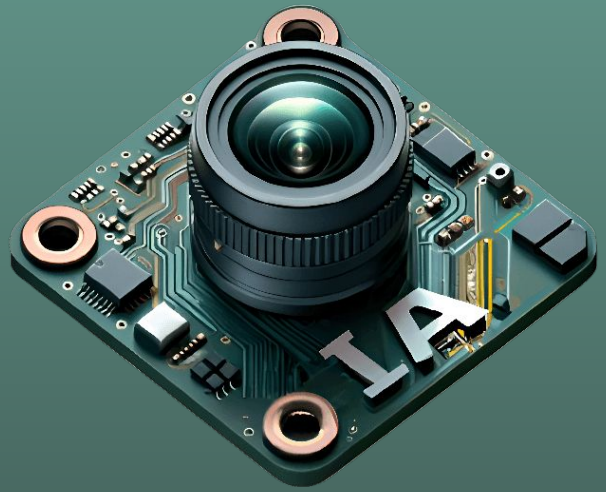
Universidad Industrial de Santander



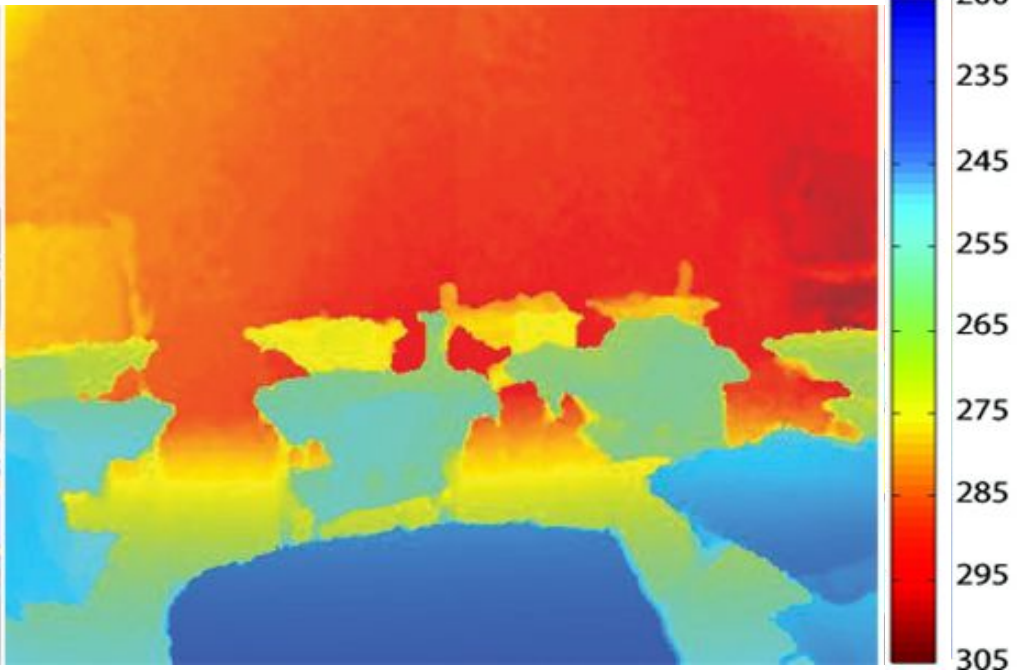
Semillero de Investigación
"Hands - on"
Computer Vision

SESIÓN 6: DEPTH ACTIVO

Contenido

1. Técnicas de estimación de profundidad
    a. Métodos pasivos vs activos
2. Luz estructurada
    a. Patrones de proyección
    b. Estimación de profundidad
3. Time of Flight (ToF)
    a. iToF y el método de cuadratura
    b. dToF (Lidar) e histogramas
4. Hands-on Active Depth Imaging

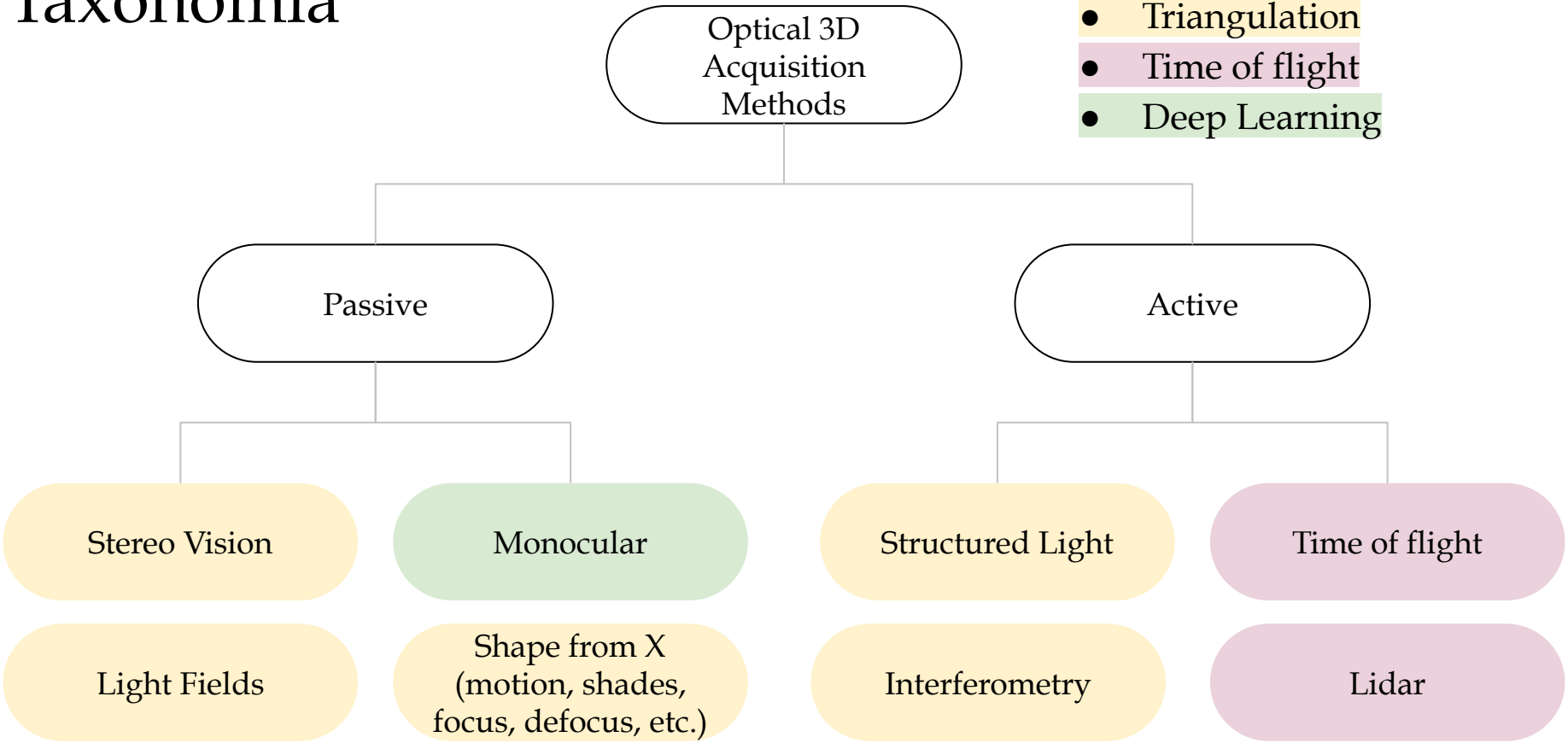
Imágenes de profundidad (x,y,z)



