Hardware systems for 3D data acquisition

Geometric Computer Vision

GCV v2021.1, Module 2

Alexey Artemov, Spring 2021

Lecture Outline

§1. Overview [10 min]

1.1. Classes and purposes of 3D acquisition systems

1.2. Contact methods vs. non-contact methods

1.3. Passive methods vs. active methods

§2. Passive multiple-view stereo [15 min]

2.1. Sparse and dense multiple-view stereopsis

Lecture Outline

§3. Active lightning [20 min]

3.1. Active stereo

3.2. Triangulation and Structured lighting systems

3.3. Time of flight (ToF) systems and LIDAR

§1. Overview

Classes and purposes of 3D acquisition systems

1.1. Classes and purposes of 3D acquisition systems

§1. Overview

- Analyze a real-world object or environment to collect data on its shape/appearance
  - Many technologies: contact, optical, computed tomography, structured light…
- **Today:** provide an overview of 3D scanning technologies
  - Taxonomy of systems
  - Physical principles and expected qualities
  - Scanners in action
  - Limitations and applications

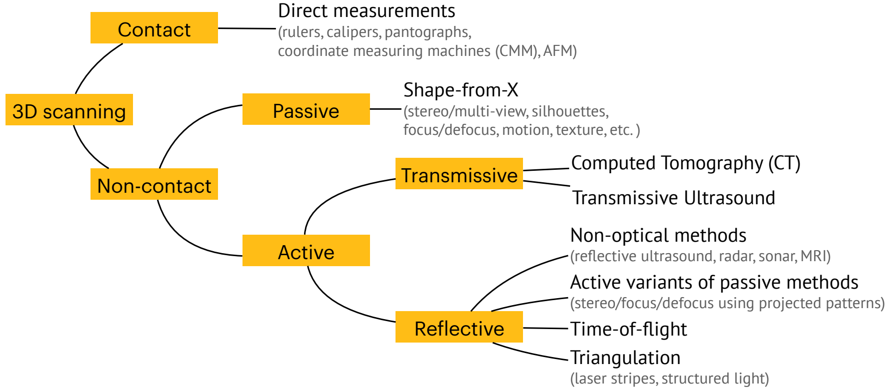
1.1. Classes and purposes of 3D acquisition systems

§1. Overview

- How do different types of 3D scanners work?
- What kinds of objects can be "scanned"?
- Spatial scale and temporal resolution
- Qualitative properties of compatible objects
- Physical principles leveraged
- Other types of variations (quantitative):
  - 3D resolution, 3D point precision

