Optical Range Sensors

Andrea Fusiello
DPIA - Università di Udine
http://www.diegm.uniud.it/fusiello/

Range sensors

- **Range sensors** (or **range finders**) are devices that capture the three-dimensional (3-D) structure of the world from the viewpoint of the sensor, usually measuring the depth to the nearest surfaces.
- These measurements could be at a single point, across a scanning plane, or a full image with depth measurements at every point.
- The benefits of this range data is that a robot can reliably find navigable routes, avoid obstacles, grasp objects, act on industrial parts, etc.

Range data

- Range data is a 2½D or 3D representation of the scene around the robot.
- The 3D aspect arises because we are measuring the (X,Y,Z) coordinates of one or more points in the scene.
- We only observe the front sides of objects - the portion of the scene visible from the robot. This is the origin of the term 2½D.

Figures of merit

- Spatial resolution (e.g. 64x64 1024x1024)
- Range (depth) resolution (0.01 mm, 10 cm)
- Operating range (10cm, 100m)
- Frame rate (30-50 fps)

- Other aspects:
  - Cost
  - Size, weight
  - Power consumption

Summary

- Stereo vision

- Active triangulation
  - Laser scanner
  - Coded light

- Time of fight
  - PW / CW scanner
  - TOF camera

} Active sensors

PW = Pulse Wave
CW = Continuous Wave

Scanner vs full-field

- Scanning: several depth measures are taken sequentially in time to cover all the scene; typ. a beam (or plane) of light is swept across the scene (**scanner**);

- Full field: measure all the scene in a single shot (**range camera**)

Active vs passive sensors

- If system only receives light rays and does not emit any radiation it is a **passive sensor** (e.g. stereo)
- If emits radiation then it is an **active sensor**

- Two principles of active systems:
  - **Triangulation:** same as stereo but a light source replaces the second camera.
  - **Time of flight:** produce a pulsed beam of light, measure distance by time light takes to return.

Summary

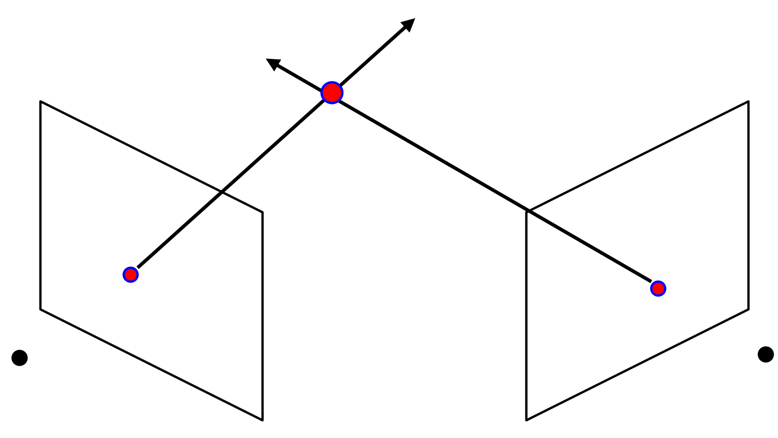
- Stereo vision

- Active triangulation
  - Laser scanner
  - Coded light

- Time of fight
  - PW / CW scanner
  - TOF camera

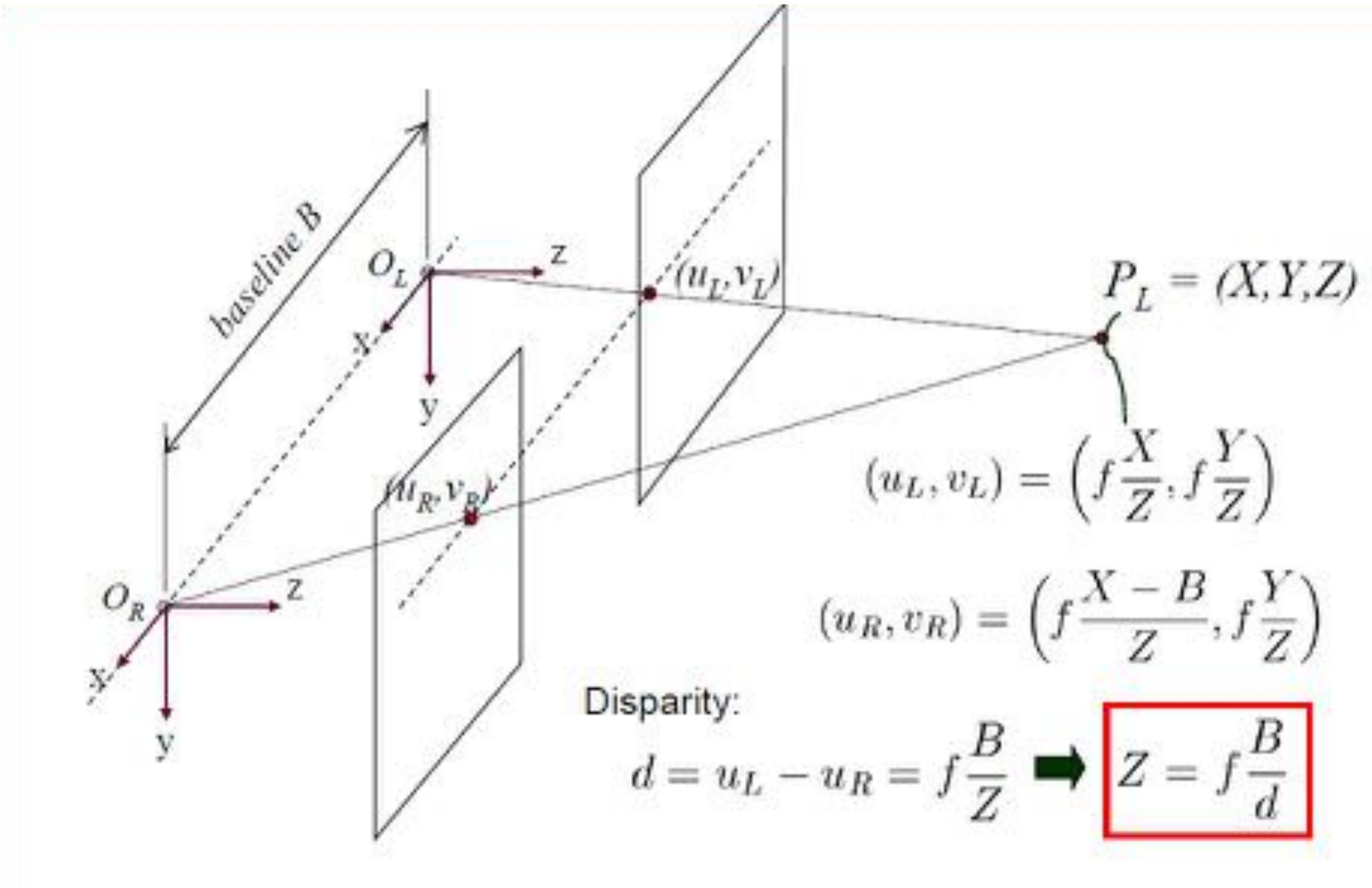
Stereo vision

- Stereo analysis uses two or more input images to estimate the distance to points in a scene.
- The basic concept is **triangulation**: a scene point and the two camera points form a triangle, and knowing the baseline between the two cameras, and the angle formed by the camera rays, the distance to the object can be determined.



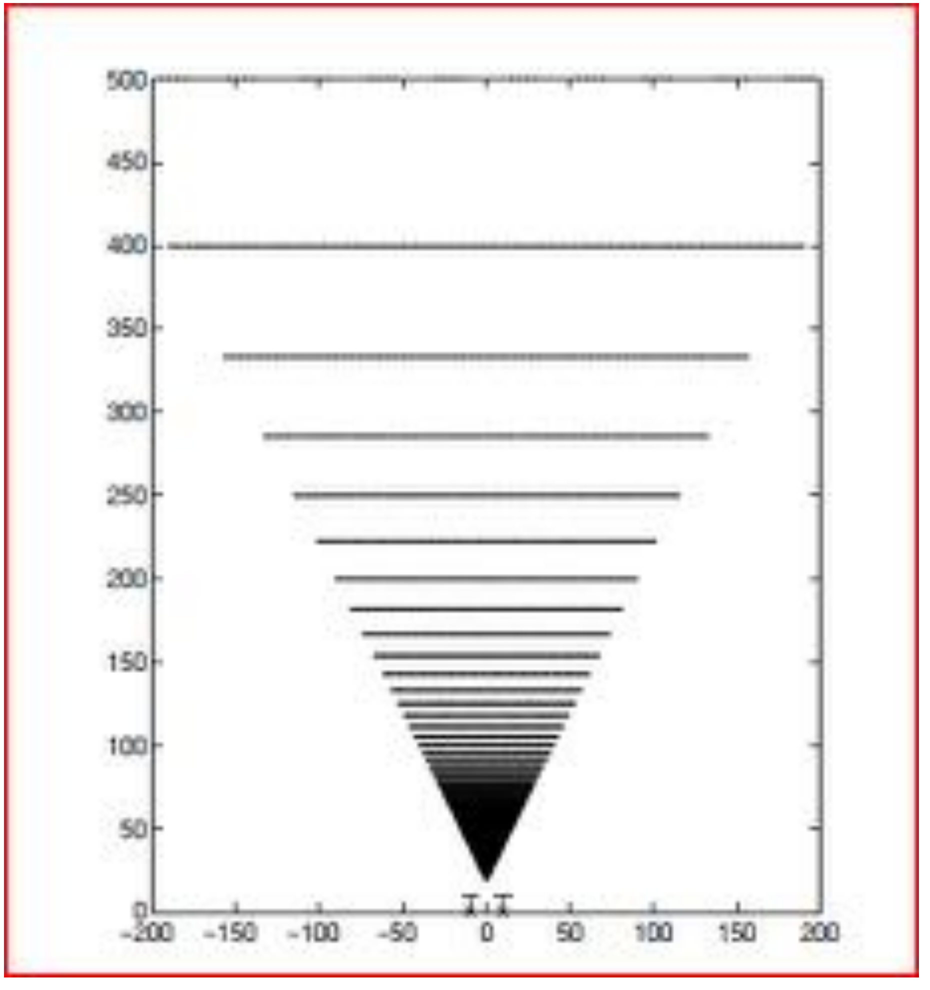
From[1]

