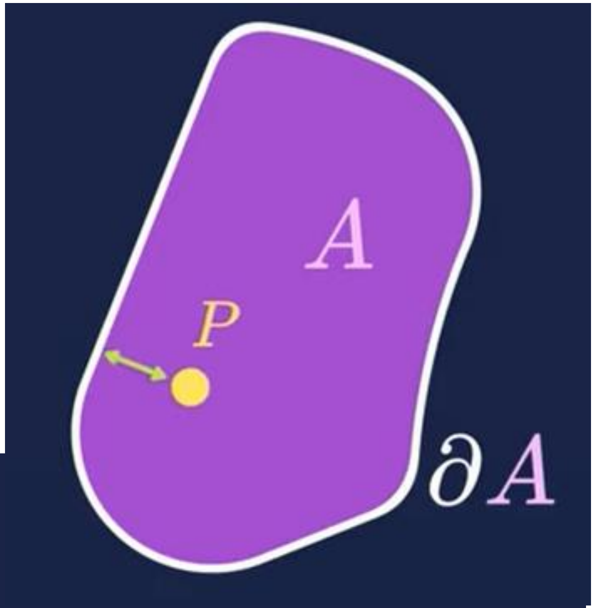
3D Representation: Voxel

- Voxel representations are defined using Voxels. A voxel is akin to a pixel but in 3D space. It's formed by dividing a 3D cube into smaller cubes, with each smaller cube called a voxel.
- Voxel representations often use Truncated Signed Distance Functions (TSDFs) to represent 3D surfaces.
- The main purpose of TSDF is to represent and reconstruct 3D surfaces or shapes from different measurements or observations, like depth images from a 3D camera or point clouds from a LiDAR sensor. By combining these measurements using TSDF, we can create a more accurate and complete representation of the 3D surface.
- An SDF is a function that can tell you how far any point in space is from the surface of that object. If the point is outside the object, the distance will be positive. If the point is inside the object, the distance will be negative. And if the point is exactly on the surface, the distance will be zero.
- Truncation (using TSDFs): In a Truncated Signed Distance Function, we limit or "truncate" the range of distances that we care about. This means that we only consider distances up to a certain threshold, and ignore any distances beyond that. Truncation is useful because it simplifies the representation of the 3D shape and reduces the amount of data we need to store
- A Signed Distance Function (SDF) is defined at each voxel as the distance between the voxel center and the nearest point on the surface. A Truncated SDF restricts values to a range between -1 and +1.



A series of voxels in a stack, with a single voxel shaded

Signed Distance Function

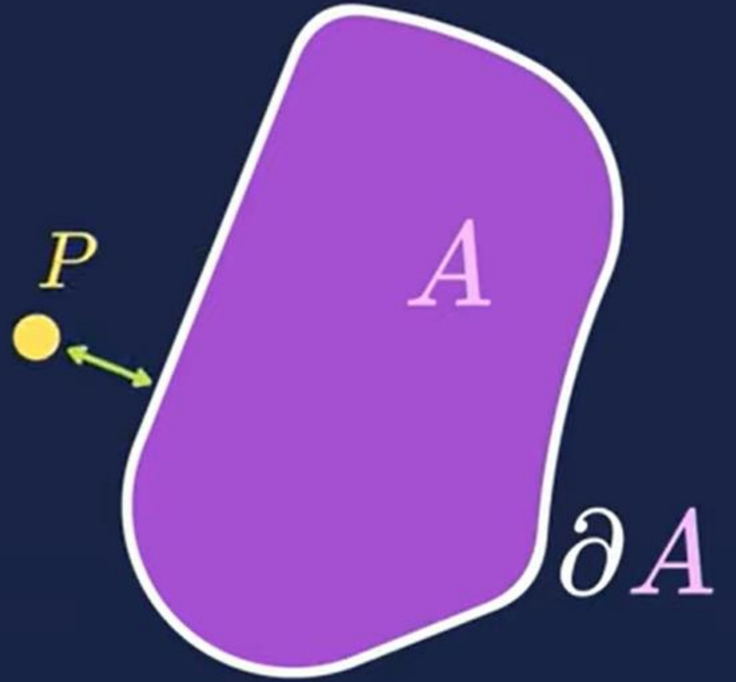


Distance:

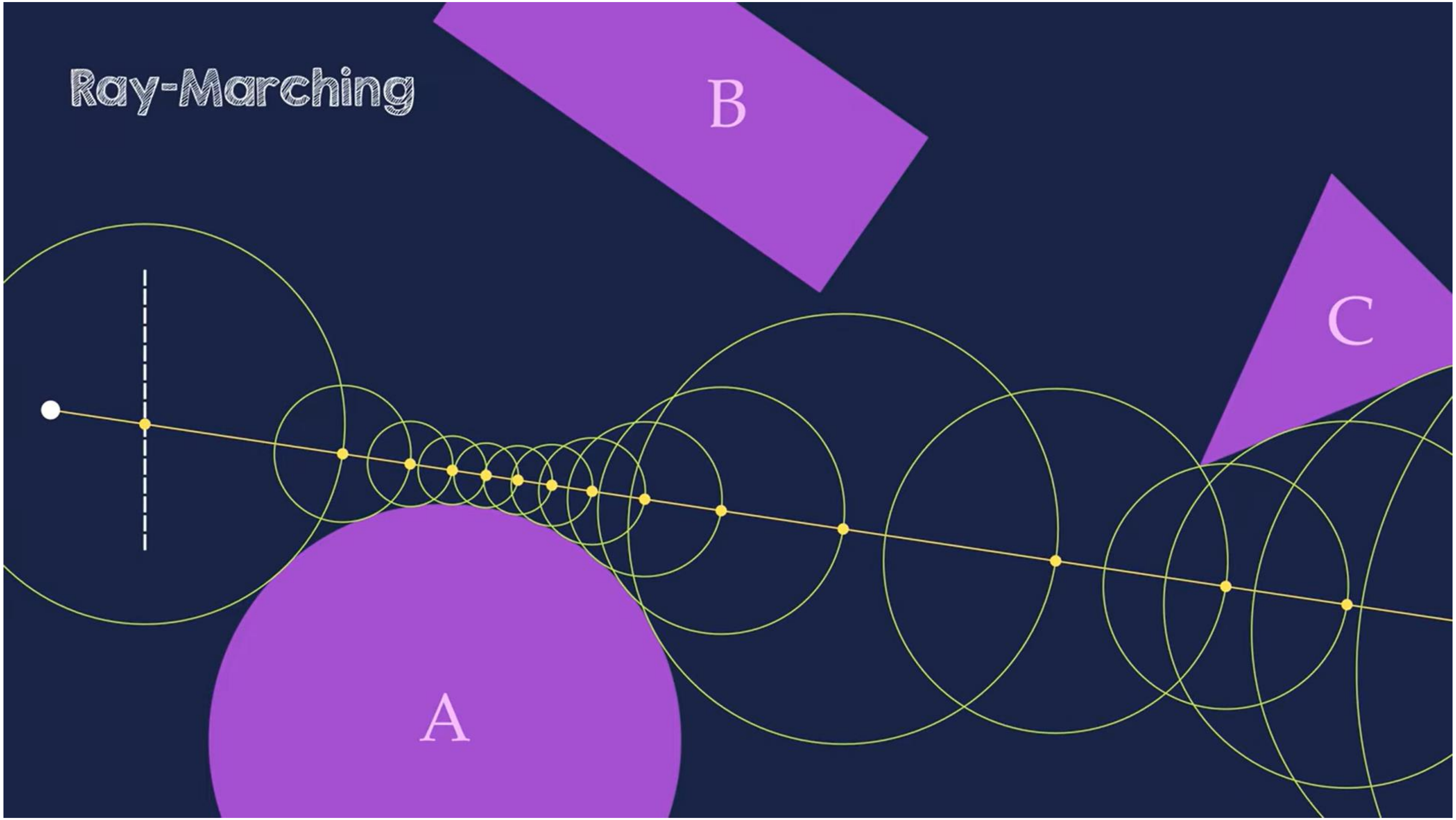
$$\mathrm{dist}(P, A) := \min\{\mathrm{dist}(P, Q) \mid Q \in A\}$$

Signed Distance:

$$d_A(P) := \begin{cases} \mathrm{dist}(P, \partial A), & P \text{ outside} \\ -\mathrm{dist}(P, \partial A), & P \text{ inside} \end{cases}$$

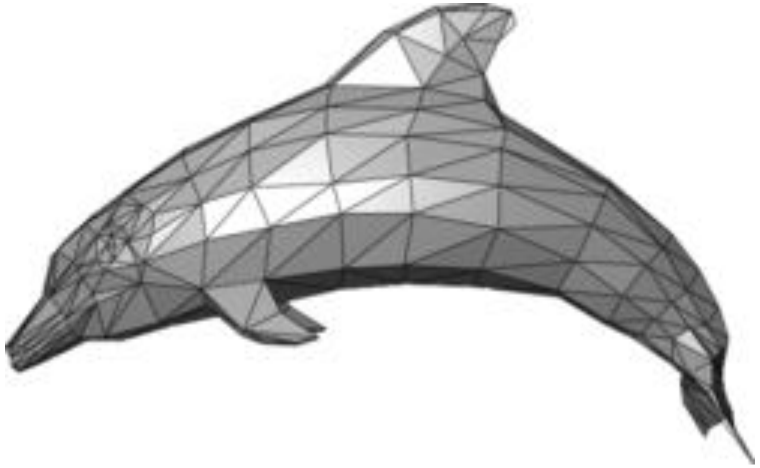


Ray Marching: Need of a Distance Function



3D Representation: Mesh

- Meshes, are popular 3D representations and they consist of vertices (3D points) and faces (polygons) defined on these vertices.
- Meshes often result from post-processing depth camera measurements or are manually designed during 3D asset creation.
- Meshes provide additional geometric, topological, and surface-normal information compared to point clouds.
- This extra information proves beneficial in training learning models. For example, graph convolutional neural networks treat meshes as graphs and use vertex neighbor information for convolutional operations.