1.1. Classes and purposes of 3D acquisition systems

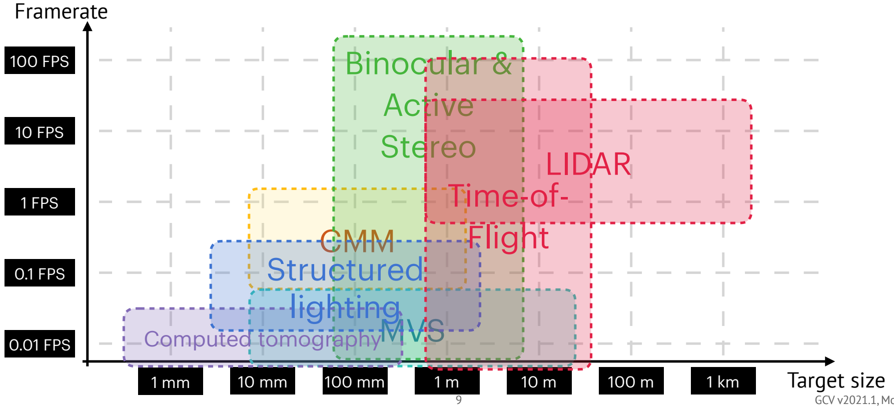
§1. Overview



Slide Credit: Lanman&Taubin

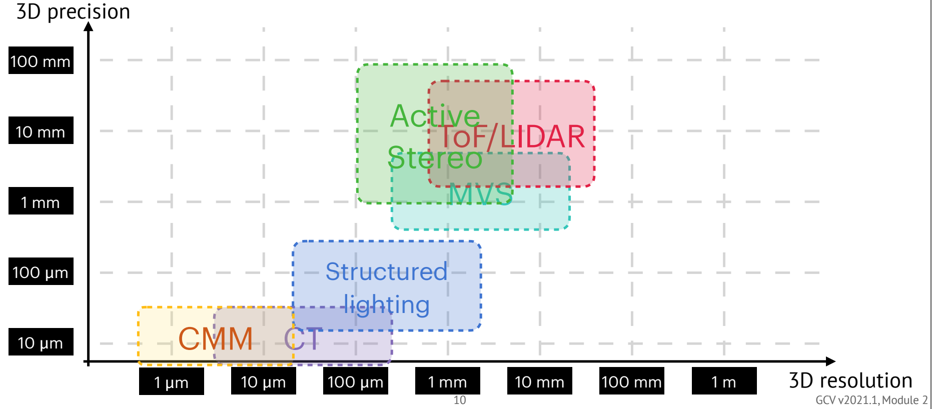
1.1. Spatial scale and temporal resolution

§1. Overview



1.1. 3D resolution and 3D point precision

§1. Overview



1.1. Compatible objects

Skolkovo Institute of Science and Technology

§1. Overview

Not compatible

- Contact: Fragile, Static only
- MVS: Reflective, Flat textures, Transparent, Periodic textures, Non-Lambertian, Noise
- Active stereo
- Structured Light: Static only
- ToF/LIDAR: multipath interference, Noise
- CT: Static only

1.1. Classes and purposes of 3D acquisition systems

§1. Overview

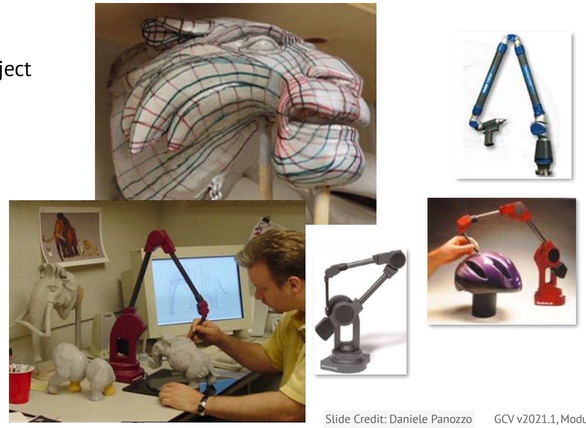
- Multiple technologies are able to deliver significantly different results
- Factors differentiating 3D scanning methods:
  - Spatial scale and temporal resolution
  - Qualitative properties of compatible objects
  - Physical principles leveraged
  - 3D resolution, 3D point precision

Contact methods vs. non-contact methods

1.2. Contact methods aka Touch probes

§1. Overview

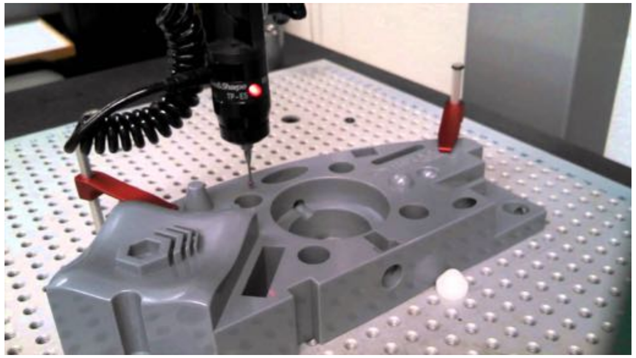
- Physical contact with the object
- Manual or computer-guided
- Advantages:
  - Can be very precise
  - Can scan **any** solid surface
- Disadvantages:
  - Slow, small scale
  - Can't use on fragile objects



Slide Credit: Daniele Panozzo

1.2. Contact methods aka Touch probes

§1. Overview

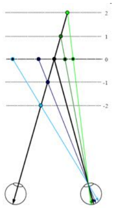


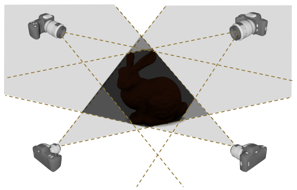
Passive methods vs. active methods

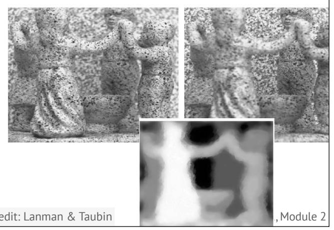
1.3. Passive methods vs. active methods

§1. Overview

- **Passive methods:** no direct control of illumination source (rely on ambient light)
- Mainstream: stereoscopic imaging (MVS, SfM, SLAM…)
  - Requires **correspondences** to be found among the various viewpoints (hard!)
- Bypass the correspondence search: silhouette, focus, defocus