Normal case for stereo



From[1]

Iso-disparity planes



Lines where the disparity value has the same value (see stereo-res.jpg)

Notice the rapid Increase in spacing

Implication is that the resolution of stereo depth depends on the disparity

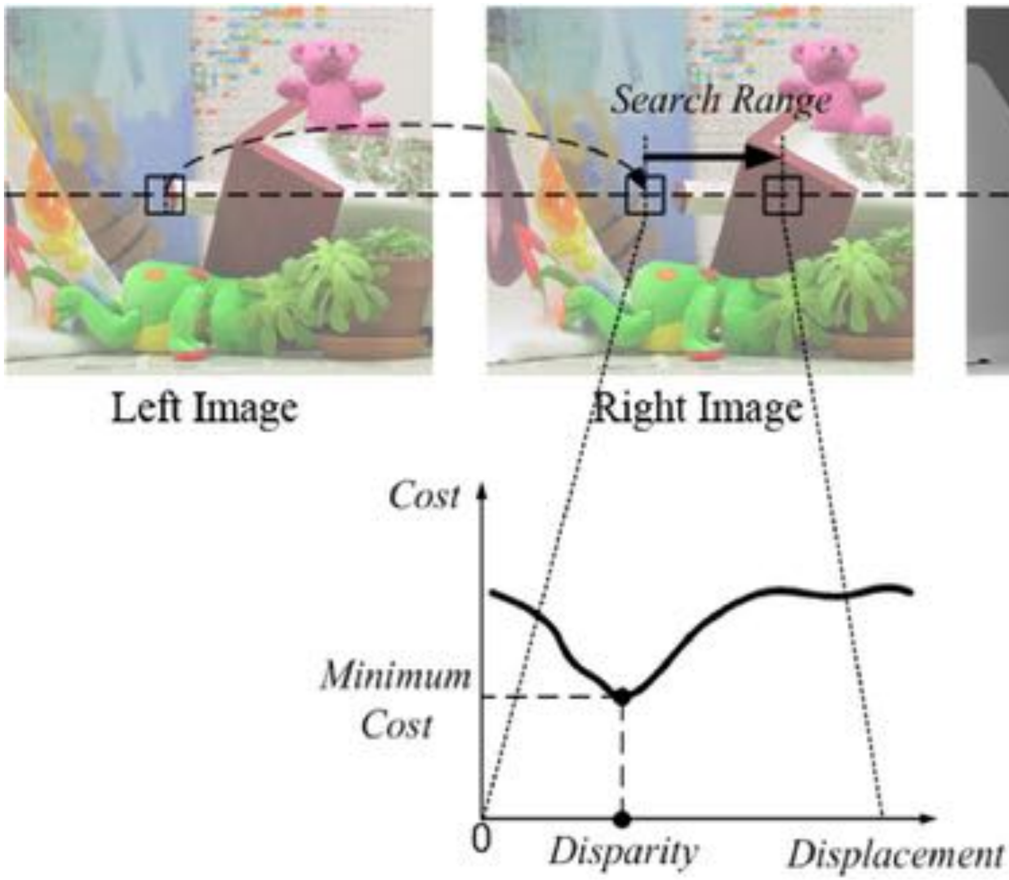
From[1]

Stereo matching

- The fundamental problem in stereo analysis is **matching** image elements that represent the same object -- or object part -- in the scene.
- Once the match is made, the range to the object can be computed using the image geometry.

Images from S. Mattoccia

Stereo matching

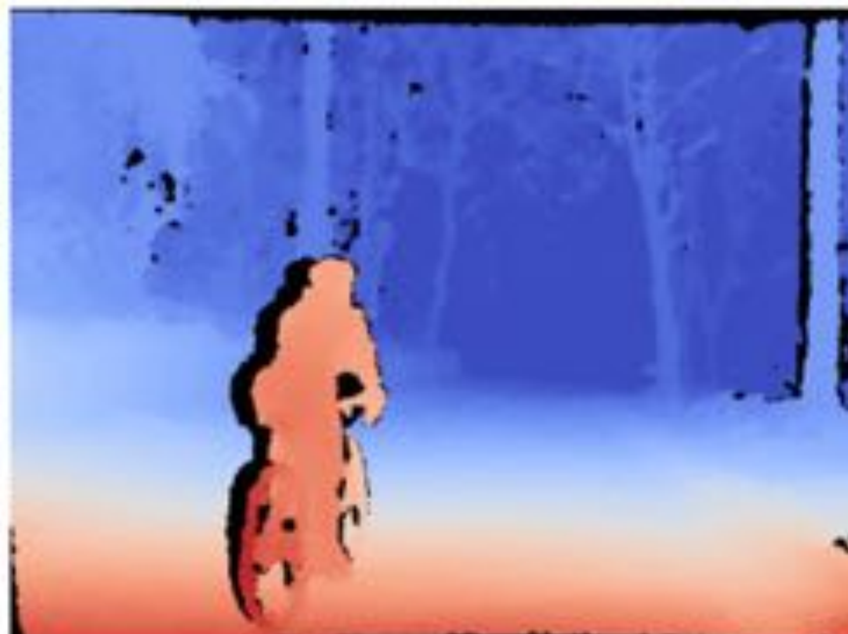


Normalized cross-correlation.

$$\frac{\sum_{x,y}[I_{x,y} - \hat{I}_{x,y}][I'_{x-d,y} - \hat{I'}_{x-d,y}]}{\sqrt{\sum_{x,y}[I_{x,y} - \hat{I}_{x,y}]^2 \sum_{x,y}[I'_{x,y} - \hat{I'}_{x-d,y}]^2}}$$

Images from [Young-Ho Seo, Ji-SangYoo, Dong-WookKim](#) A new parallel hardware architecture for high-performance stereo matching calculation

Nerian SceneScan + Karmin2



Nerian SceneScan + Karmin2

Summary

- Stereo vision
- Active triangulation
  - Laser scanner
  - Coded light
- Time of fight
  - PW / CW scanner
  - TOF camera

Active triangulation

- Active triangulation range sensors are based on principles similar to stereo, but with **one camera and a light source.**
- A laser beam is projected from one position onto the observed surface. The light spot that this creates is observed by a camera from a second position
- Still use **triangulation** to find the depth
- Many possible light sources variations

Laser: pros…

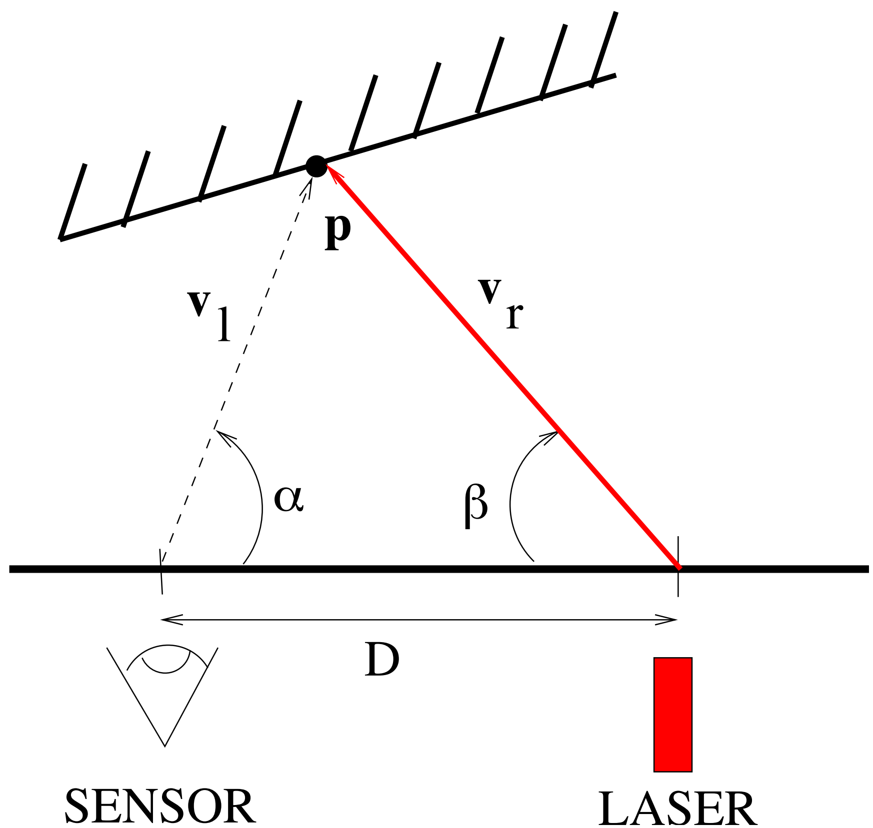
- Laser can easily generate bright beams with lightweight sources,
- infrared beams can be used unobtrusively,
- focus well to give narrow beams,
- single frequency sources allow easier rejection filtering of unwanted frequencies,
- single frequency sources do not disperse from refraction as much as full spectrum sources,
- semiconductor devices can more easily generate short pulses, etc

… and cons

- One disadvantage of all active sensor types is specular reflections. The normal assumption is that the observed light is a diffuse reflection from the surface. If the observed surface is specular, such as polished metal or water, then the source illumination may be reflected in unpredictable directions.
- A second problem is the laser ‘footprint’. Because the laser beam has a finite width, when it strikes at the edge of a surface, part of the beam may be actually lie on a more distant surface.

Active triangulation

- Knowing the **relative positions and orientations of the laser and sensor**, some simple trigonometry allows calculation of the 3D position of the illuminated surface point.