Example of a low poly triangle mesh representing a dolphin

3D Representation: RGB-D?

- RGB-D image. It is like a stack of single valued images. Each pixel has four properties, Red, Green, Blue, Depth.
- With normal pixel-based images, we can locate any pixel by the (x, y) coordinates, then we can obtain three properties (R, G, B) colour respectively.
- While with RGB-D image, each (x, y) coordinates would be corresponding to four properties (depth, R, G, B).
- The only difference between RGB-D and point cloud is that with point cloud, the (x, y) coordinates reflect the actual values in the real world rather than the simple integer values.
- Point cloud can be constructed from RGB-D images. If you have a RGB-D scanning and also known the intrinsics of your scanning camera, you will be able to create the point clouds from the RGB-D image, by simply calculating the real world (x, y) with your camera intrinsics. It is called camera calibration.
- RGB-D image are grid-aligned images while the point cloud is in a more sparse structure.

RGB-D image. It is like a stack of single valued images.
Each pixel has four properties, Red, Green, Blue, Depth.

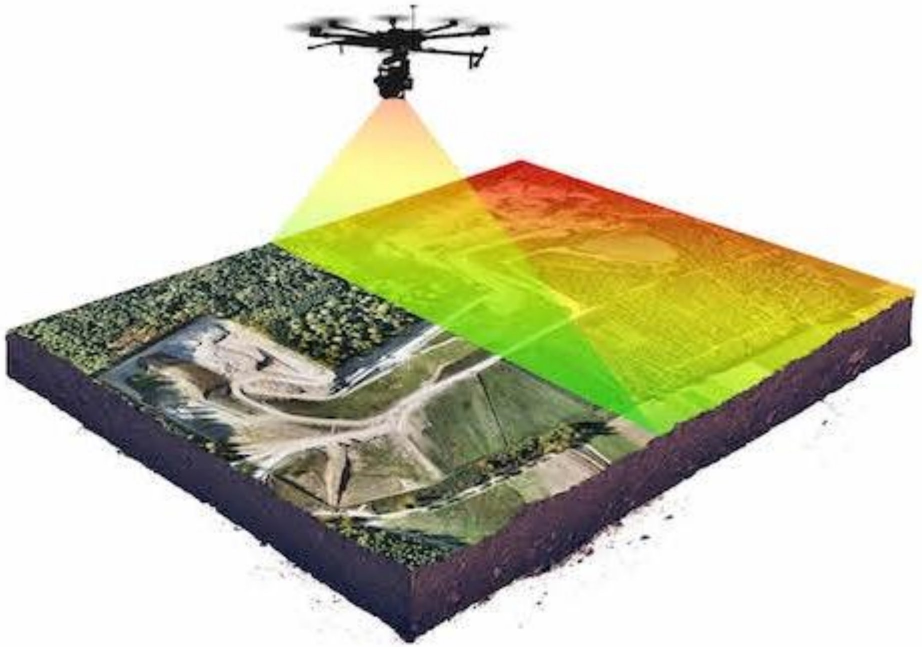
3D data representations Comparison

	Characteristics	Advantages	Challenges
Point Clouds	- Unordered and irregular data	- Direct representation of raw 3D measurements	- Lack of clear neighboring point definitions
	- Each point is a 3D tuple (x, y, z)	- Simple and intuitive	- Handling varying point counts in training data
	- Used in depth camera measurements	- Suitable for specific deep learning models (e.g., PointNet)	
Voxel	- Grid-like structure with voxels	- Regular structure allows convolutional filters	- Potential memory consumption issues
	- Truncated Signed Distance Functions (TSDFs)	- Suitable for deep learning models	- May require advanced memory optimization techniques
	- Positive/negative values indicate inside/outside	- Can represent detailed spatial information	
Mesh	- Vertices and faces defined on vertices	- Additional geometric, topological, and surface-normal info	- Mesh creation can be complex and time-consuming
	- Result from depth camera post-processing	- Enables graph-based deep learning models	- Handling varying vertex and face counts
	- Often designed manually during asset creation	- Represents detailed geometry	- More data to manage and process

Two ways to capture data from a drone: lidar (right)