Taxonomía

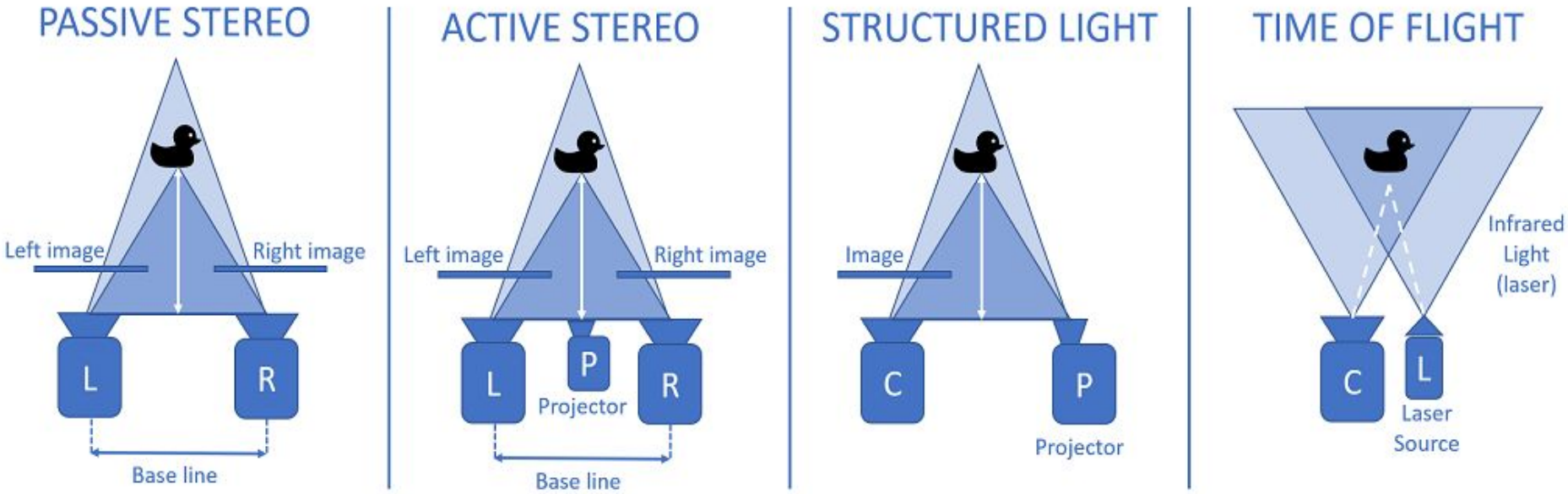
Depth estimation techniques:

- Triangulation
- Time of flight
- Deep Learning

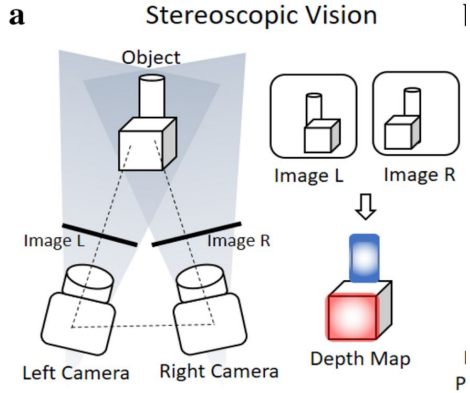


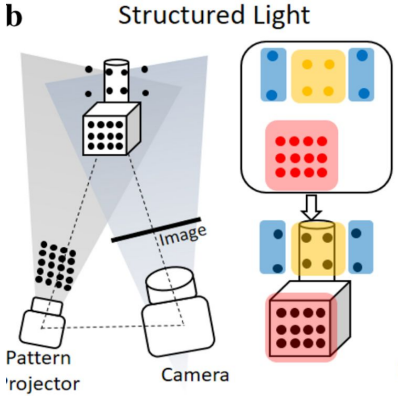
Pasivo: Estimación de la profundidad no requiere la emisión activa de señales o radiación hacia la escena