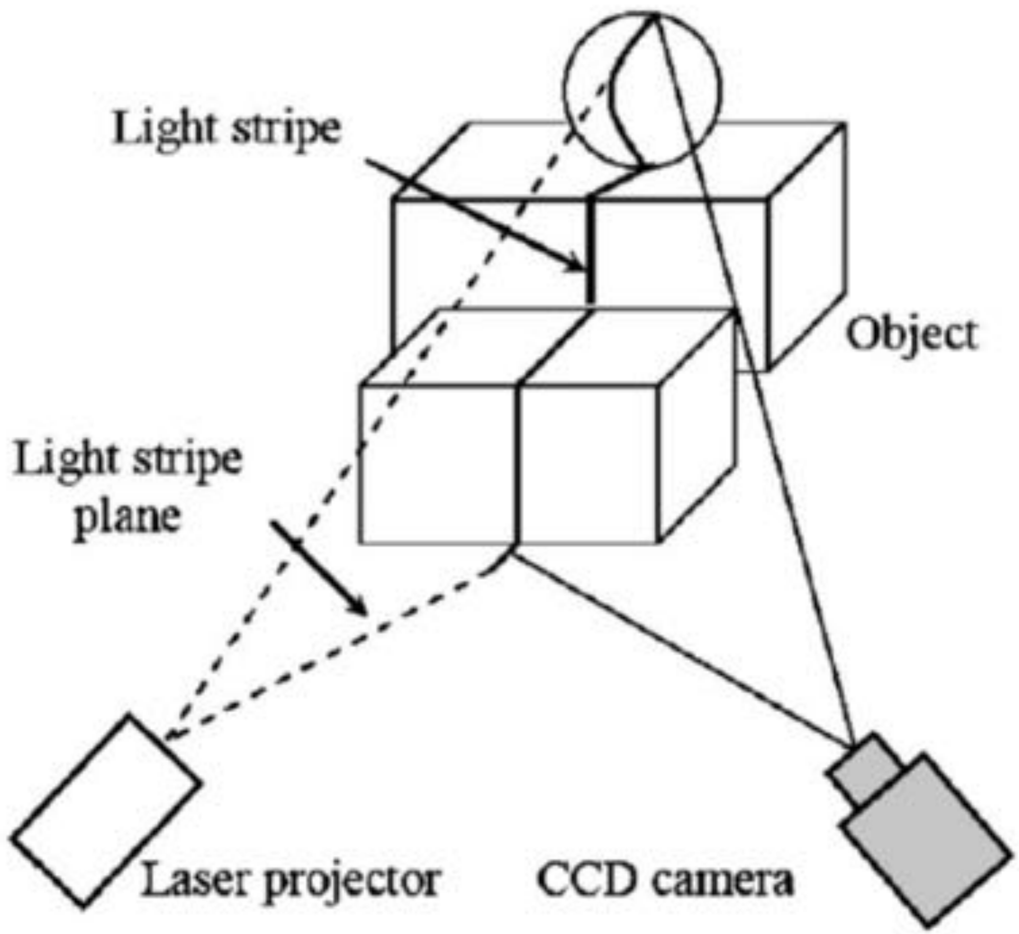
Spreading out the spot

- Move the laser spot over the scene with a **spinning mirror** to gather (e.g.) thousands of points per second. → Scanner
- The laser spot can be reshaped with lenses or mirrors to create multiple spots or stripes, thus allowing a measurement of multiple 3D points simultaneously.
- **Stripes** are commonly used because these **can be swept across the scene**. → Scanner
- Other illumination patterns are also commonly used, such as parallel lines, concentric circles, cross hairs and dot grids. → Structured/coded light

Depth from a laser line (structured light)

Process each horizontal line in the camera

Works, but still only for one line at a time



From[1]

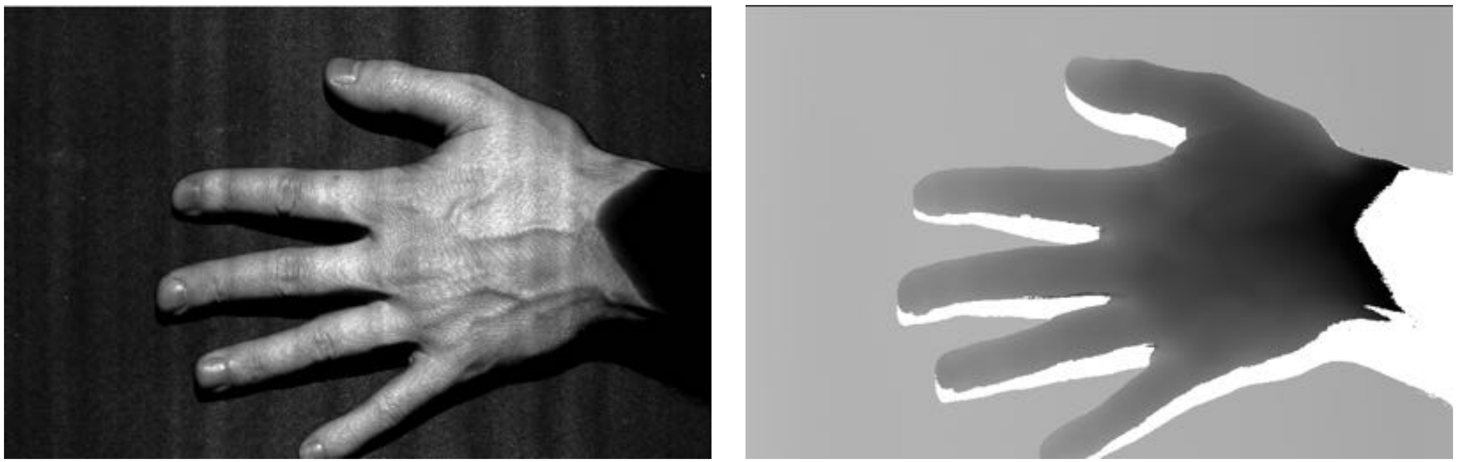


Figura 102. – Immagine di intensità ed immagine range ottenute da un sistema a triangolazione attiva commerciale con lama laser. Le parti mancanti nella immagine range (in bianco) sono dovute alla diversa posizione della sorgente laser e della fotocamera. Per cortesia di S. Fantoni

Triangulation laser scanner

Shape Grabber (Vitana)

Minolta

Mensi

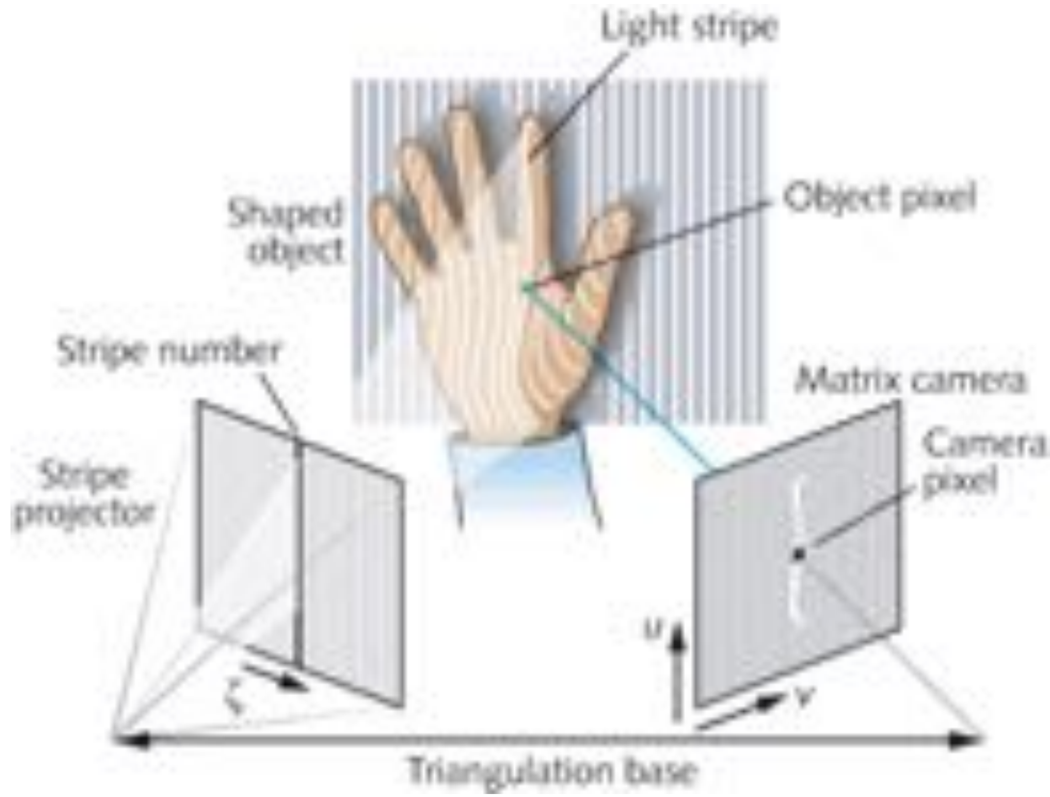
From[3]

Minolta Vivid 910

- A commercial scanner, with high precision (0.2-0.3mm), but high cost (>30K euros).
- A laser line is swept over the object: 300K points are measured in 2.5 seconds.

From[2]

Multi stripe projector

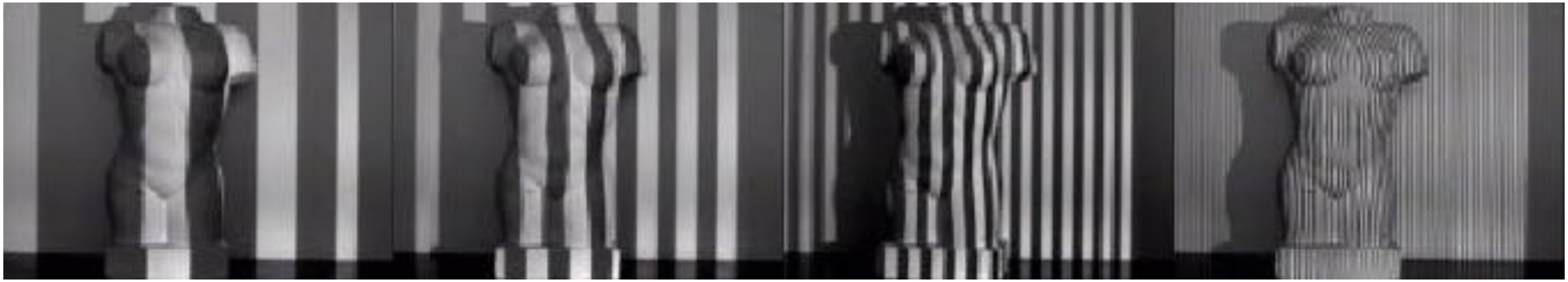


- To go faster project multiple stripe, a.k.a. structured light
- But which stripe is which?
- Need some coding to differentiate the stripes