Photogrammetry

Lidar

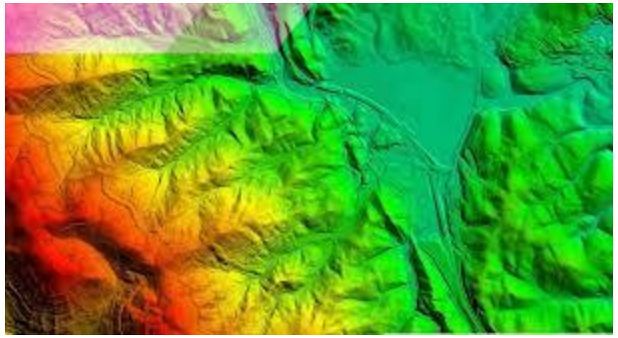


- A point cloud is a set of points in space that represent the surface of an object, which can be used for many different things.
- The points are typically created by a 3D scanning method, such as LiDAR or photogrammetry, which measures the distance between the scanner and nearby objects. These two methods measure the objects in question with different tools – LiDAR uses lasers, while photogrammetry uses photographs.
- Point clouds are also often used to create high-quality maps of city environments or to provide accurate topographical information about natural terrain. This data type is also often used as the foundation for creating digital versions of real-world objects – large and small.

3D Point Cloud vs Photo

Photogrammetry vs LIDAR

- Photogrammetry is the art of making a three dimensional scale model from a set of photographs taken from different angles.
- With the data gathered with photogrammetry technology, the end-user can:
  - Create a 3D model of the object being photographed
  - Calculate the size and shape of an object
  - Measure distances between objects
- Advantages
  - Low cost + Low Entry barrier + Useful for Fast moving projects such as those that involve cars or airplanes
  - Photorealistic interpretation of the original model – color, texture, shape
  - Photogrammetry also has lower error rates than many traditional measuring methods.
- Disadvantage
  - Relative inaccuracy when used with close-up photos (less than about 0.5 meters).
  - Significantly slower than LiDAR mapping
  - Lots of photos needed to create a 3D model



- LiDAR stands for Light Detection and Ranging .
- LiDAR uses lasers to measure distances from the sensor on the LiDAR device to objects in the environment.
- LiDAR uses laser light in the form of a narrow pulsed beam emitted from a single laser source.
- LiDAR has many significant advantages:
  - Can quickly process data and provides high-resolution 3D images of the surrounding environment, which is useful for mapping and surveying
  - detects objects or hazards in the environment that are not visible to the naked eye.
  - identifies obstacles or hazards that block the view of the road
  - creates highly accurate 3D maps which are used for localization and navigation. This is possible because LiDAR provides detailed information about surrounding objects on ground level
  - can be used as a scanning tool to inspect objects or infrastructure for damage or defects
- LiDAR disadvantages
  - Expensive technology to use. Now, low cost ($200 solid state lidars)
  - LiDAR setup is difficult to transport and set up in the field.
  - inaccurate distance measurements for terrestrial LiDAR comes from distortions in the laser beam caused by atmospheric turbulence and they cause the laser light to spread out as it travels through the atmosphere.
  - Fog and dust can affect the quality of the data collected making it harder to obtain accurate measurements
  - Rain water negatively affects the accuracy as the droplets cause the lasers to bounce back

Point Clouds: Usecases

❖ **Use cases for point clouds:**

- Land surveying: point clouds can help surveyors create a digital representation of a stretch of land, which holds all of the key elements – infrastructure, natural formations, roads, real estate and more
- Environmental monitoring: point clouds can help scientists and environmentalists detect the change in a piece of land over time by comparing point clouds of an eroding beach or glacier
- Structural inspections: insurance and real estate companies can review the deterioration of a building through the months or years, always keeping a digital copy on hand for comparison
- Historical site surveying: historians and archeologists can collect data and create digital copies of delicate ruins with remote data capture via LiDAR or photogrammetry, so as not to disrupt the site

References

- https://medium.com/@phamtdong0406/1-3d-ai-data-representation-99c98ec402c6
- https://www.creaform3d.com/sites/default/files/assets/technological-fundamentals/ebook1_an_introduction_to_3d_scanning_en_26082014.pdf (Good)
- SDF (signed distance function)
- Signed Distance Functions & Ray-Marching (very good)
- The SDF of a Box
- https://iquilezles.org/articles/distfunctions/ (Multiple Example 3D distance functions)
- https://iquilezles.org/articles/distfunctions2d/
- https://stanford.edu/~rqi/pointnet/
- https://stanford.edu/~rqi/pointnet/docs/cvpr17_pointnet_slides.pdf
- https://medium.com/@sim30217/truncated-signed-distance-function-f765a0f1d432
- https://avantierinc.com/resources/technical-article/3d-medical-imaging-revolution-optical-coherence-tomography-part-2/