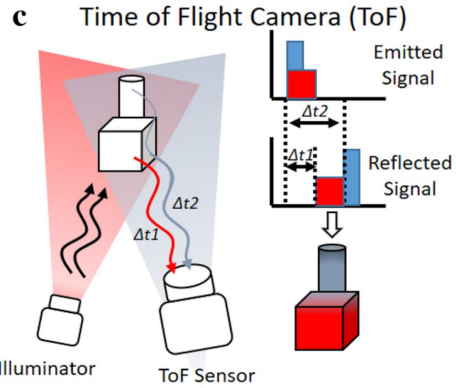
Basadas en Triangulación

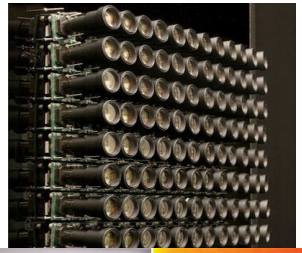


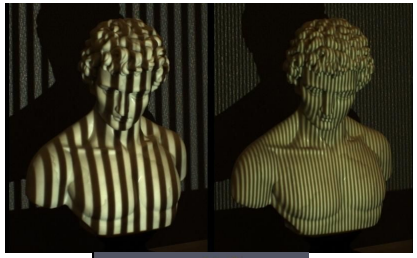
Métodos de Estimación de Profundidad

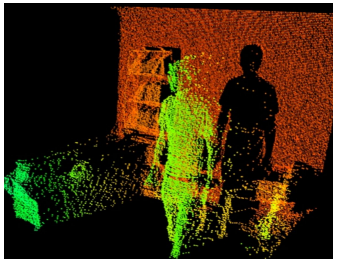


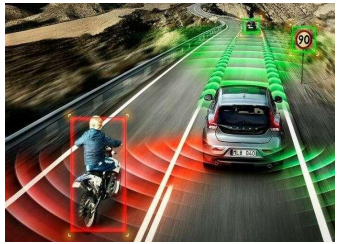


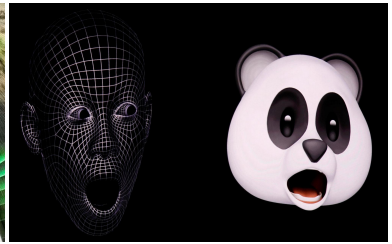












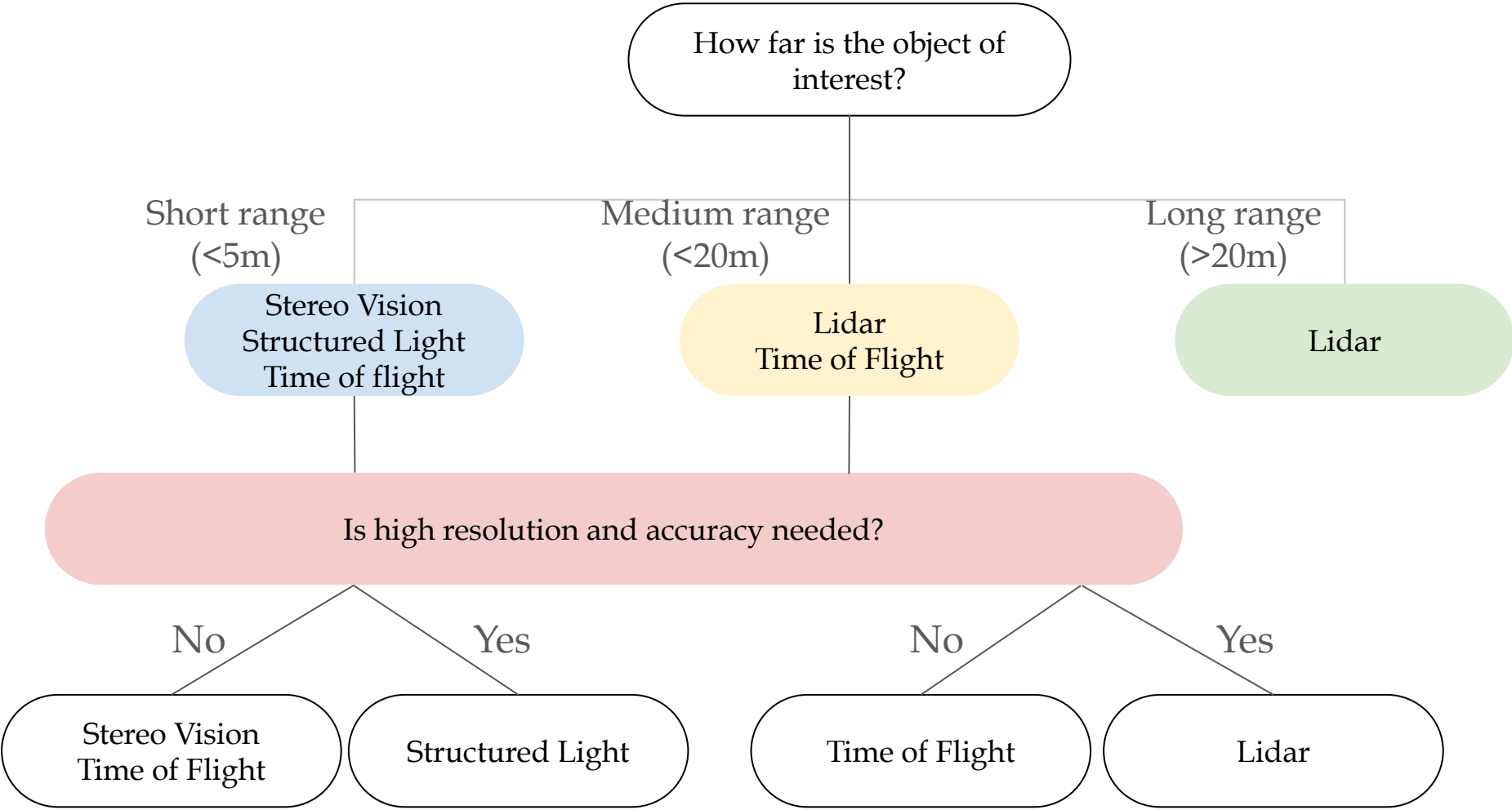
Tecnología Pasiva vs. Activa

Parameter	Stereo vision	Structured light	Time-of-flight
Range	Limited	Can be adapted	Can be adapted
Cost	Low	High	Medium
Software complexity	High	Medium	Low
Depth accuracy	Low	High	Medium
Low-light performance	Weak	Good	Good
Sunlight robustness	Good	Weak	Good

Best

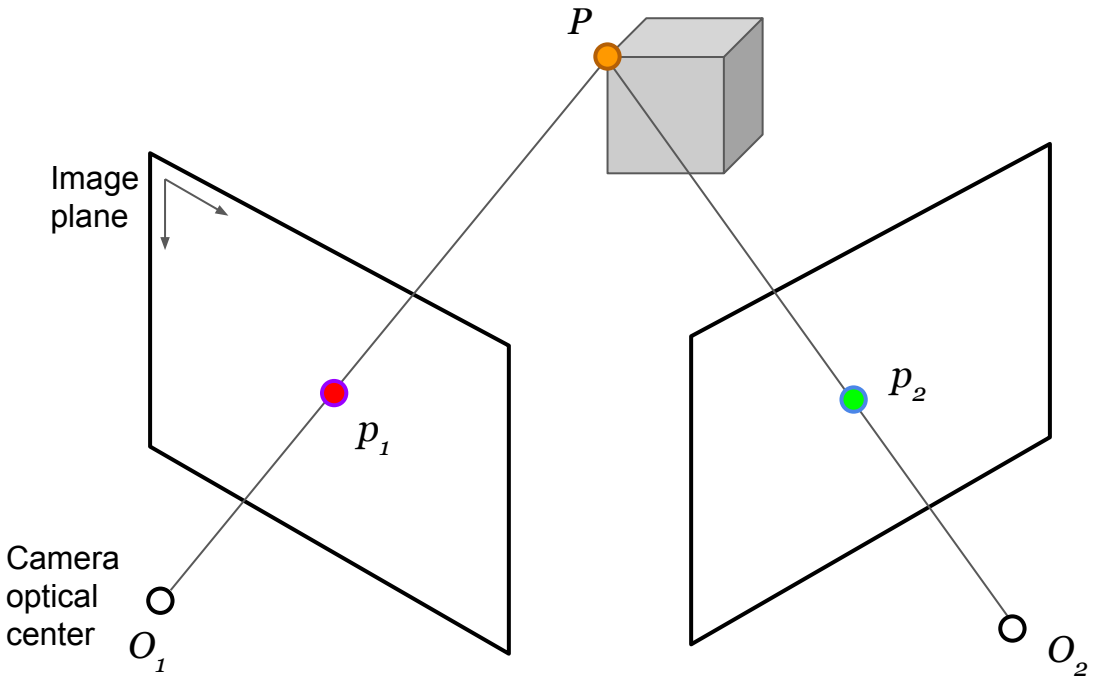
Worst

Best trade-off


How far is the object of interest?
Short range (<5m)
Medium range (<20m)
Long range (>20m)
Stereo Vision
Structured Light
Time of flight
Lidar
Time of Flight
Lidar
Is high resolution and accuracy needed?
No
Yes
No
Yes
Stereo Vision
Time of Flight
Structured Light
Time of Flight
Lidar

2. Luz Estructurada

Recordando: Passive Stereo Vision



Principio de Operación (Un punto a la vez)

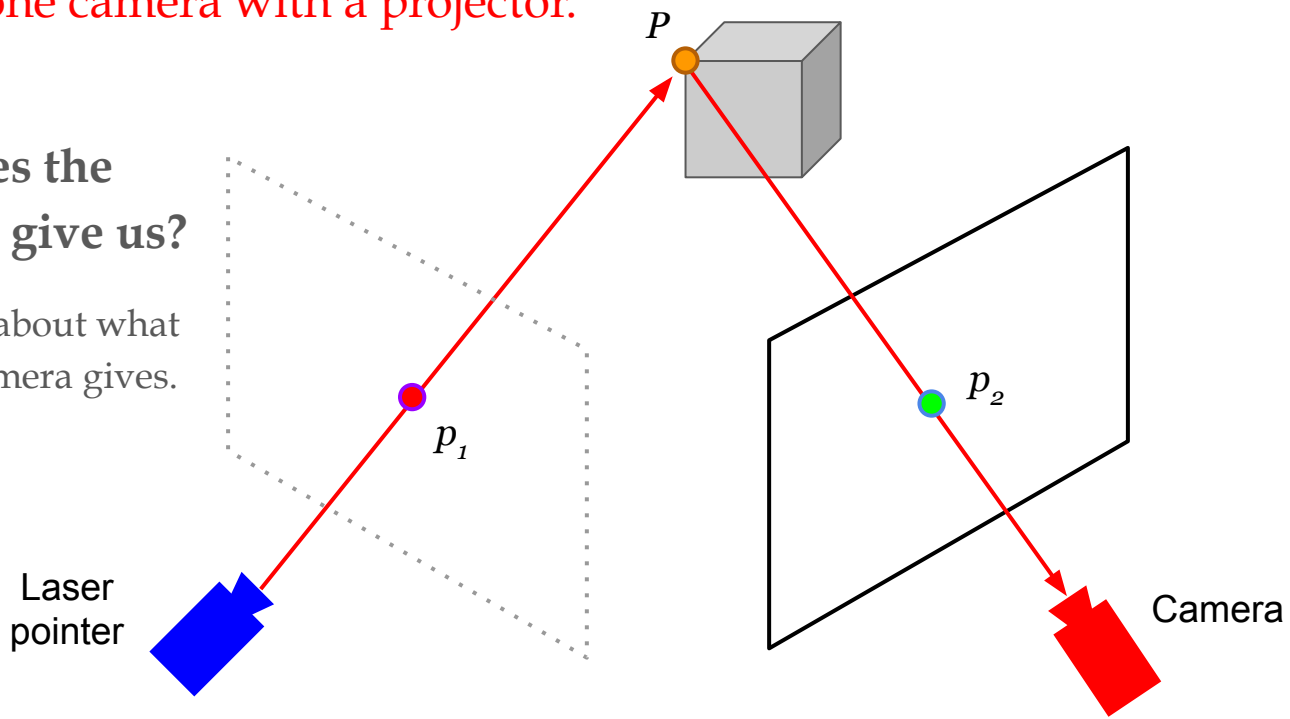
"Active Stereo"