Structured/coded light

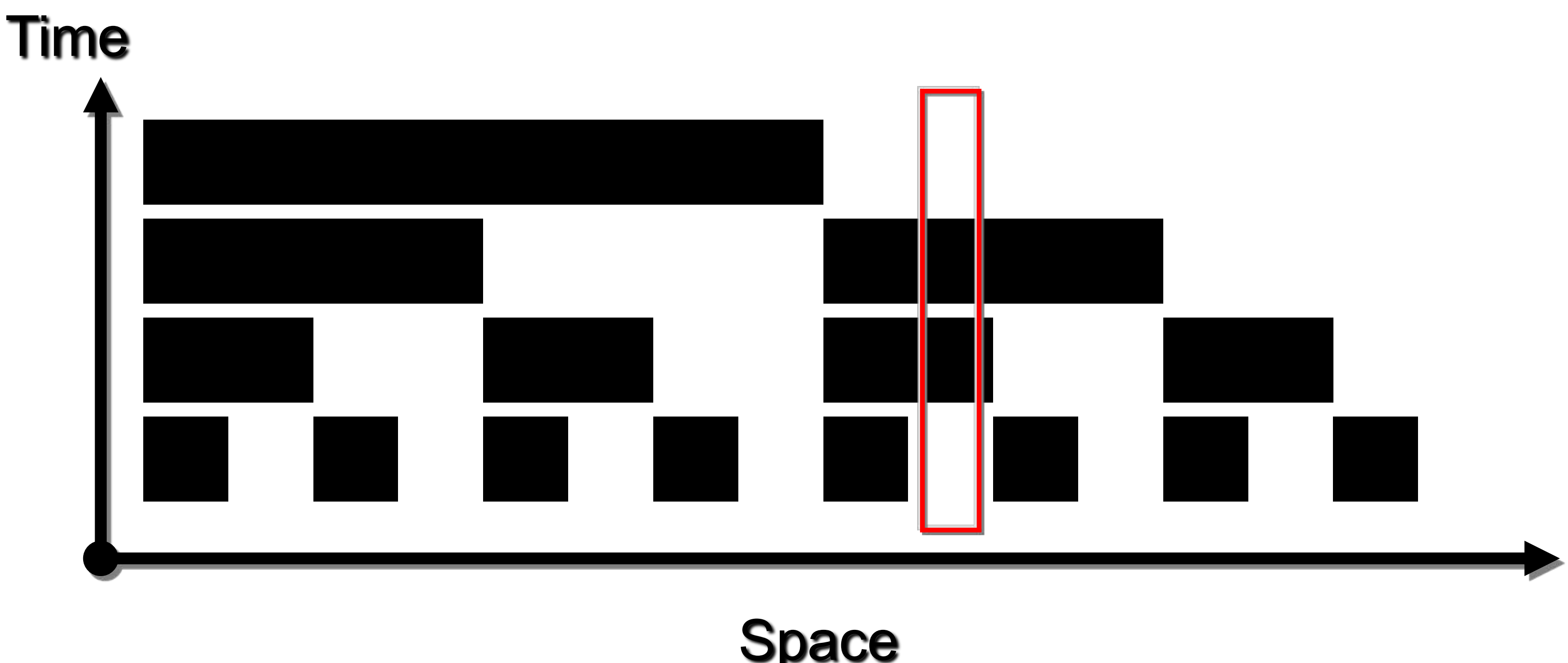
- Solutions:
  - No coding, assume surface continuity
  - Color coding

  - Temporal coding



Time-Coded Light Patterns

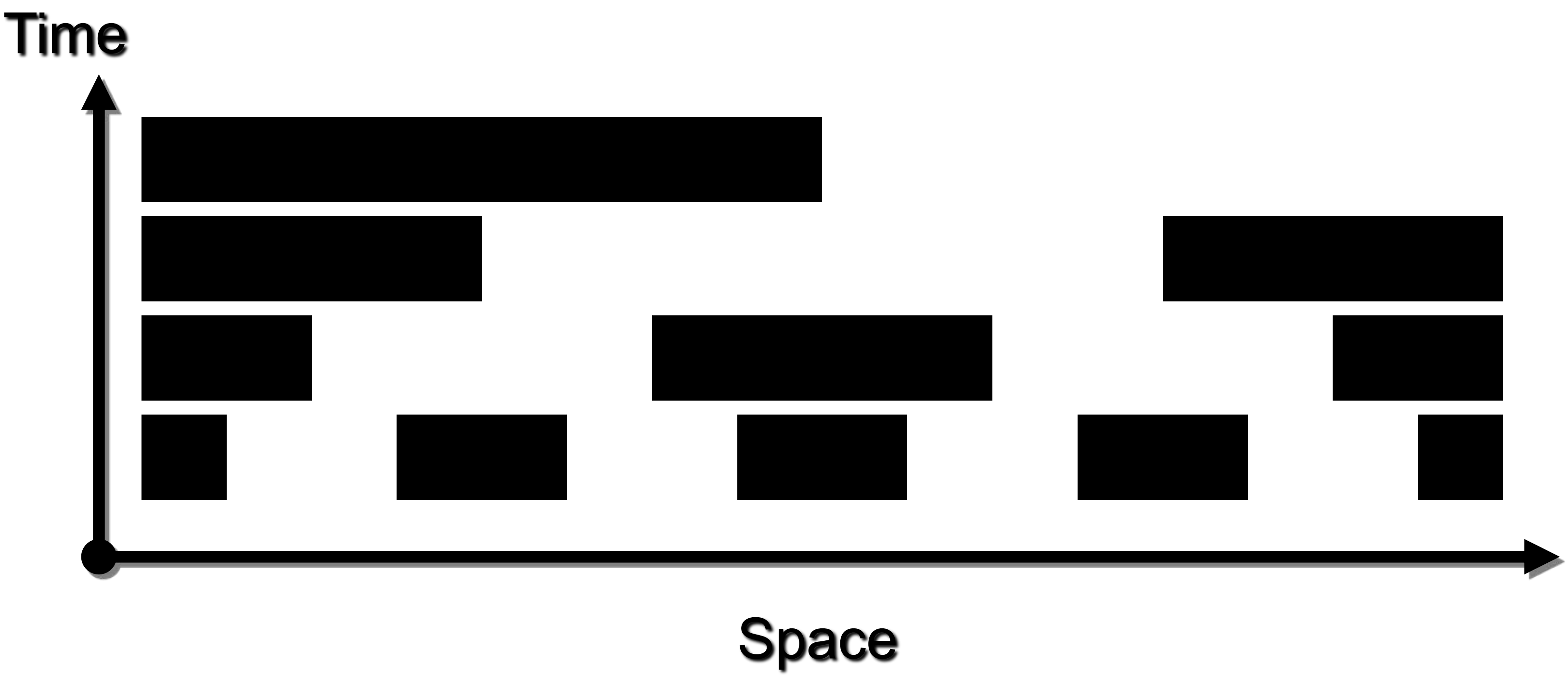
Assign each stripe a unique illumination code over time [Posdamer 82]



From[1]

Gray-Code Patterns

To minimize effects of quantization error:
each point may be a boundary only once

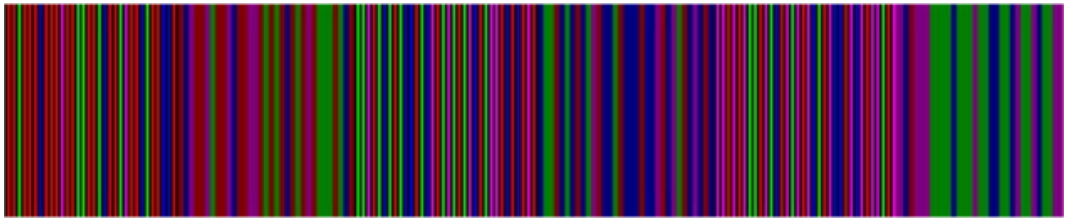


From[1]