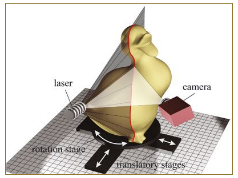


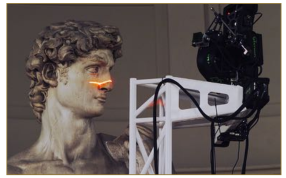


1.3. Passive methods vs. active methods

§1. Overview

- **Active methods:** control illumination in some form
  - Project visible or invisible textures / coded images
  - Measurement of time taken by light to reach target objects





§2. Passive multiple-view stereo

MVS/SfM: motivation

§2. Passive multiple-view stereo

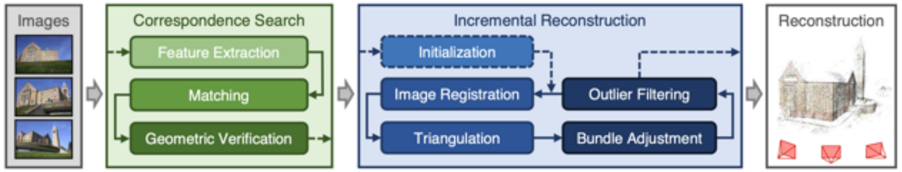
- Passive, no special equipment, anyone can use it
- Leverage large-scale Internet photo collections (uncalibrated, unstructured)
- Calibrated setups with controlled/known lighting: high-quality reconstruction

- **Multiple-view stereo (MVS)/Structure-from-Motion (SfM):** computer vision terms
- **Simultaneous localisation and mapping (SLAM):** robotics term

MVS/SfM: pipeline overview

Skoltech
Skolkovo Institute of Science and Technology

§2. Passive multiple-view stereo



Picture Credit: SfM Revisited

- Reconstruct 3D structure from projection into multiple images (fully passive)
- **Correspondence search:** find projections of the same points in overlapping images
- **Reconstruction (sparse):** estimate camera poses and 3D feature locations
- **Reconstruction (dense):** estimate consistent (across views) per-pixel depth

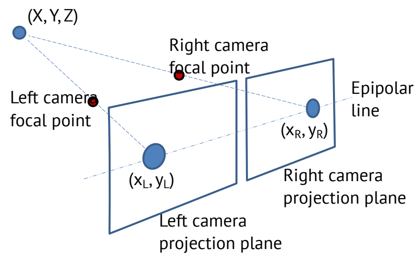
MVS/SfM: triangulation

Skoltech
Skolkovo Institute of Science and Technology

§2. Passive multiple-view stereo

- (X, Y, Z): 3D point in scene
- (x_L, y_L), (x_R, y_R): 2D points in left and right images, respectively
- Epipolar constraints: $y_L = y_R$
- Calibration: known baseline B, focal length f
- Disparity: $d = x_R - x_L$
- Estimate depth from similar triangles:

$$Z = \frac{B \cdot f}{d}$$

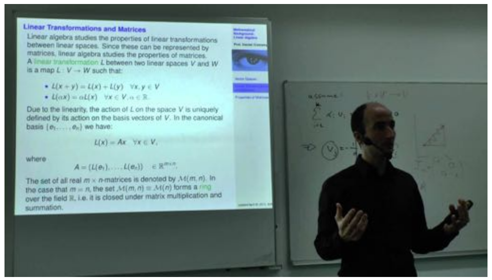


MVS/SfM: triangulation

Skoltech
Skolkovo Institute of Science and Technology

§2. Passive multiple-view stereo

- Prof. D. Cremers: Multiple View Geometry (YouTube playlist with 14 lectures on SfM)

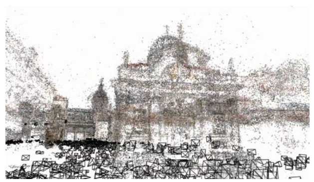


MVS/SfM: example reconstructions

Skoltech
Skolkovo Institute of Science and Technology

§2. Passive multiple-view stereo

- [See accompanying video for a real-life demo]
- [Sparse reconstruction: the city of Dubrovnik]
- [Dense reconstruction: Moscow historical cites]