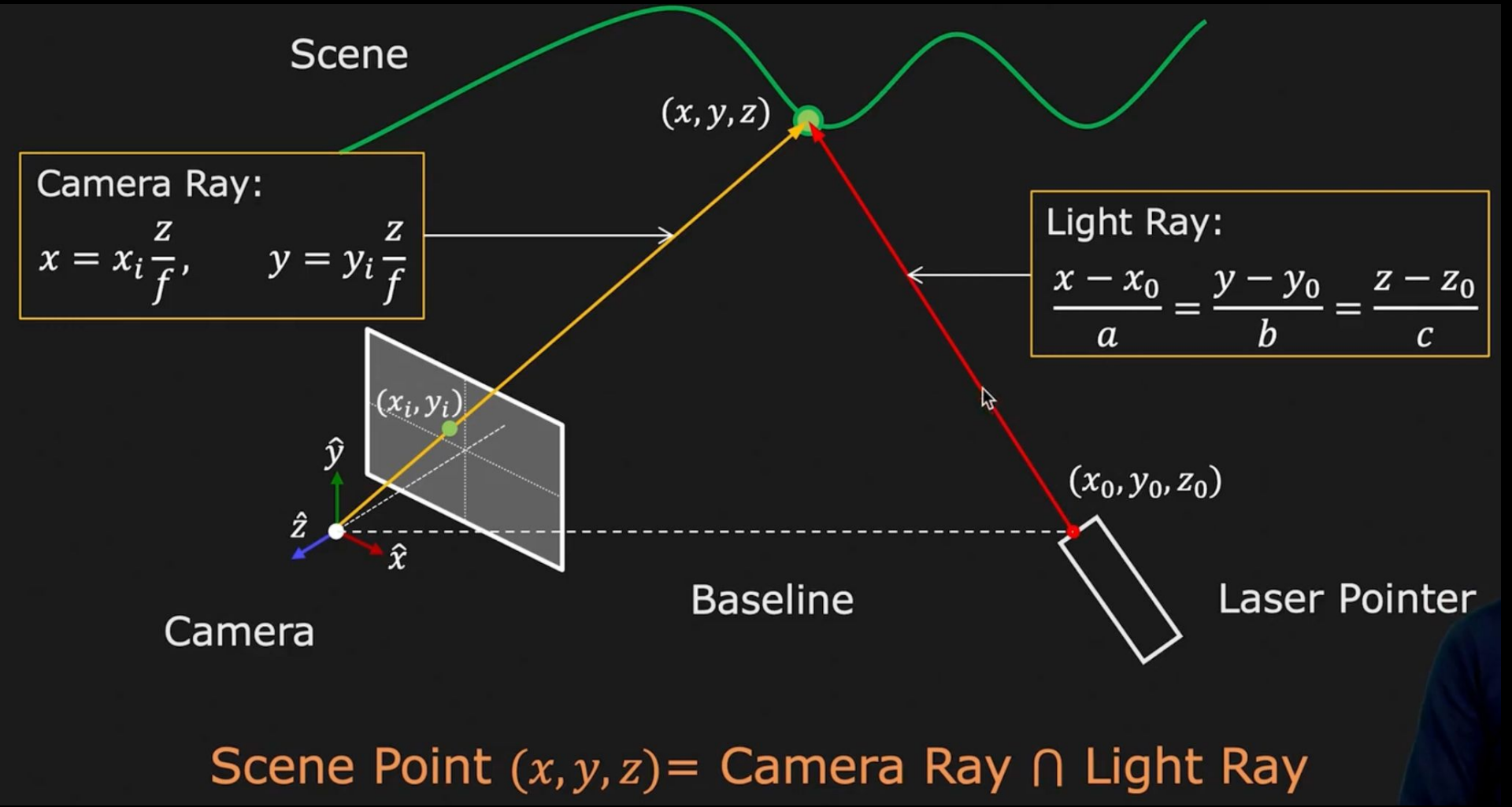
Replace one camera with a projector.

But:

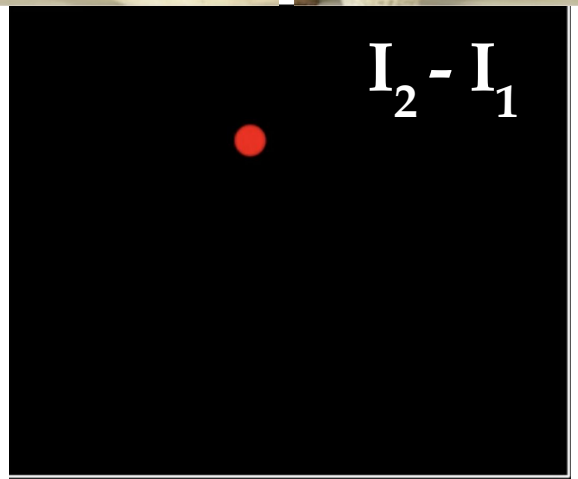
What does the projector give us?

Hint: think about what a second camera gives.




Scene
(x, y, z)
Camera Ray:
$x = x_i \frac{z}{f}, \quad y = y_i \frac{z}{f}$
Light Ray:
$\frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c}$
(x_i, y_i)
$\hat{y}$
$\hat{z}$
$\hat{x}$
(x_0, y_0, z_0)
Camera
Baseline
Laser Pointer
Scene Point (x, y, z)= Camera Ray ∩ Light Ray