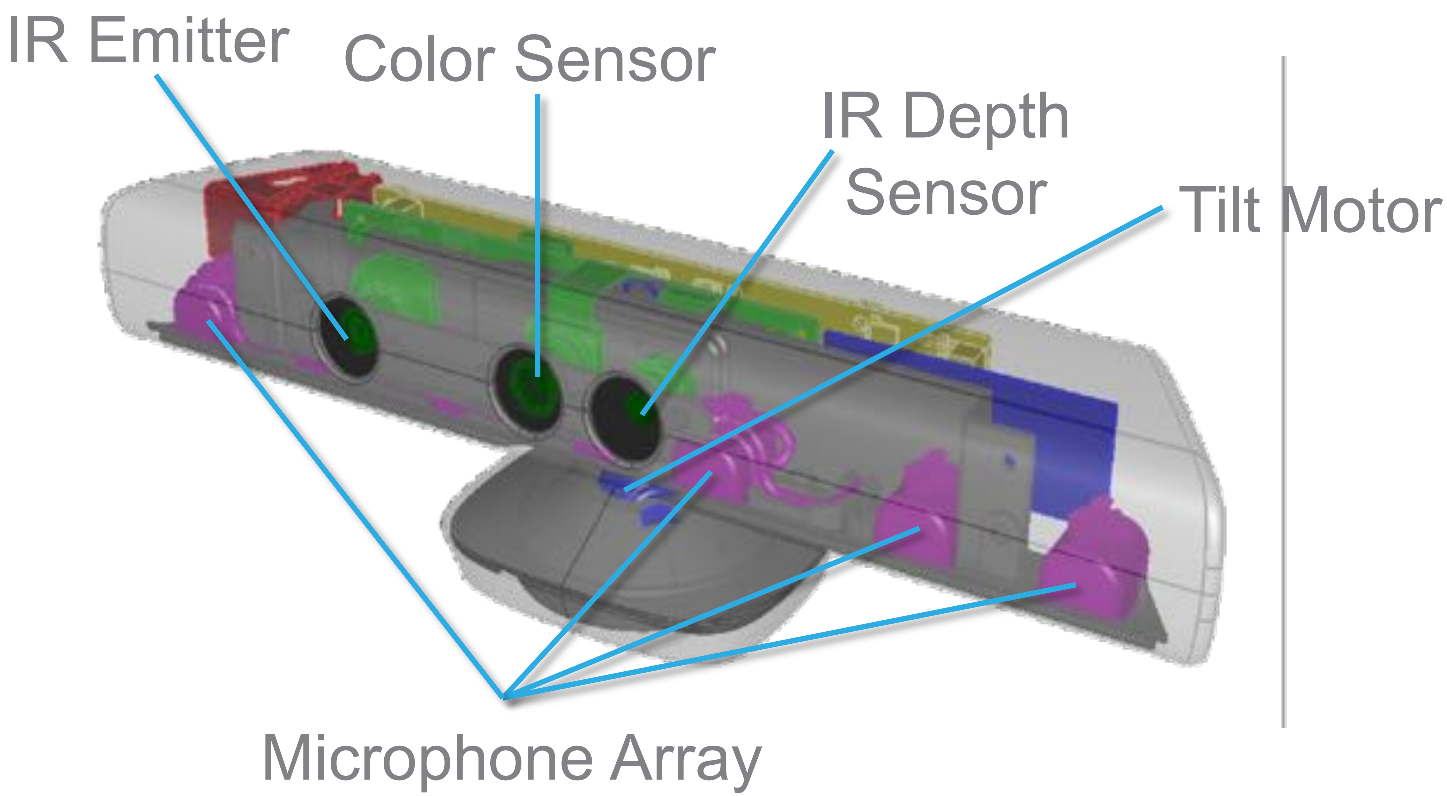
De Bruijn sequence

- A «de Bruijn sequence» of order n on a size k alphabet A, denoted by B(k, n), is a cyclic sequence in which every possible length n string on A occurs exactly once as a substring (i.e., as a contiguous subsequence).

- Eg. B(2, 3): 00010111



Kinect Hardware



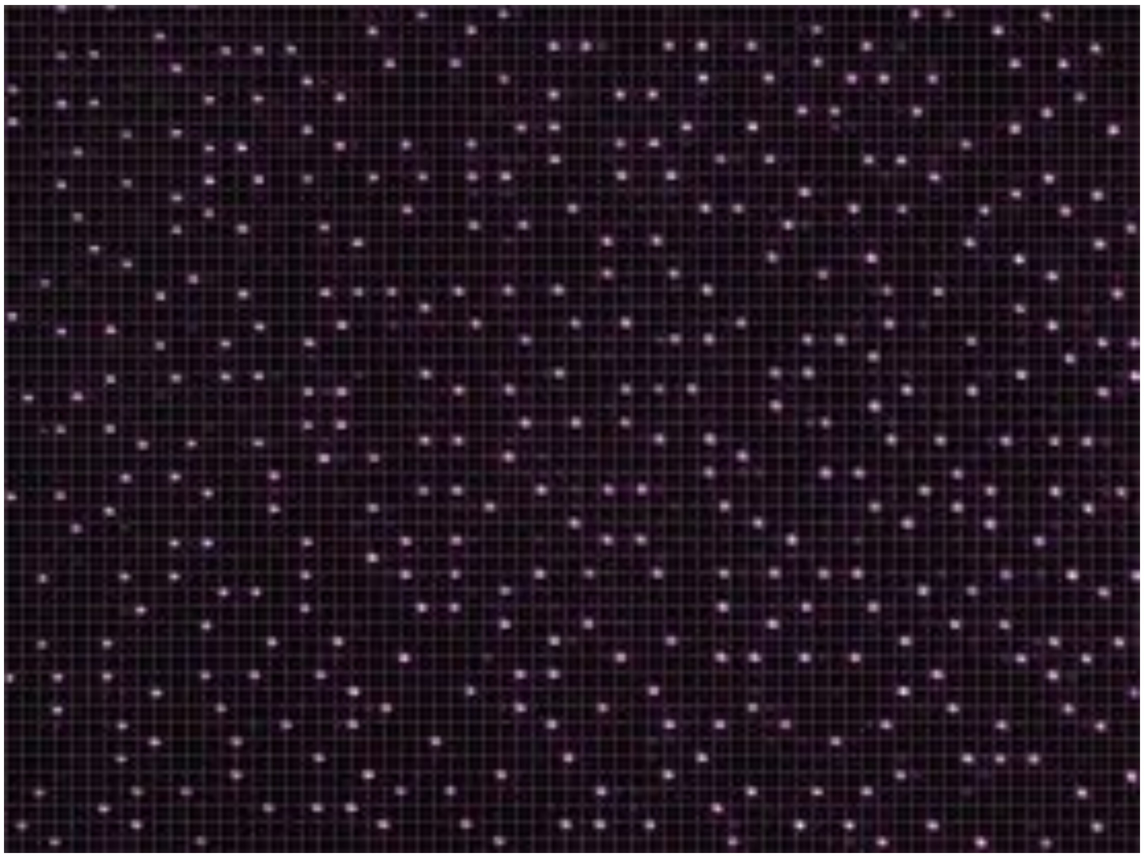
From[1]

Kinect Projects dots which are glyphs

From[1]

Kinect Glyphs – almost unique

- Local pattern identifies location of projection
- Find local identifier by looking in a small region around a given point => code



From[1]

Kinect Depth Acquisition Summary

- There is a projector for the laser dots and a sensor just for these dots (infrared)
- We can recognize the glyph in the infrared image so can triangulate to find the depth
- This requires a prior calibration process so that we know the rays for the laser dots
- Still just ordinary triangulation process
- There is a another camera that produces a separate and distinct intensity image
- The Kinect returns both a depth map and the overlayed intensity image

From[1]

Figura 99. – Immagine di colore ed immagine range dello stesso soggetto, catturate da un dispositivo Kinect (©Microsoft). Per cortesia di U. Castellani

Coded-light triangulation

Breuckmann

Steinbichler

GOM

From[3]

Breuckmann GmBh

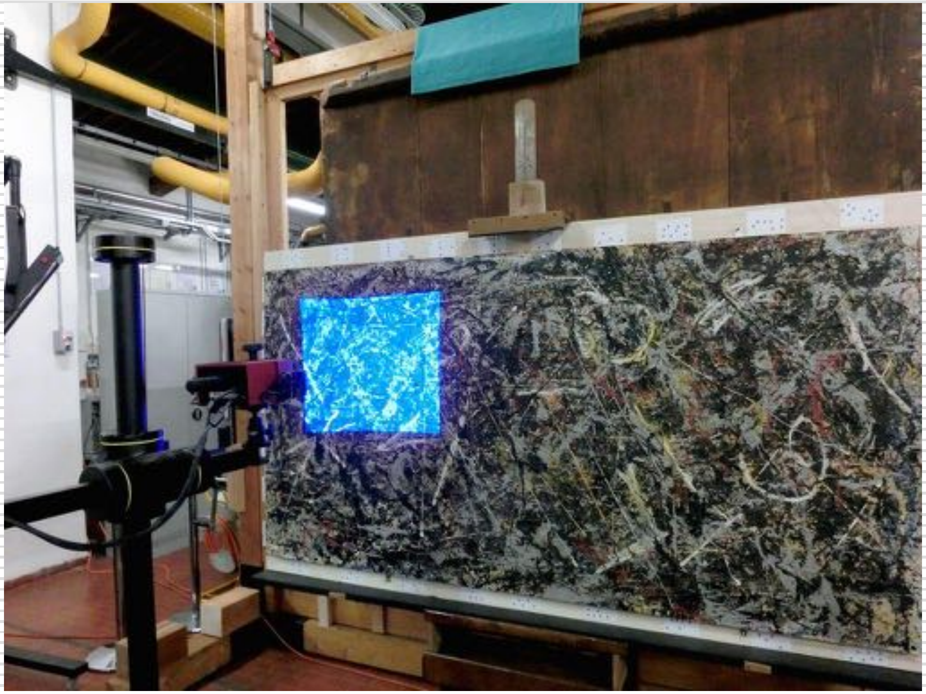
Industrial sensor, designed for optical metrology

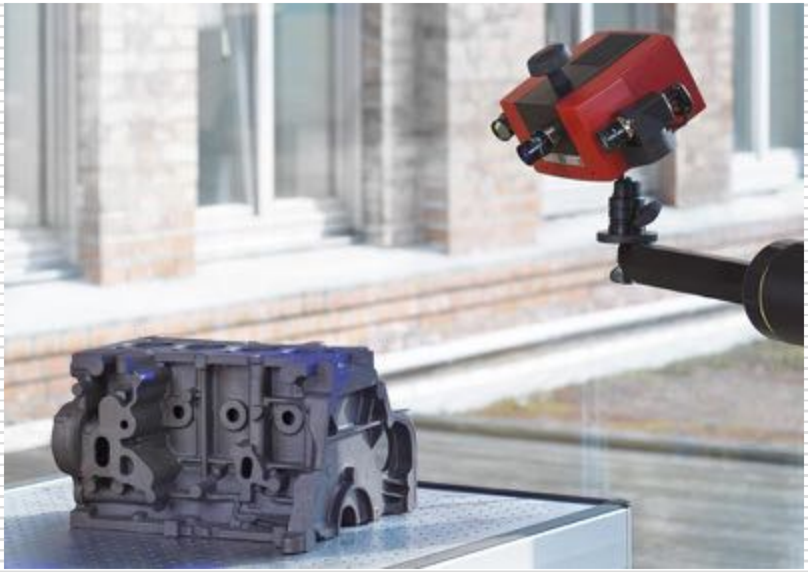
Accuracy: 0.1 mm (or less)	
Cost: 70-80k Euro	

From[2]

GOM

Industrial sensor, designed for optical metrology





From[2]

Accuracy: 0.1 mm (or less)	
Cost: 70-80k Euro	

Sensori a triangolazione attiva: sintesi

- Sistemi a scansione (scanner):
  - Punto laser
  - Lama laser

- Sistemi a campo intero (coded light)
  - Multi-stripe (color/time coding)
  - Altri pattern (Kinect)

Summary

- Stereo vision
- Active triangulation
  - Laser scanner
  - Coded light
- Time of fight
  - PW / CW scanner
  - TOF camera

Pulsed Time of Flight

Basic idea: send out pulse of light (usually laser), time how long it takes to return

Advantages:

- Large working volume (up to from 20 to 1000 m.)