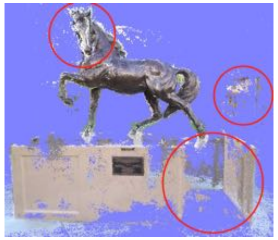


MVS/SfM: limitations

Skoltech
Skolkovo Institute of Science and Technology

§2. Passive multiple-view stereo

- **Problems:** textureless or specular regions
  - Nothing to match between different views, or something to match looks differently from different views



Reconstruction by [Yu and Gao, CVPR'20]

MVS/SfM: discussion

Skoltech
Skolkovo Institute of Science and Technology

§2. Passive multiple-view stereo

- Infer the geometry from light reflectance
- **Advantages:**
  - Less invasive than touch
  - Fast, large scale possible
- **Disadvantages:**
  - Difficulty with transparent and shiny objects

Slide Credit: Daniele Panozzo

Skoltech
Skolkovo Institute of Science and Technology

§3. Active lighting

Skoltech
Skolkovo Institute of Science and Technology

Active stereo

3.1. Active stereo

§3. Active lighting

- **Active stereo:** use an optical projector to overlay the scene with a semi-random texture
- **Enhance stereo matching** by providing a rich source of correspondences for dense stereo methods
  - Flat/smooth surfaces
  - Textureless surfaces (e.g. white walls)
- Related to **unstructured light**
- Outputs: **dense range-images**

3.1. Active stereo: unstructured light

§3. Active lighting

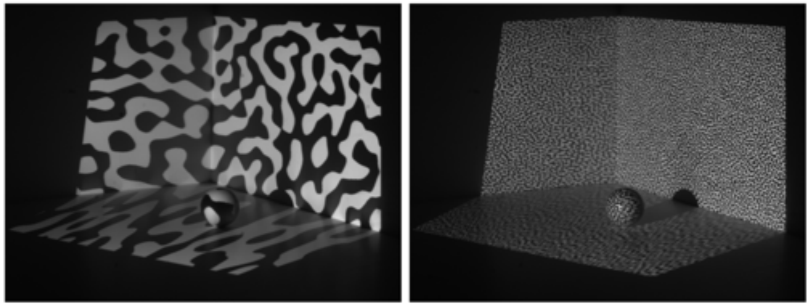


Figure Credit: UMontreal

- Two patterns with different vertical frequencies projected on a scene
- Dense texture, photometrically consistent, no repetition along X (effectively random)

3.1. Active stereo: unstructured light

§3. Active lighting

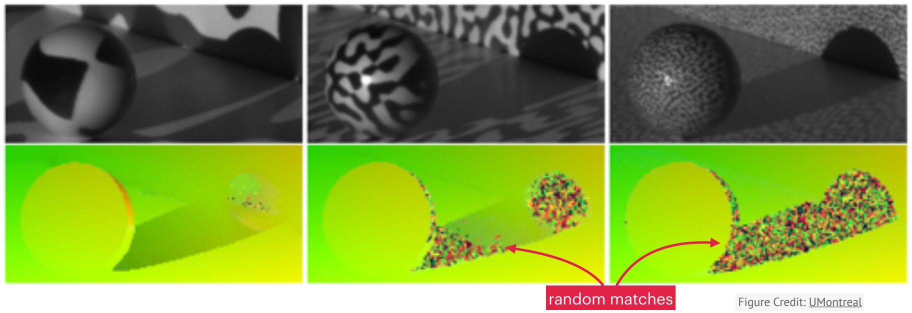


Figure Credit: UMontreal

- Correspondence from unstructured patterns at frequencies 8 (left), 32 (middle) and 128 (right)

3.1. Active stereo: Intel RealSense

§3. Active lighting

Figure Credit: Intel

- Example device: **Intel RealSense D435**
- Mass-market, ~200 USD
- https://dev.intelrealsense.com/docs/stereo-depth-camera-d400
- Technology overview:
  - IR projector (850–854 nm band)
  - Left-right cameras: 10bit, VGA (640x480), 30–90 Hz
  - Color camera: Full HD (1920x1080)
  - Stereo algorithm: Census cost function, 64 disparities