Un punto a la vez



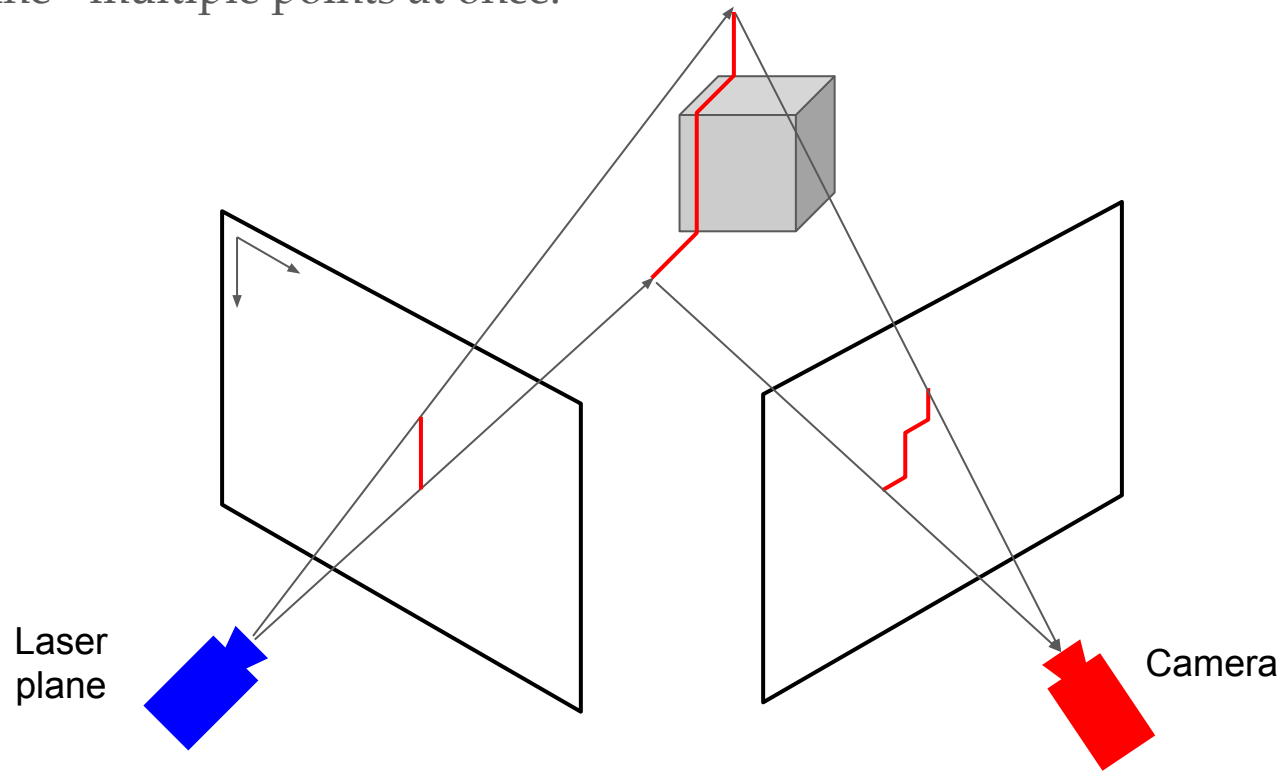
Para una imagen pequeña:
640x480

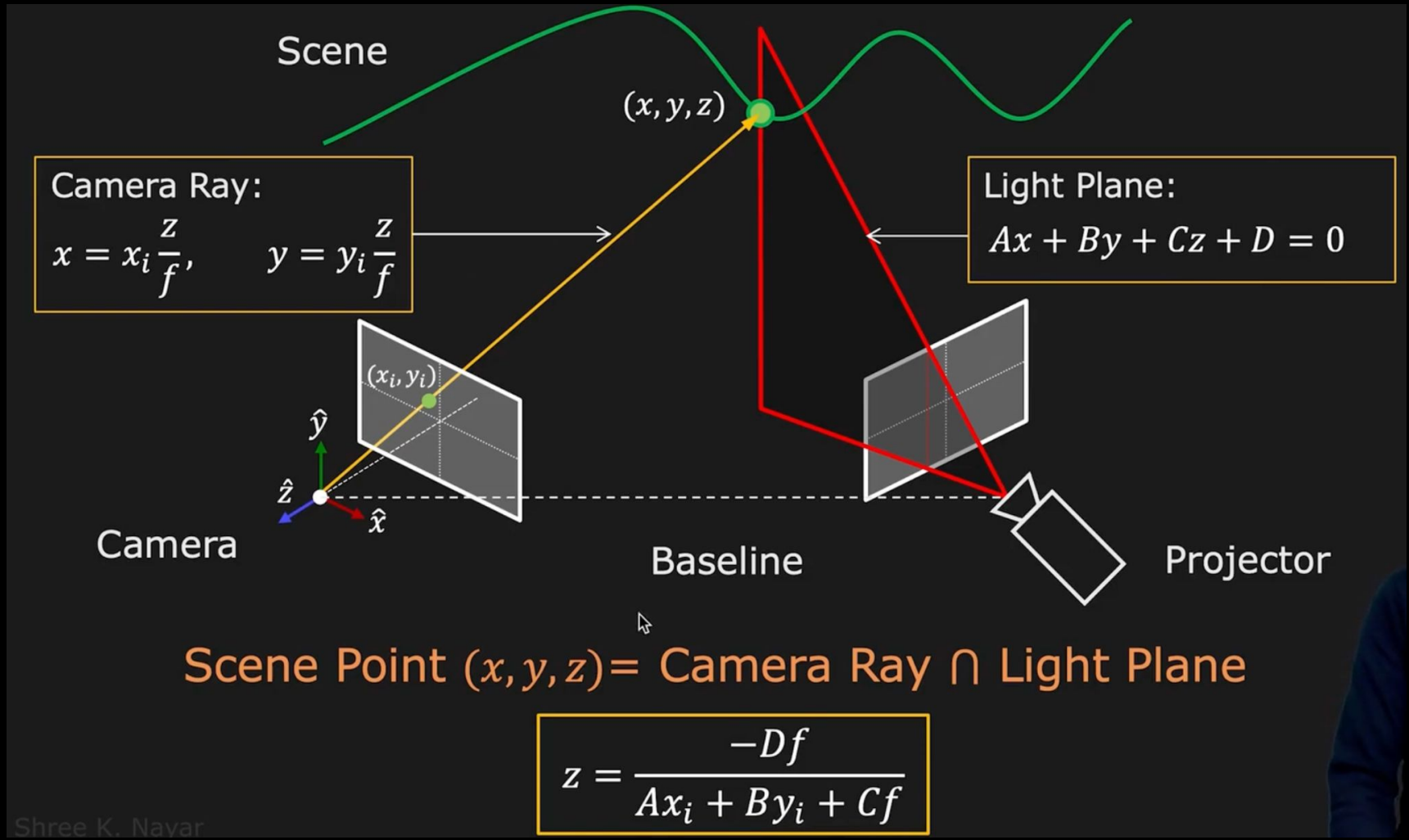
Número de imágenes:
>300.000

Tiempo de adquisición:
(Asumiendo 30 fps)
~2,77 horas

Principio de Operación (Una línea a la vez)

Project a line - multiple points at once.





Scene Point (x, y, z)= Camera Ray ∩ Light Plane

$$z = \frac{-Df}{Ax_i + By_i + Cf}$$

Una línea a la vez

Lo que el proyector "ve"

Lo que la cámara "ve"

Para una imagen pequeña:
640x480

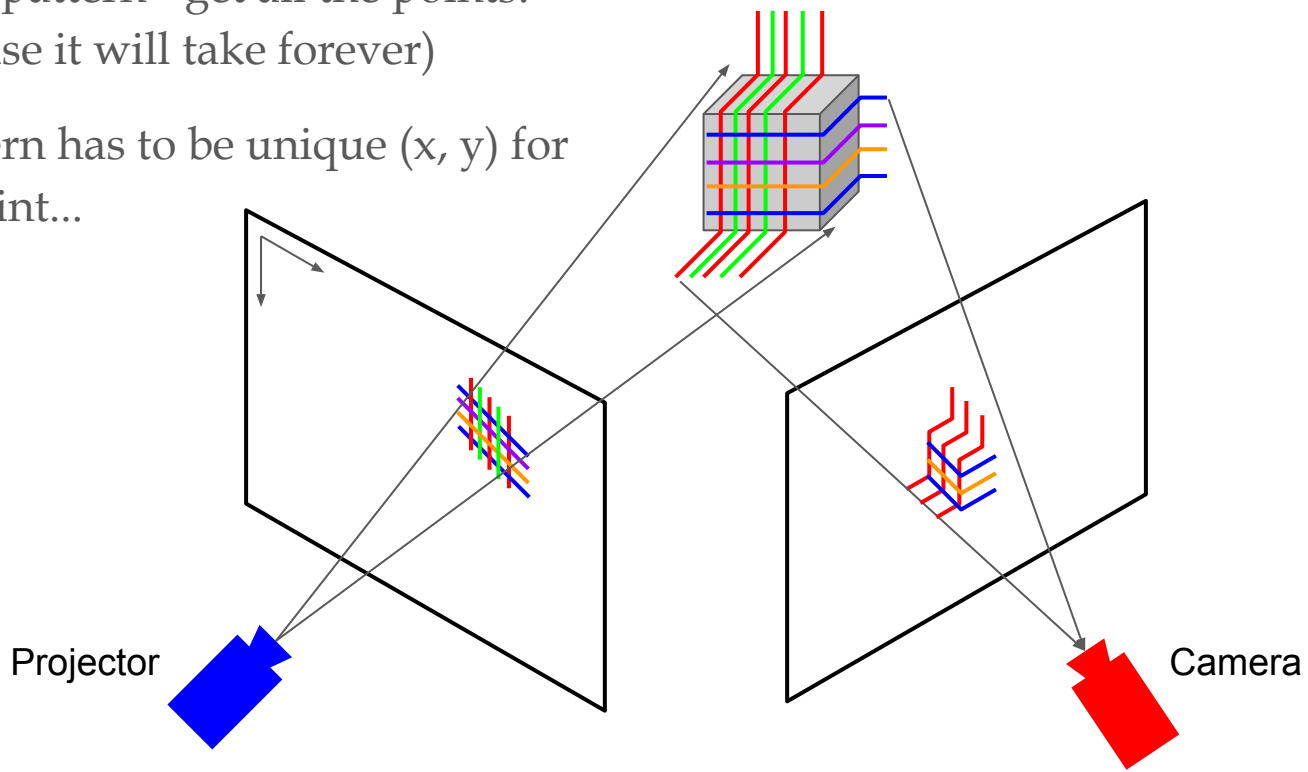
Número de imágenes:
640

Tiempo de adquisición:
(Asumiendo 30 fps)
~21 segundos

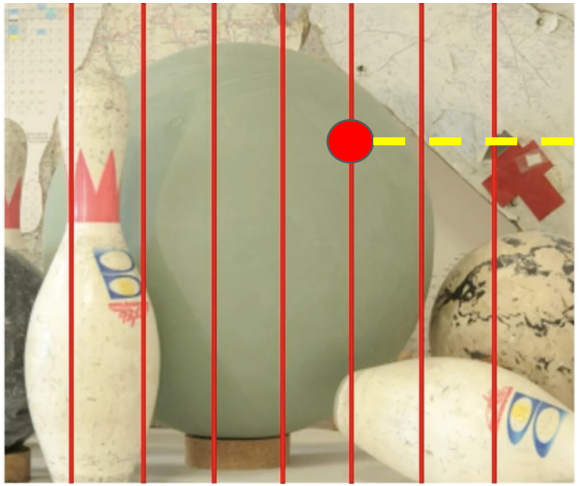
Principio de Operación (Múltiples líneas a la vez)

Project a pattern - get all the points!
(Otherwise it will take forever)

But pattern has to be unique (x, y) for every point...