Disadvantages:

- Not-so-great accuracy (at best ~5 mm.)
  - Requires getting timing to ~30 picoseconds

Often used for scanning buildings, rooms, archeological sites, etc.

The only practical long range measuring technology (triangulation fails over 20 meters)

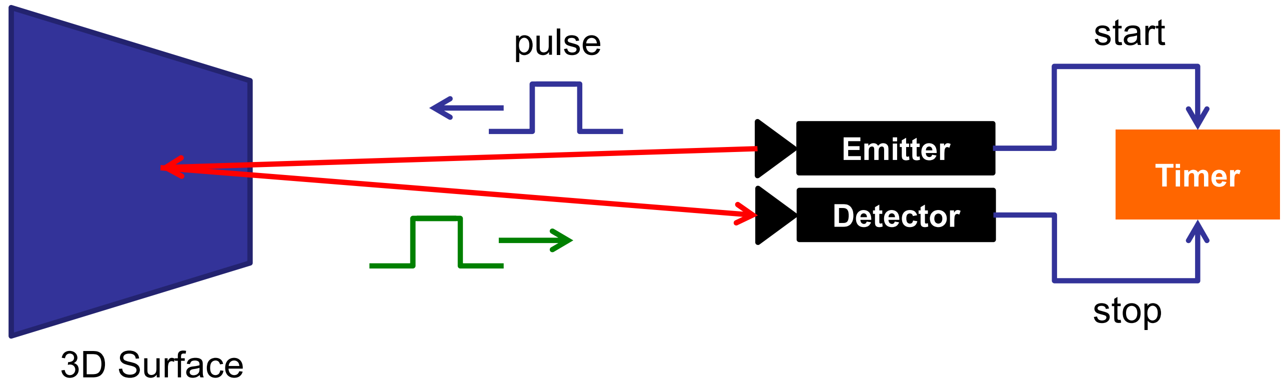
Time of flight range sensors are also caller **LIDAR** (LIght Detection And Ranging)

From[1]

Principles of ToF imaging

Pulsed Wave (PW)

- Measure distance to a 3D object by measuring the absolute time a light pulse needs to travel from a source into the 3D scene and back, after reflection
- Speed of light is constant and known, $c = 3 \cdot 10^8$m/s



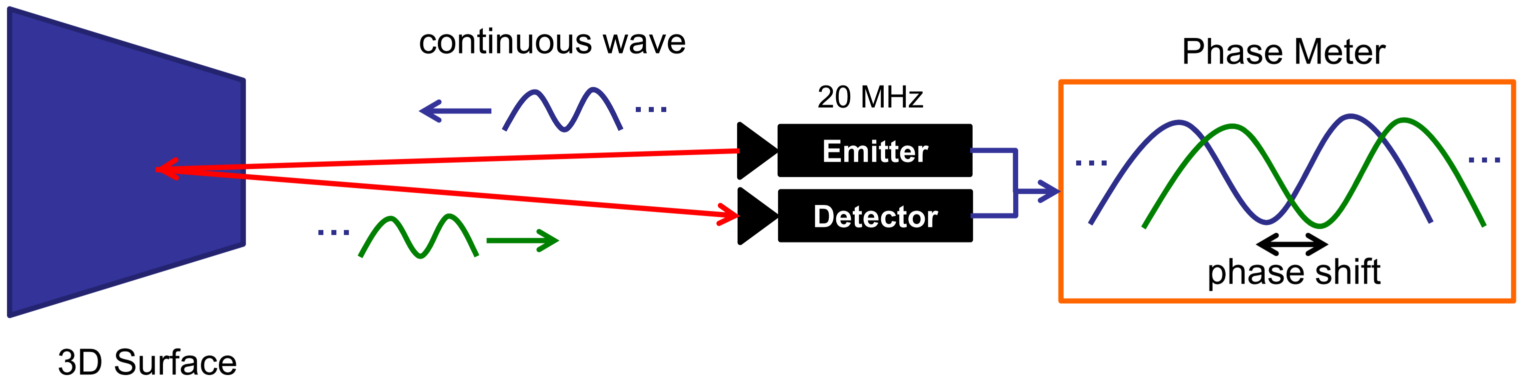
La luce percorre 1 mm ogni 3.3ps ($3.3\ 10^{-12}$ s)

From[4]

Principles of ToF imaging

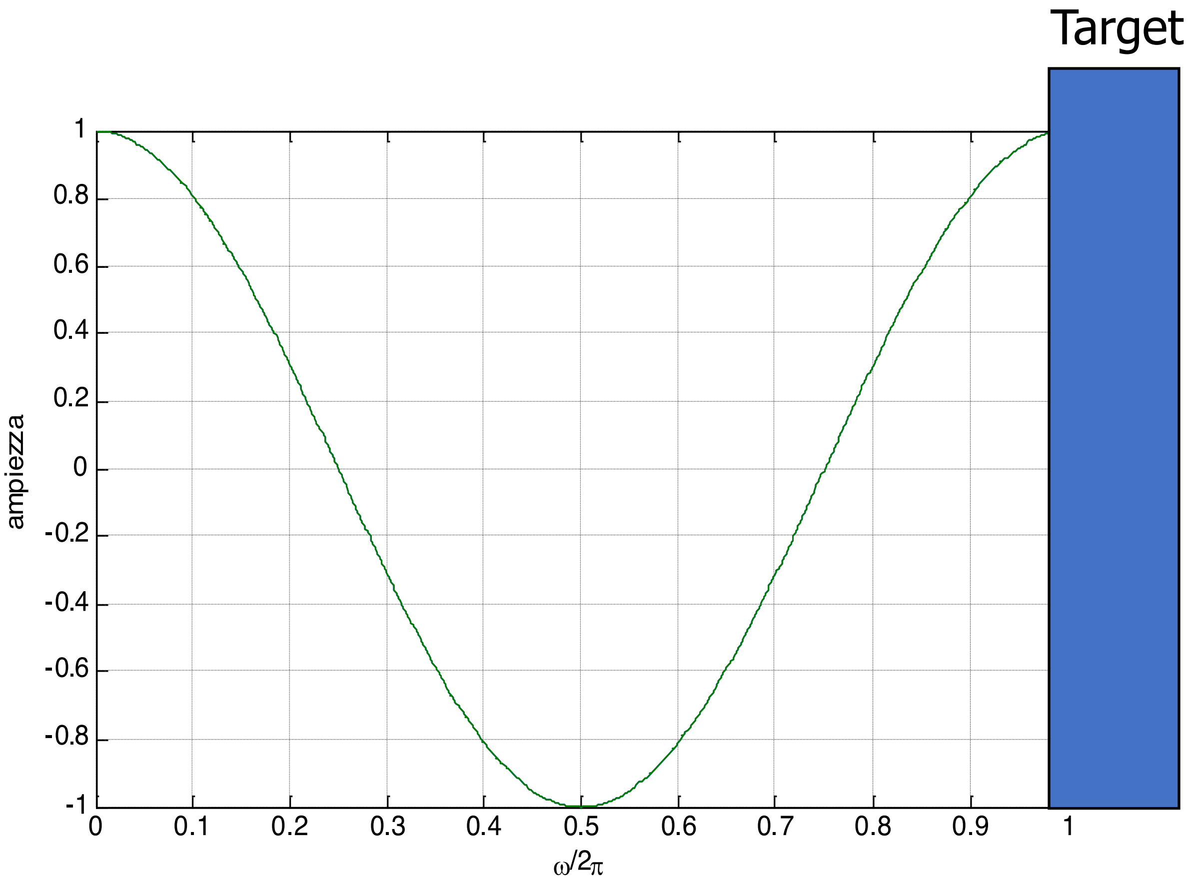
Continuous Wave (CW)

- Continuous light waves instead of short light pulses
- Modulation in terms of frequency of sinusoidal waves
- Detected wave after reflection has shifted phase
- Phase shift proportional to distance from reflecting surface



From[4]

STIMA INDIRETTA DEL TOF TRAMITE LA FASE



$$s_T(t) = \cos(\omega t)$$

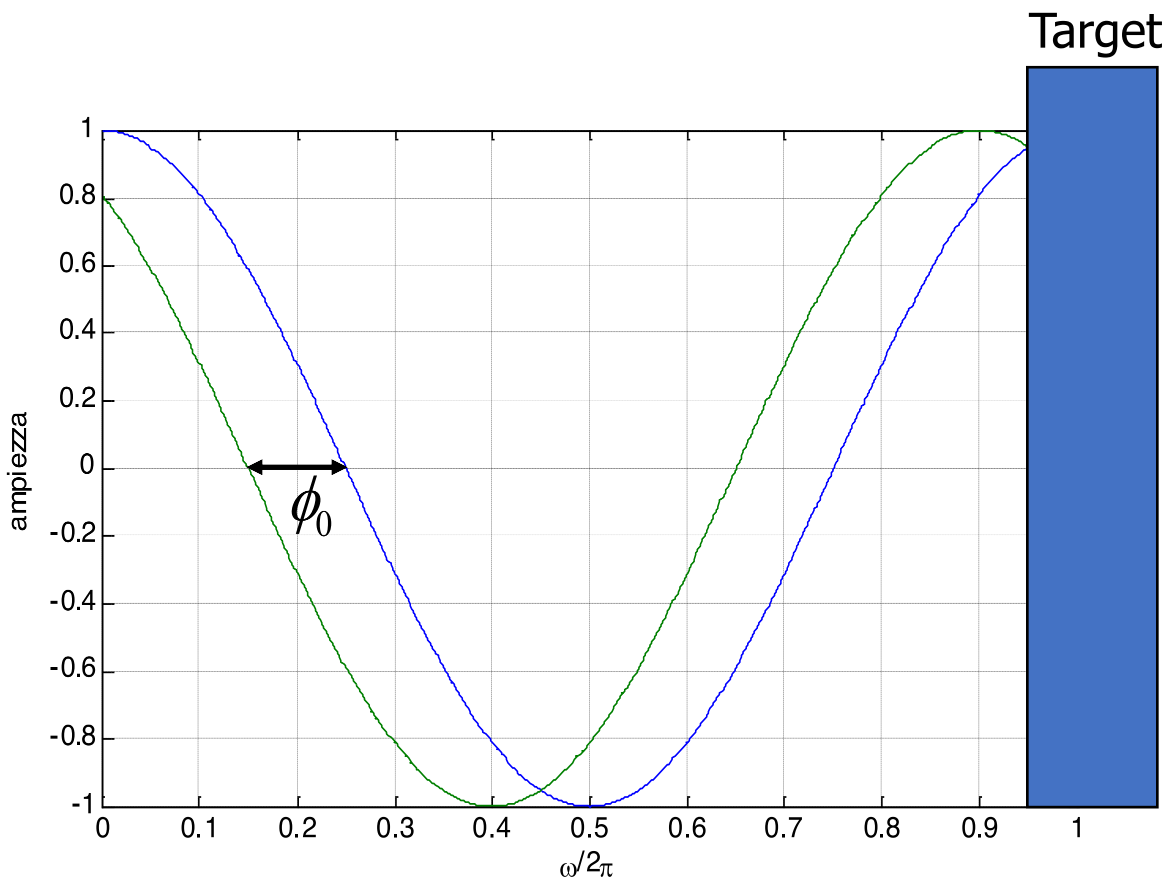
$$s_R(t) = \cos(\omega t)$$

$$\Delta\phi = 0$$

- Nel caso di ostacolo posto ad un ciclo esatto di distanza, l'onda viene riflessa identica a se stessa