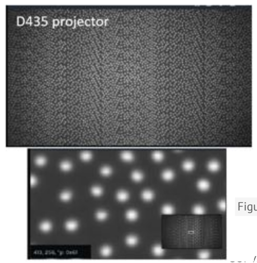
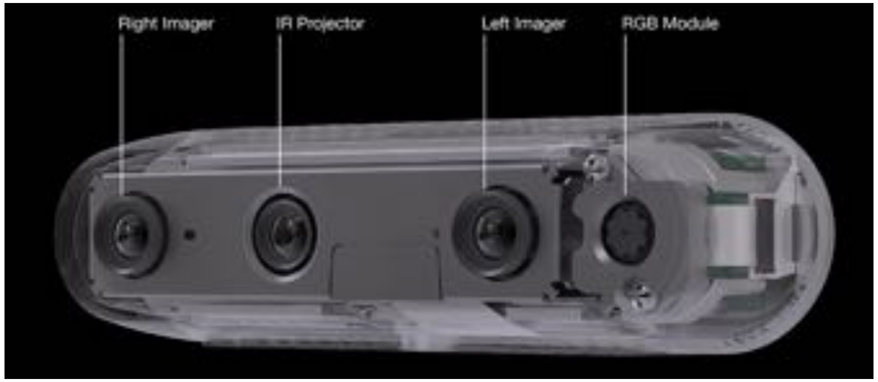


Figure Credit: Intel

3.1. Active stereo: Intel RealSense

Skoltech
Skolkovo Institute of Science and Technology

§3. Active lighting



3.1. Active stereo: example

Skoltech
Skolkovo Institute of Science and Technology

§3. Active lighting

- [See accompanying video for a real-life demo]

3.1. Active stereo: discussion

Skoltech
Skolkovo Institute of Science and Technology

§3. Active lighting

- **Active stereo:** use an optical projector to overlay the scene with a semi-random texture
- **Advantages:**
  - improved correspondence search
  - works in 0 lux environments and in broad daylight
- **Disadvantages:**
  - laser speckle pattern inconsistent between views → noise
  - unable to function without a strong infrared light source

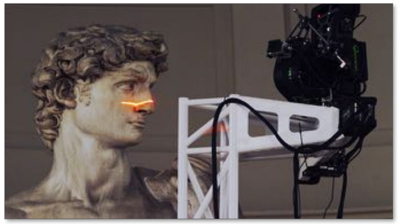
Skoltech
Skolkovo Institute of Science and Technology

Triangulation

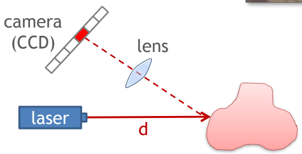
3.2. Triangulation

§3. Active lighting

- Laser beam and camera
- Laser dot is photographed
- The location of the dot in the image allows triangulation: we get the distance to the object



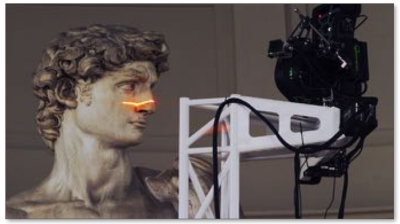
Digital Michelangelo Project



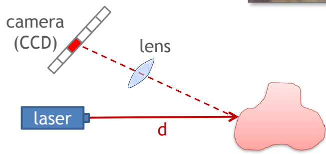
3.2. Triangulation

§3. Active lighting

- Laser beam and camera
- Laser dot is photographed
- The location of the dot in the image allows triangulation: we get the distance to the object



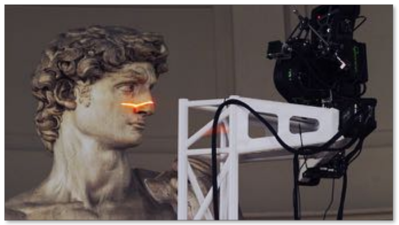
Digital Michelangelo Project



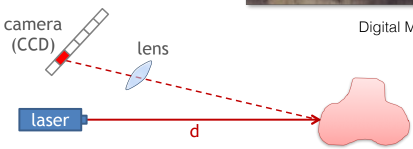
3.2. Triangulation

§3. Active lighting

- Laser beam and camera
- Laser dot is photographed
- The location of the dot in the image allows triangulation: we get the distance to the object