Múltiples líneas a la vez



Lo que el proyector "ve"

Lo que la cámara "ve"

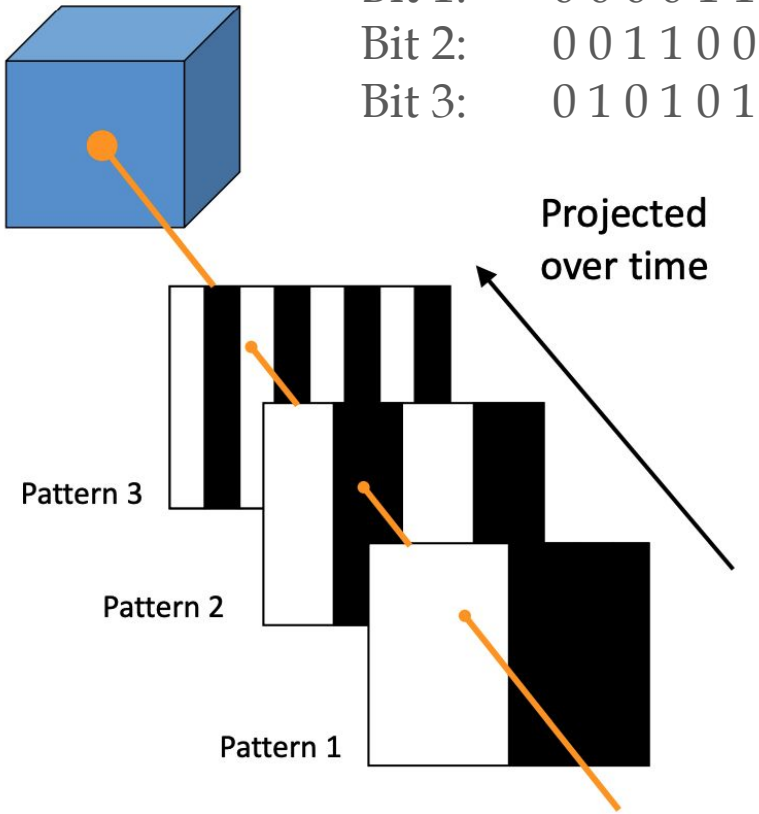
¿Cómo identificamos cuál línea es cuál?

Time multiplexing

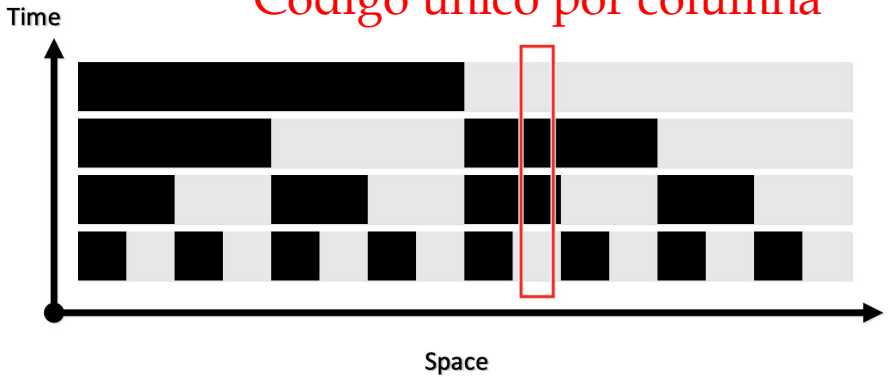
Resolviendo “ambigüedad” en múltiples líneas

Con 3 bits:

Bit 1: 0 0 0 0 1 1 1 1
Bit 2: 0 0 1 1 0 0 1 1
Bit 3: 0 1 0 1 0 1 0 1



Código único por columna



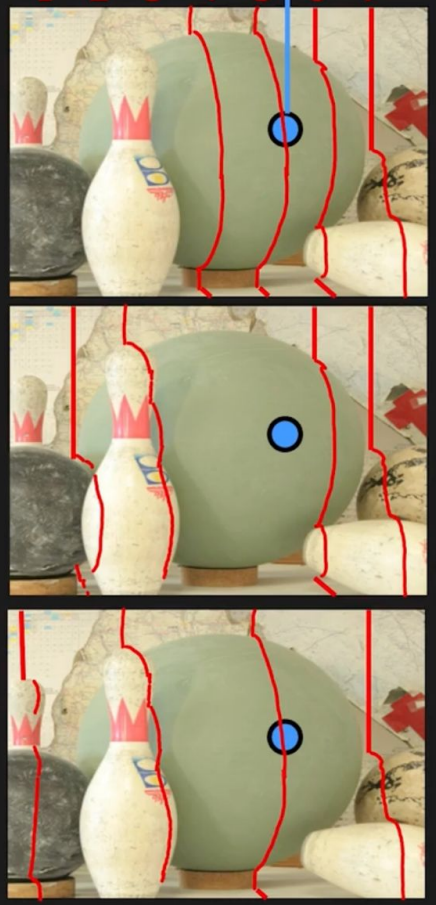


Image		Projection Pattern						
1	Bit 1	0	0	0	1	1	1	1
2	Bit 2	0	1	1	0	0	1	1
3	Bit 3	1	0	1	0	1	0	1
(Binary)		(001)	(010)	(011)	(100)	(101)	(110)	(111)
Stripe Numbers		1	2	3	4	5	6	7

7 stripes in 3 images!

In general, we can do

$2^n - 1$ stripes in n Images

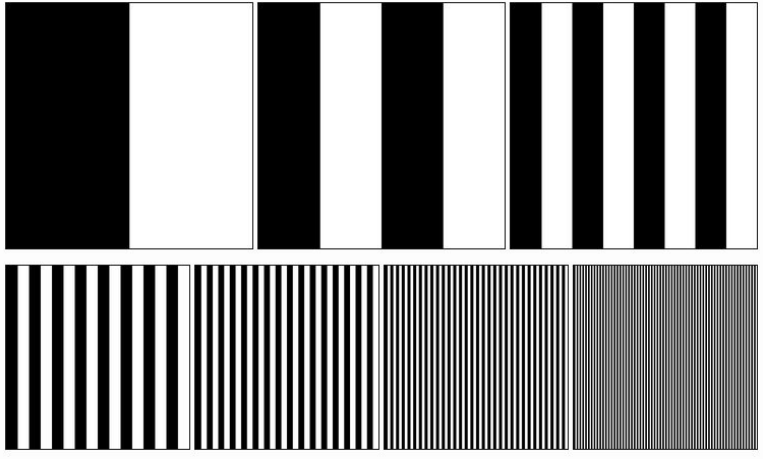
Note: (000) is not an option. Hence, the

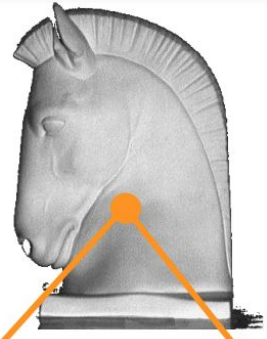
[Posdamer 1981]