From[3]

STIMA INDIRETTA DEL TOF TRAMITE LA FASE



$$s_T(t) = \cos(\omega t)$$

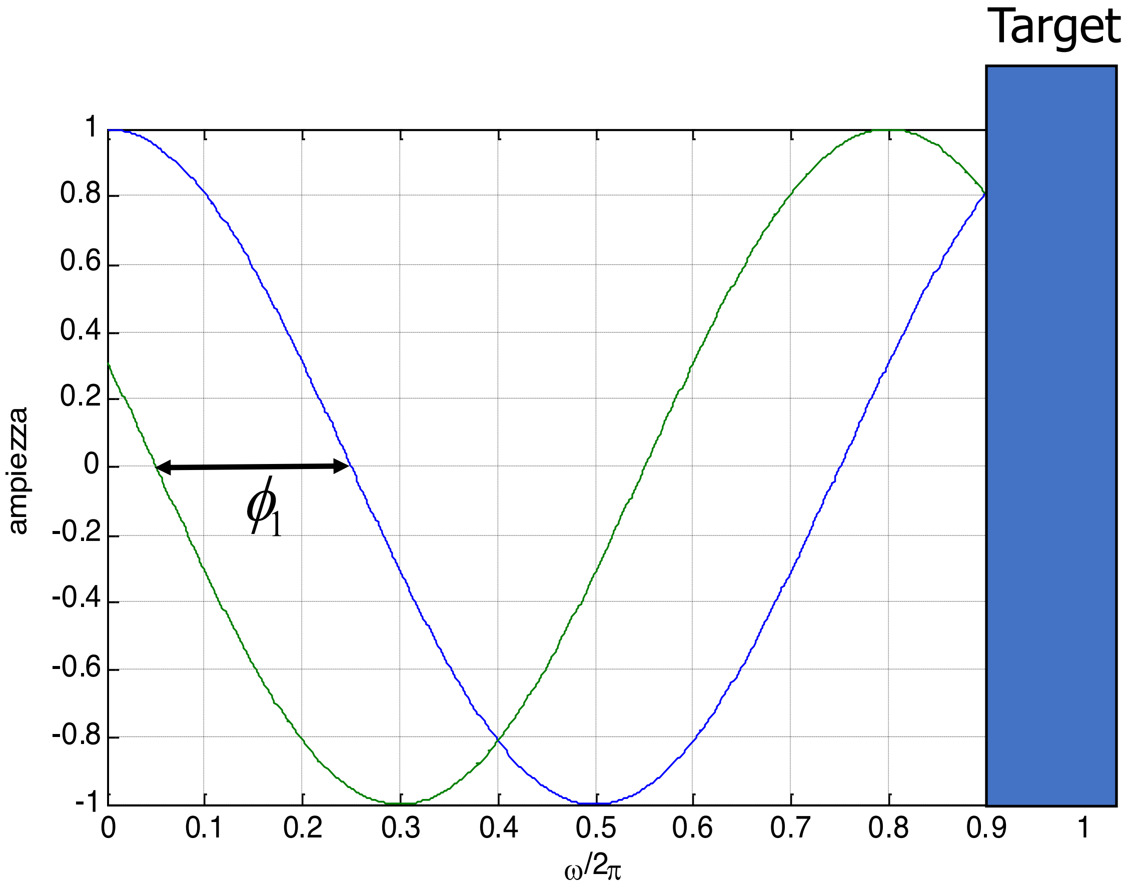
$$s_R(t) = \cos(\omega t + \phi_0)$$

$$\Delta\phi = \phi_0$$

- Se l'ostacolo viene leggermente arretrato il segnale riflesso è uguale a quello che continuerebbe, ma invertito nel tempo.

From[3]

STIMA INDIRETTA DEL TOF TRAMITE LA FASE



$$s_T(t) = \cos(\omega t)$$

$$s_R(t) = \cos(\omega t + \phi_1)$$

$$\Downarrow$$

$$\Delta\phi = \phi_1$$

- Se l'ostacolo viene ulteriormente arretrato il segnale riflesso è caratterizzato da uno sfasamento via via crescente.

From[3]

Scanners

- Most time of flight sensors transmit only a single beam, thus range measurements are only obtained from a single surface point.
- To obtain these denser representations, the laser beam is swept across the scene → **scanner**. Normally the beam is swept by a set of mirrors rather than moving the laser and detector themselves (mirrors are lighter and less prone to motion damage).

TOF laser scanner (PW)

From[3]

TOF laser scanner (CW)

From[3]

Faro

Zoller+Fröhlich

Leica

ToF cameras (Flash LIDAR)

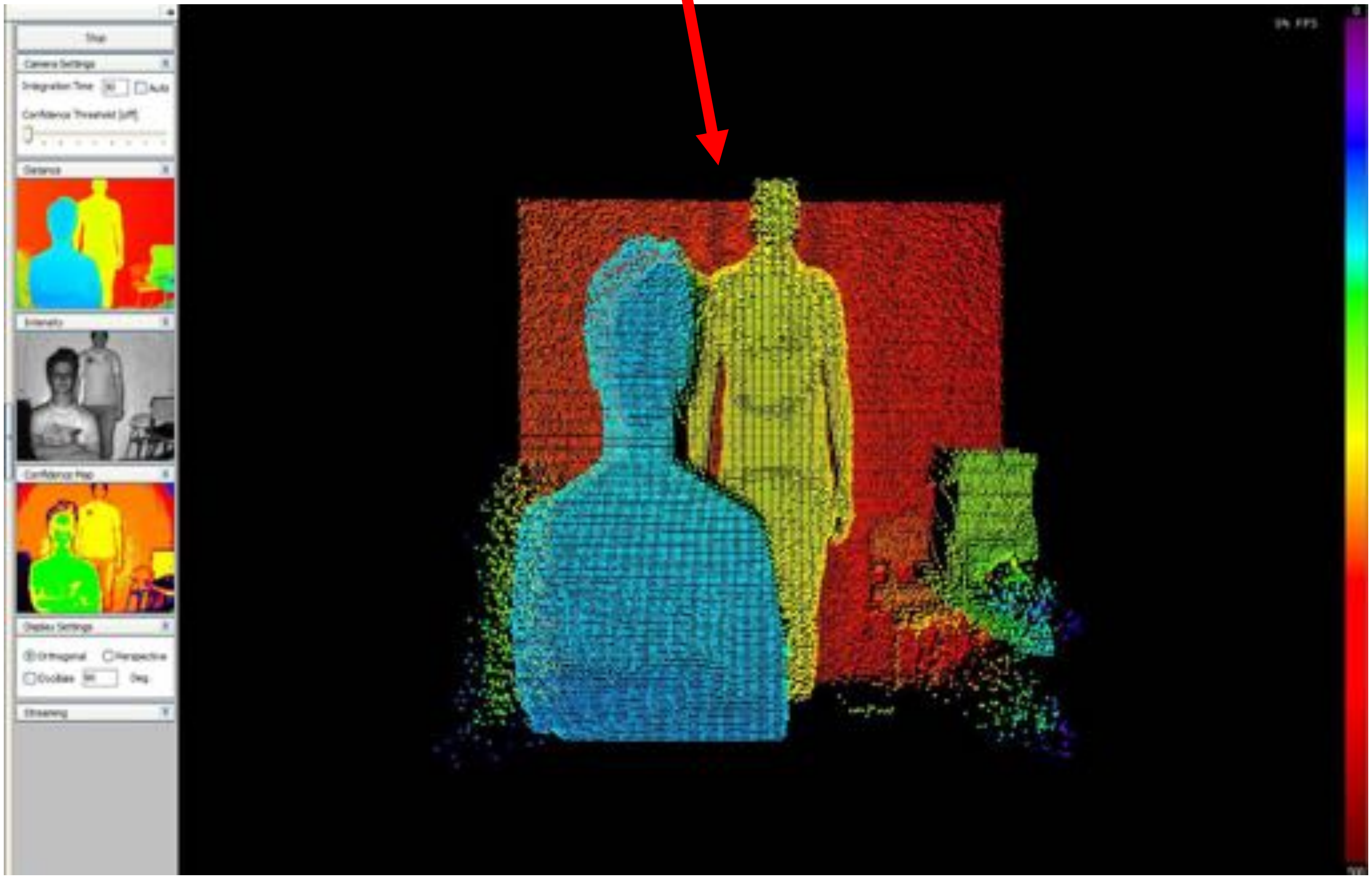
- Recently, a type of ToF range sensor called the "Flash LIDAR" of ToF cameras has been developed.
- VLSI timing circuits at each pixel of the sensor chip. Thus, each pixel can measure the time at which a light pulse is observed from the line of sight viewed by that pixel. This allows simultaneous calculation of the range values at each pixel.
- The light pulse now has to cover to whole portion of the scene that is observed, so sensors typically use an array of infrared laser LEDs.
- While spatial resolution is smaller than current cameras (e.g. 64×64, 160×124, 128×128), the data can be acquired at video rates (30-50 fps), which provides considerable information usable for robot feedback

Mesa Swiss Ranger 4000

- Range measurements from 0.3m to 5 m (Mod. freq. 30 MHz)

Cost ~$9,000.