Digital Michelangelo Project



3.2. Triangulation

§3. Active lighting

- Very precise (tens of microns)
- Small distances (meters)

Skoltech
Skolkovo Institute of Science and Technology

Structured lighting

Skoltech
Skolkovo Institute of Science and Technology

3.2. Structured lighting

§3. Active lighting

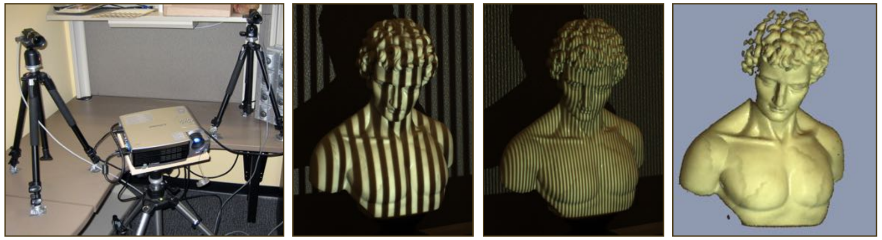
- Swept-plane scanning recovers 3D depth using ray-plane intersection
- **Disadvangate:**
  - Slow scanning due to manually-/mechanically-swept laser/shadow planes
  - Needs hundreds of images
- **Use a data projector to replace mechanics!**

Skoltech
Skolkovo Institute of Science and Technology

3.2. Structured lighting

§3. Active lighting

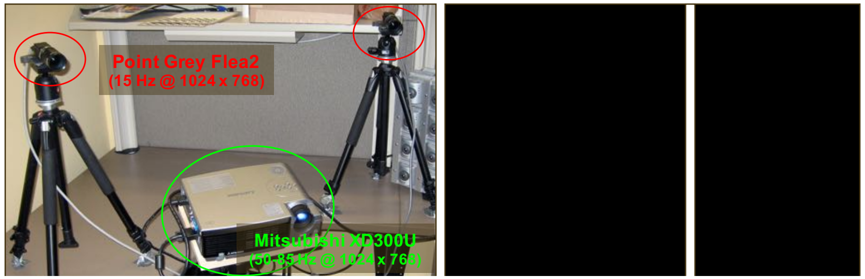
- **Structured-lighting:** measure the 3D shape using projected light and a camera system
  - Display 2-dimensional patterns
  - Extract more information → less images, faster scanning



Skoltech
Skolkovo Institute of Science and Technology

3.2. Structured lighting: swept-planes

§3. Active lighting



- How to assign correspondence from projector planes to camera pixels?
- Solution: Project a spatially- and temporally-encoded image sequence
- **What is the optimal image sequence to project?**

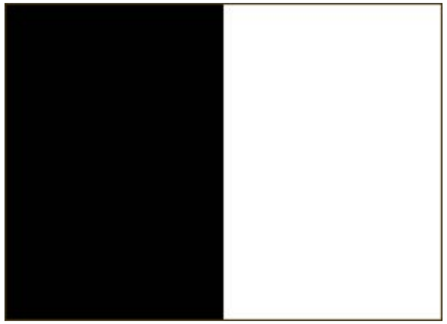
3.2. Structured lighting: binary codes

Skoltech
Skolkovo Institute of Science and Technology

§3. Active lighting

Binary Image Sequence [Posdamer and Altschuler 1982]

- Each image is a bit-plane of the binary code for projector row/column
- Minimum of 10 images to encode 1024 columns or 768 rows
- In practice, 20 images are used to encode 1024 columns or 768 rows
- Projector and camera(s) must be synchronized



3.2. Structured lighting: Gray codes

Skoltech
Skolkovo Institute of Science and Technology

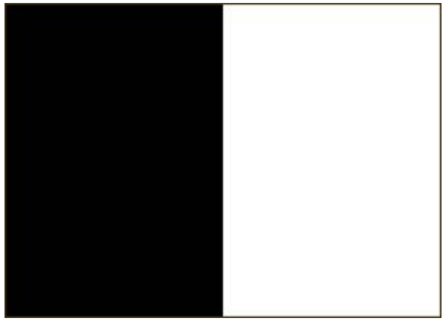
§3. Active lighting

Gray Code Image Sequence [Inokuchi 1984]