Shree K. Nayar

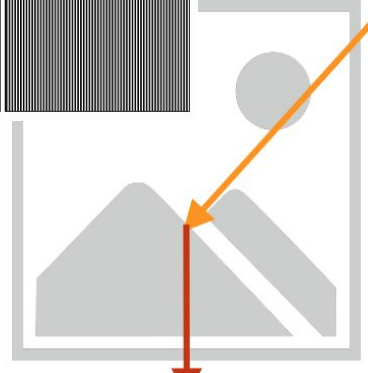
Más resolución

Con 7 bits:

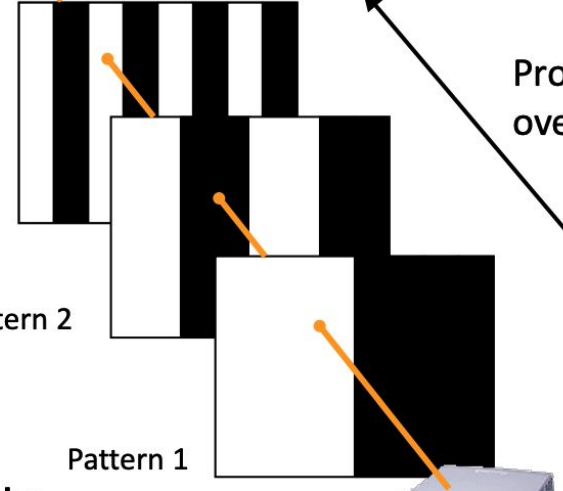




Example: 7 binary patterns proposed by Posdamer & Altschuler



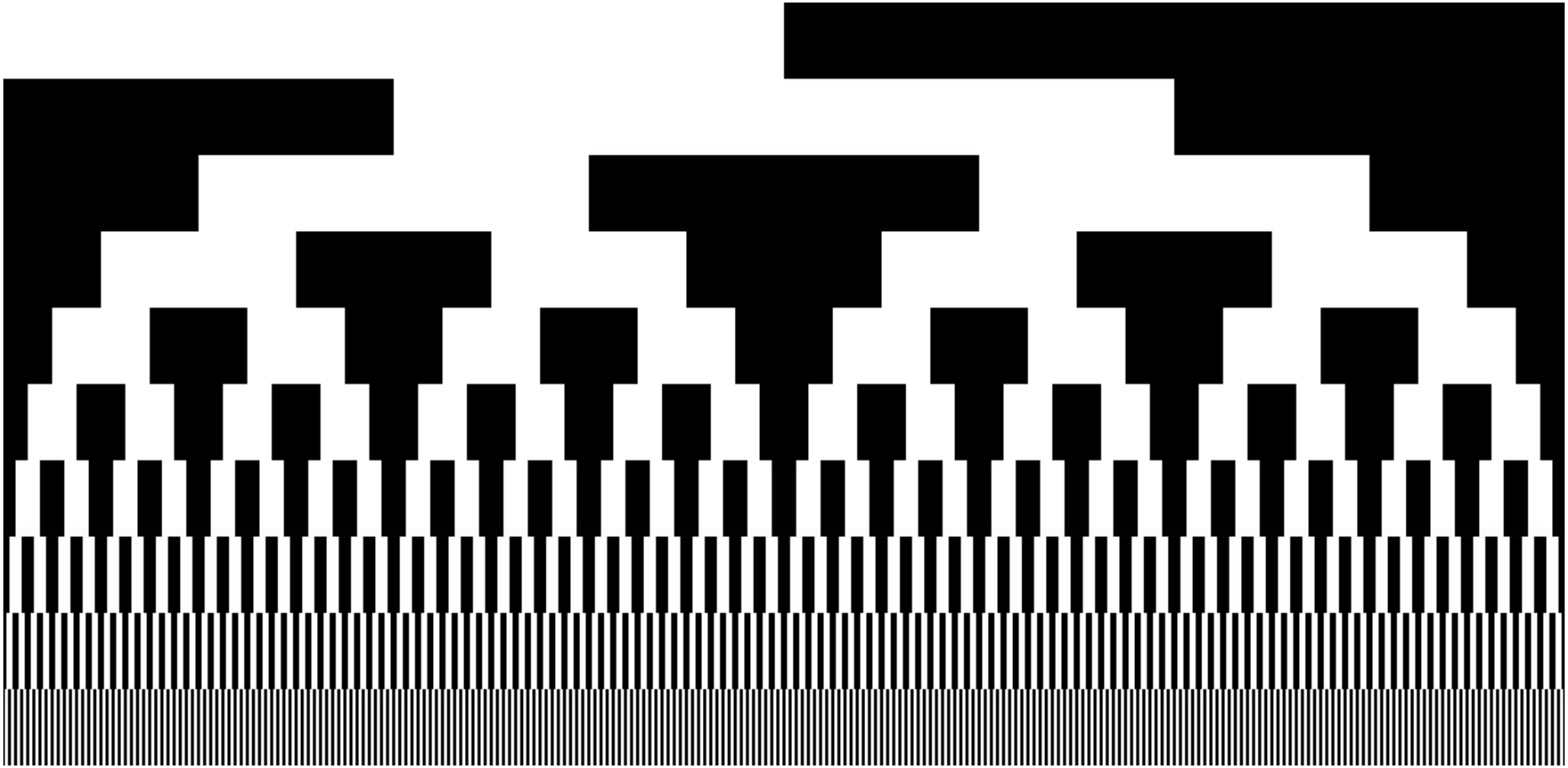
...



Projected over time

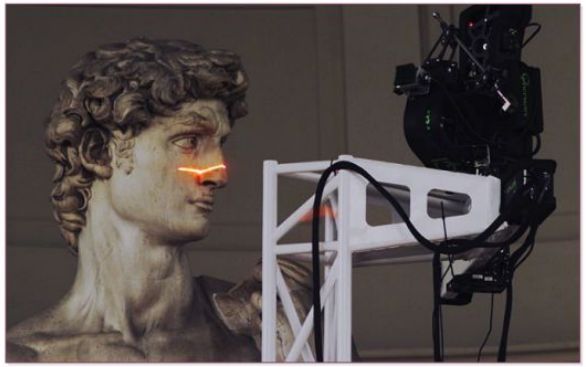
Codeword of this píxel: 1010010 → identifies the corresponding pattern stripe

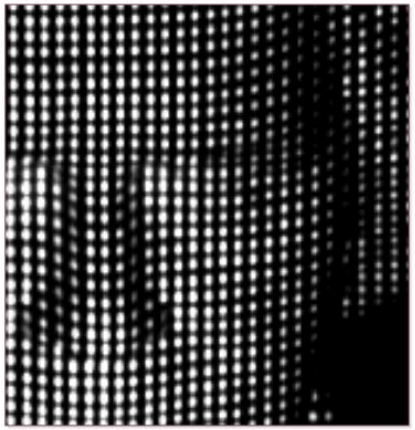
Con 10 bits:



Depende de la precisión que quieras, debes aumentar la resolución (#bits) de las líneas