Depth map



Depth map

Reflectance

Confidence map

From P. Zanuttigh and G. Cortelazzo

Range sensor + CMM

- Problem:
  - Range sensors capture the 3D structure of the world from the viewpoint of the sensor;
  - Several viewpoints are needed to obtain a complete 3D model;
  - Multiple observations of the scene need to be aligned in a common reference frame.
- Post-processing solution: alignment is done by a software
- On-line solution: couple the range sensor with a CMM (**Coordinate Measuring Machine**) to get direct and automatic alignment

Triangulation laser Scanner + CMM

From[3]

Leica T-scan – optical CMM

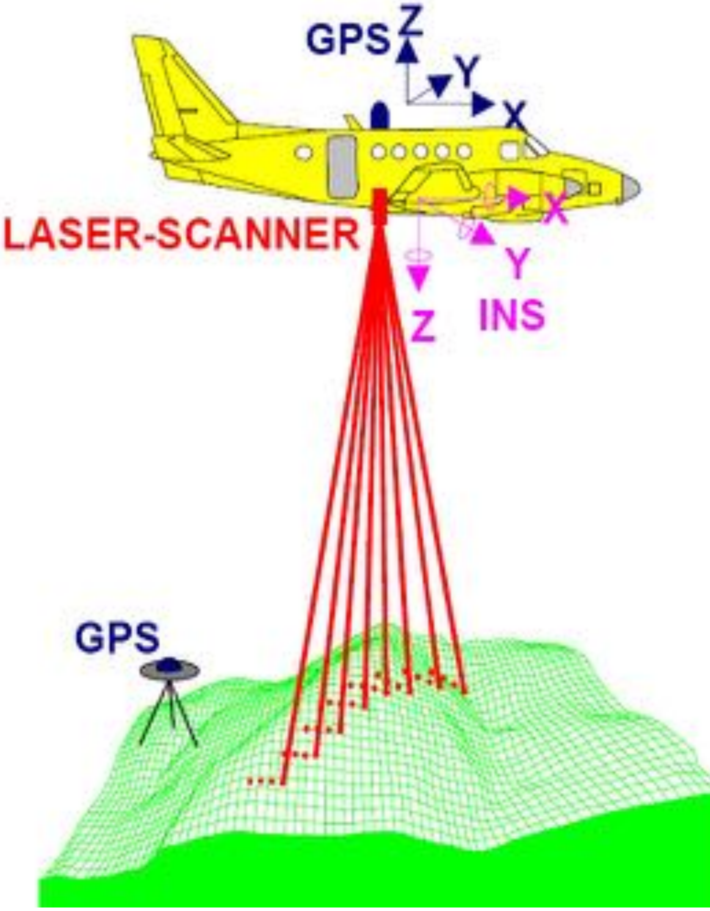
Polhemus – magnetic CMM

INO – ultrasound CMM

Leica T-scan 5

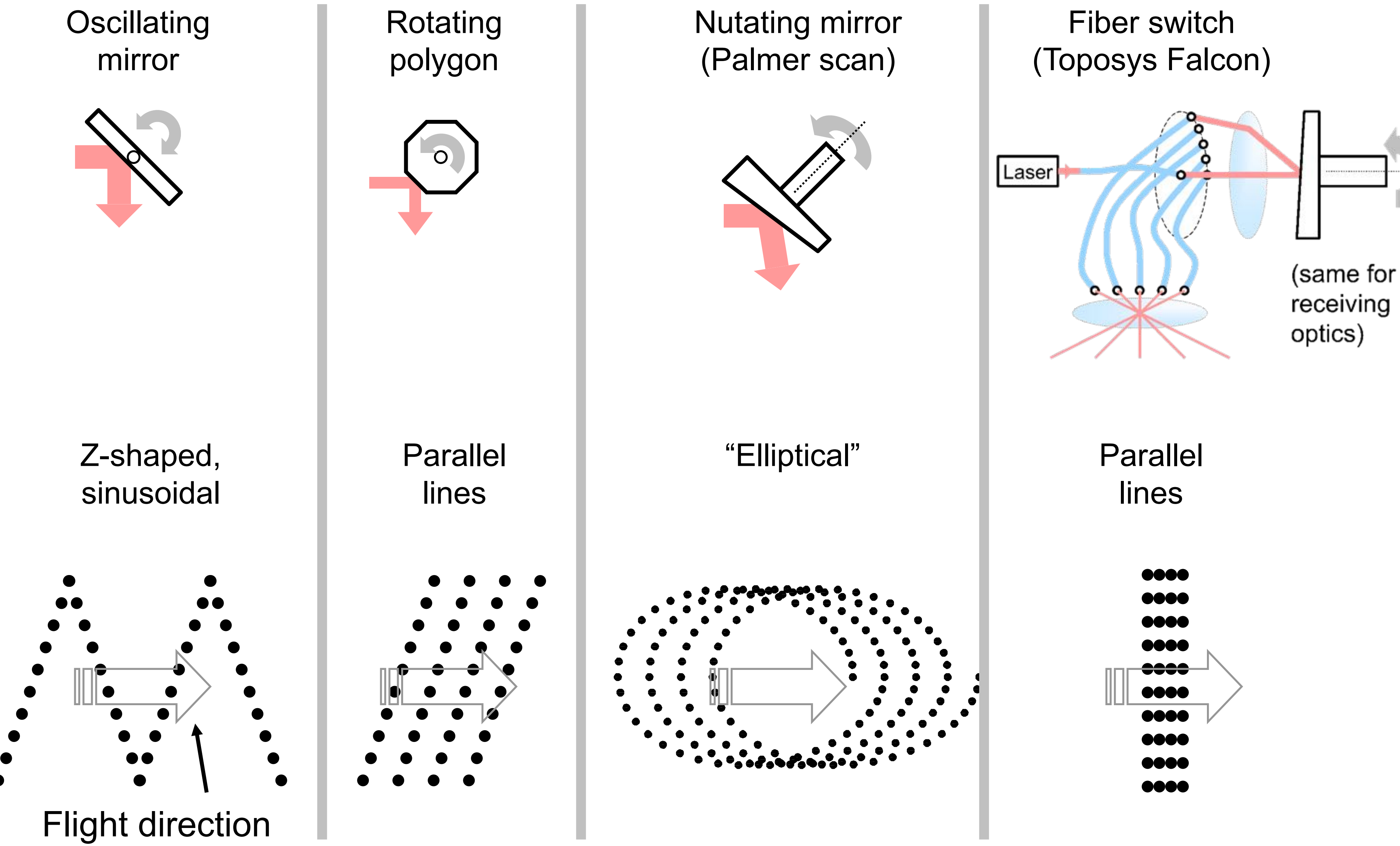
Airborne Laser Scanner

- Sistema TOF aviotrasportato o ALS (Airborn Laser Scanner)
- In analogia ai sistemi terrestri accoppiati ai CMM, questo è un sistema di scansione con misura in un unico sistema di coordinate, grazie all'uso di sistemi di misura accessori:
  - GNSS (*Global Navigation Satellite System*) (e.g. GPS) per la misura della posizione
  - IMU *(Inertial Measurement Unit*) per la misura dell'orientamento



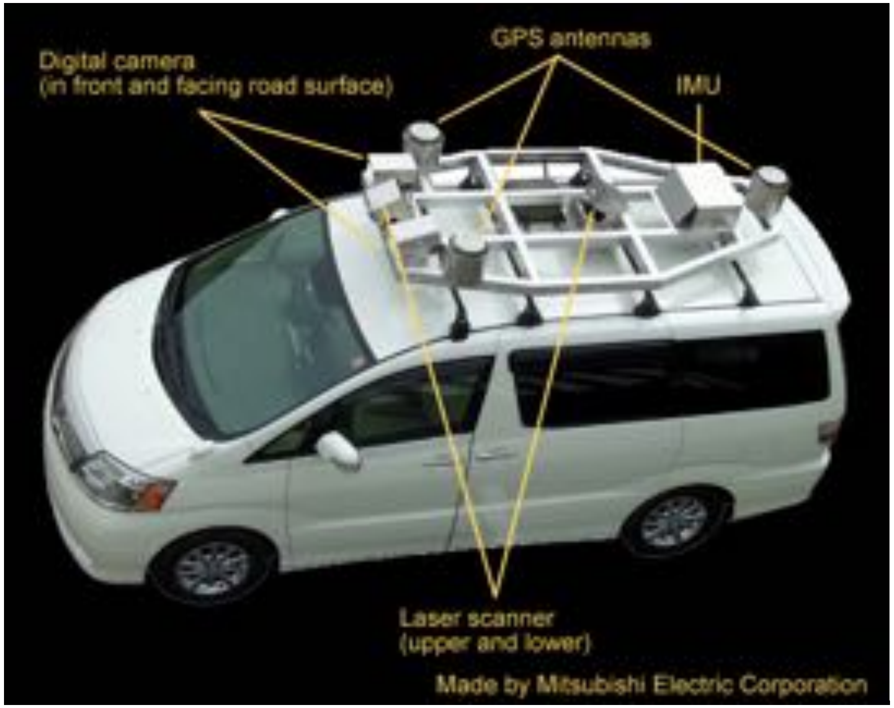
From[3]

MECCANISMI DI SCANSIONE NEGLI ALS



From[3]

Mobile mapping systems (MMS)



- Same principles of ALS applies to different carriers such as vehicles or humans.
- In GNSS-denied areas SLAM techniques are used

Credits

1. R.B Fisher, K. Konolige. Handbook of Robotics Chapter 22 - Range Sensors (2008)
2. Gerhard Roth. Active Stereo Vision (Winter 2014)
3. M. Callieri, M. Dellepiane . A 3D scanning primer (Febbraio 2016).
4. G. Guidi. Il paradigma della scansione laser: principi e storia (Aprile 2010).
5. V. Castaneda, N. Navab. Time-of-Flight and Kinect Imaging (June 2011).