- Each image is a bit-plane of the Gray code for each projector row/column
- Requires same number of images as a binary image sequence, but has better performance in practice

```
BIN2GRAY(B)
n ← length[B]
G[1] ← B[1]
for i ← 2 to n
    do G[i] ← B[i − 1] xor B[i]
return G
```

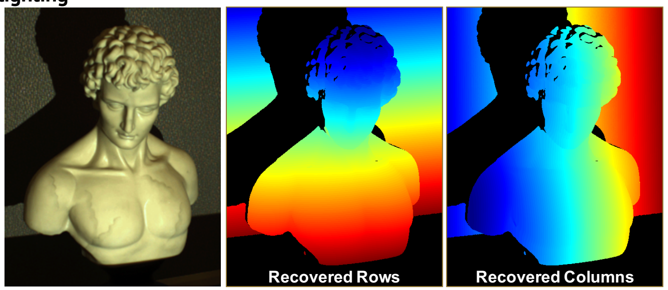
```
GRAY2BIN(G)
n ← length[G]
B[1] ← G[1]
for i ← 2 to n
    do B[i] ← B[i − 1] xor G[i]
return B
```



3.2. Structured lighting: swept-planes

Skoltech
Skolkovo Institute of Science and Technology

§3. Active lighting



- 42 images (2 to measure dynamic range, 20 to encode rows, 20 to encode columns)
- Individual bits assigned by detecting if bit-plane (or its inverse) is brighter
- Decoding algorithm: Gray code → binary code → integer row/column index

3.2. Structured lighting: example

Skoltech
Skolkovo Institute of Science and Technology

§3. Active lighting

- [See accompanying video for a real-life demo]

3.2. Structured lighting: discussion

Skoltech
Skolkovo Institute of Science and Technology

§3. Active lighting

- Pattern of visible (or infrared) light is projected onto the object
- The distortion of the pattern, recorded by the camera, provides geometric information
- **Advantages:**
  - Very fast – 2D pattern at once
  - Even in real time, like Kinect 1.0
- **Disadvantages:**
  - Complex distance calculation, prone to noise
  - Affected by indirect illumination scattered from directly-illuminated points
  - Not real-time, ill-suited for dynamic scenes

Skoltech
Skolkovo Institute of Science and Technology

Time of flight

3.3. Time-of-flight measurements

Skoltech
Skolkovo Institute of Science and Technology

§3. Active lighting

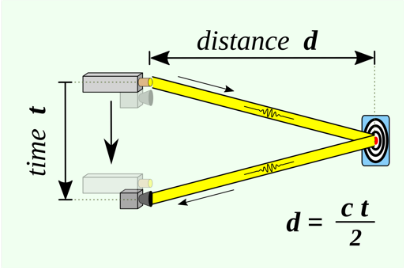
- Previous scanners:
  - Unable to scan dynamic scenes (needed by robotics)
  - Occluded regions not recovered (light source away from camera)
- Time-of-flight: estimate distance to a surface from a single center of projection

3.3. Time-of-flight

Skoltech
Skolkovo Institute of Science and Technology

§3. Active lighting

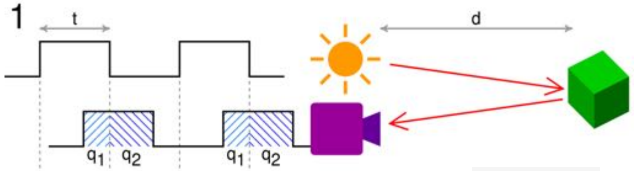
- **Radar: RA**dio **D**etection **A**nd **R**anging
  - Emit radio wave signal and receive reflected signal to estimate range
- **LIDAR: Li**ght **D**etection **A**nd **R**anging
  - Use **laser** light (spatially coherent – focused to a tight spot)
- **Time-of-flight (ToF):** measure the time taken to travel a distance through a medium
- Sensors: scanning LIDAR, ToF camera



Picture Credit: Wikipedia

3.3. Time-of-flight: light pulse

§3. Active lighting



Picture Credit: Wikipedia

- **Light pulse ToF:**
  - switch illumination on for a short time, read light reflected by objects

$$d = \frac{ct}{2} \frac{q_2}{q_1 + q_2}$$

- Need very precise delay measurement!

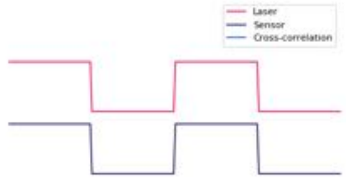
$$t_D = 2 \cdot \frac{D}{c} = 2 \cdot \frac{2.5\ \mathrm{m}}{300\ 000\ 000\ \frac{\mathrm{m}}{\mathrm{s}}} = 16.66\ \mathrm{ns}$$

Skoltech
Skolkovo Institute of Science and Technology

3.3. Time-of-flight: light pulse

§3. Active lighting

- **Time delay measurement:** cross-correlate the reference waveform with the measured optical signal and find the peak

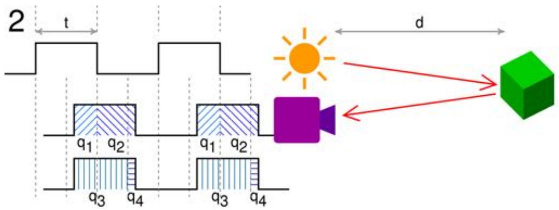


Picture Credit: Medium

Skoltech
Skolkovo Institute of Science and Technology

3.3. Time-of-flight: amplitude modulated

§3. Active lighting



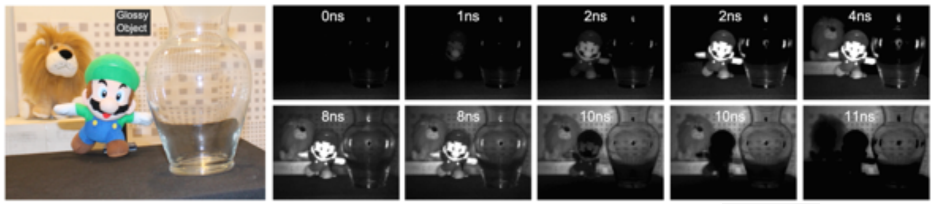
- **Amplitude modulation of a continuous wave (AMCW) ToF:** operate continuously, modulate amplitude
- Compute phase difference between the emitted and reflected signals

Skoltech
Skolkovo Institute of Science and Technology

3.3. Time-of-flight: nanophotography

§3. Active lighting

- Perform visualization of light sweeping over the scene
- Time resolution: 69.4 picoseconds (calculated theoretical best-case)



Picture Credit: MIT

3.3. Time-of-flight: nanophotography

§3. Active lighting

Video Credit: MIT

3.3. Time-of-flight: example

§3. Active lighting

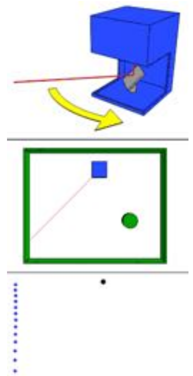
- [See accompanying video for a real-life demo]

LIDAR

3.3. LIDAR

§3. Active lighting

- Use laser illumination:
  - Tight focus over long distances → longer scanning range, higher spatial resolution
  - Narrow radiation spectrum → immune to ambient illumination in the environment
- Common technical architecture: a system of rotating mirrors

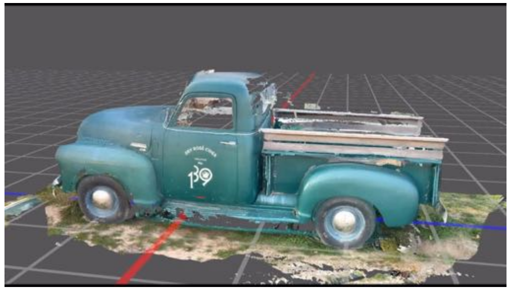


Picture Credit: Wikipedia

3.3. LIDAR: applications

§3. Active lighting

Autonomous driving



3D scanning (iPad 12 Pro)

3.3. Time-of-flight and LIDAR: discussion

§3. Active lighting

- Radar-based: compute time taken by radio waves / laser pulse to travel to/from objects
- **Advantages:**
  - Real-time possible
  - (Very) Large-scale possible
  - Immune to ambient illumination: day and night enabled
- **Disadvantages:**
  - In real-time: too sparse for 3D scanning
  - Prone to reflections from dust / smoke / snow / rain

References

1. Taubin, G., Moreno, D., & Lanman, D. (2014). *3d scanning for personal 3d printing: build your own desktop 3d scanner.* In ACM SIGGRAPH 2014 Studio (pp. 1-66).