Research topic: Reducir el número de patrones





Multi-stripe
Multi-frame

Single-stripe

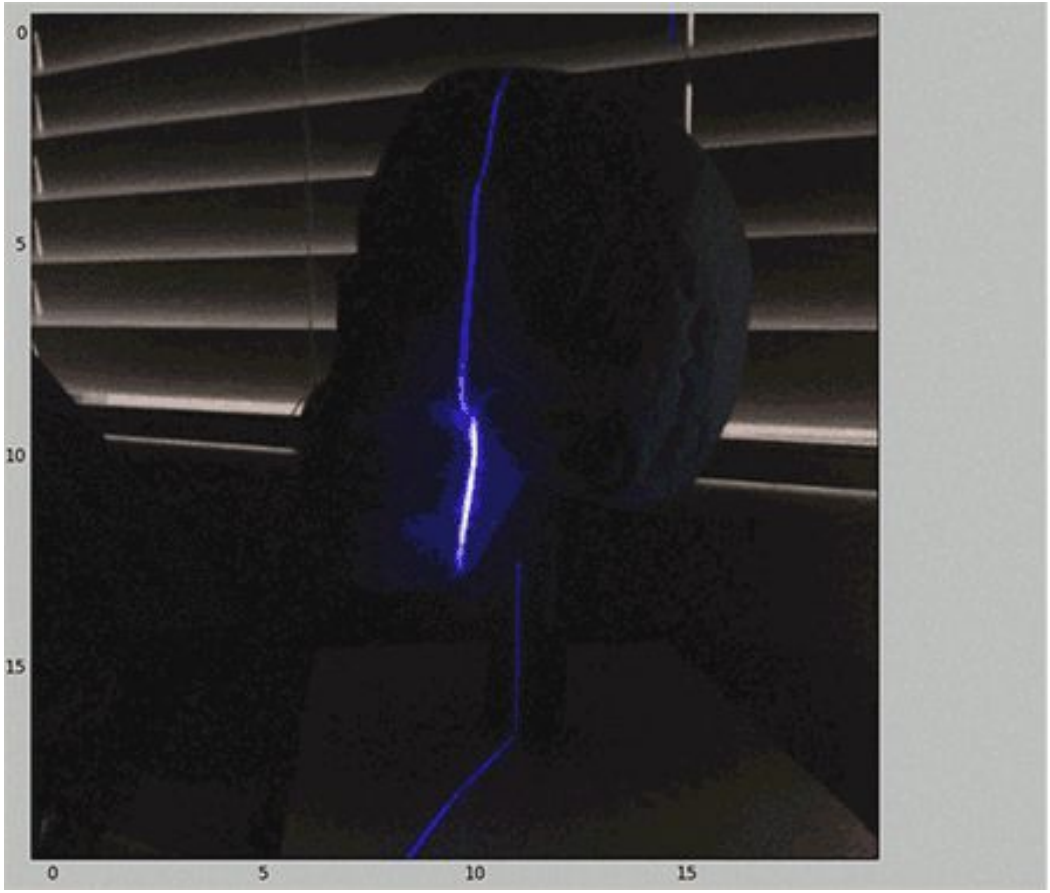
Single-frame

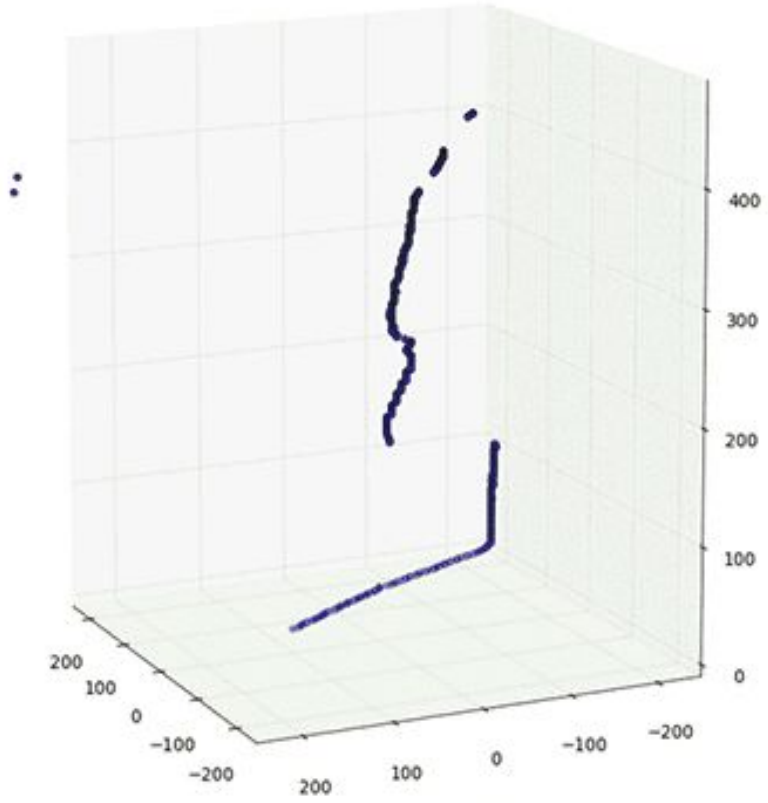
Slow, robust

Fast, fragile

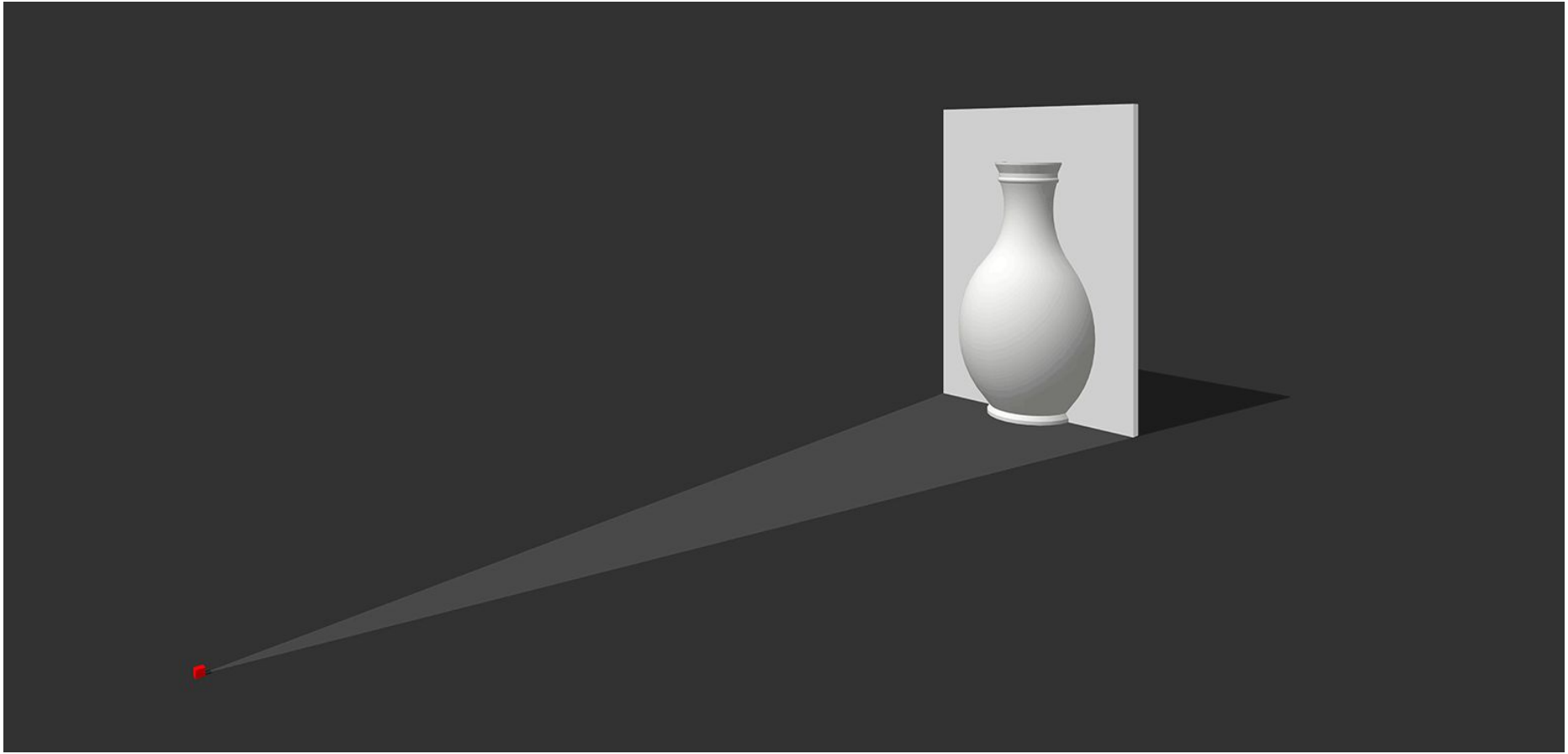
Method	Number of Images
Point based Structured Light	NM
Line based Structured Light	N
Binary Coded Structured Light	$\lceil \log_2(N+1) \rceil$
k-ary (Color) Coded Structured Light	$\lceil \log_k(N+1) \rceil$
Intensity Ratio Method	2
Phase Shifting Method	3

Proyectando 1 sola línea



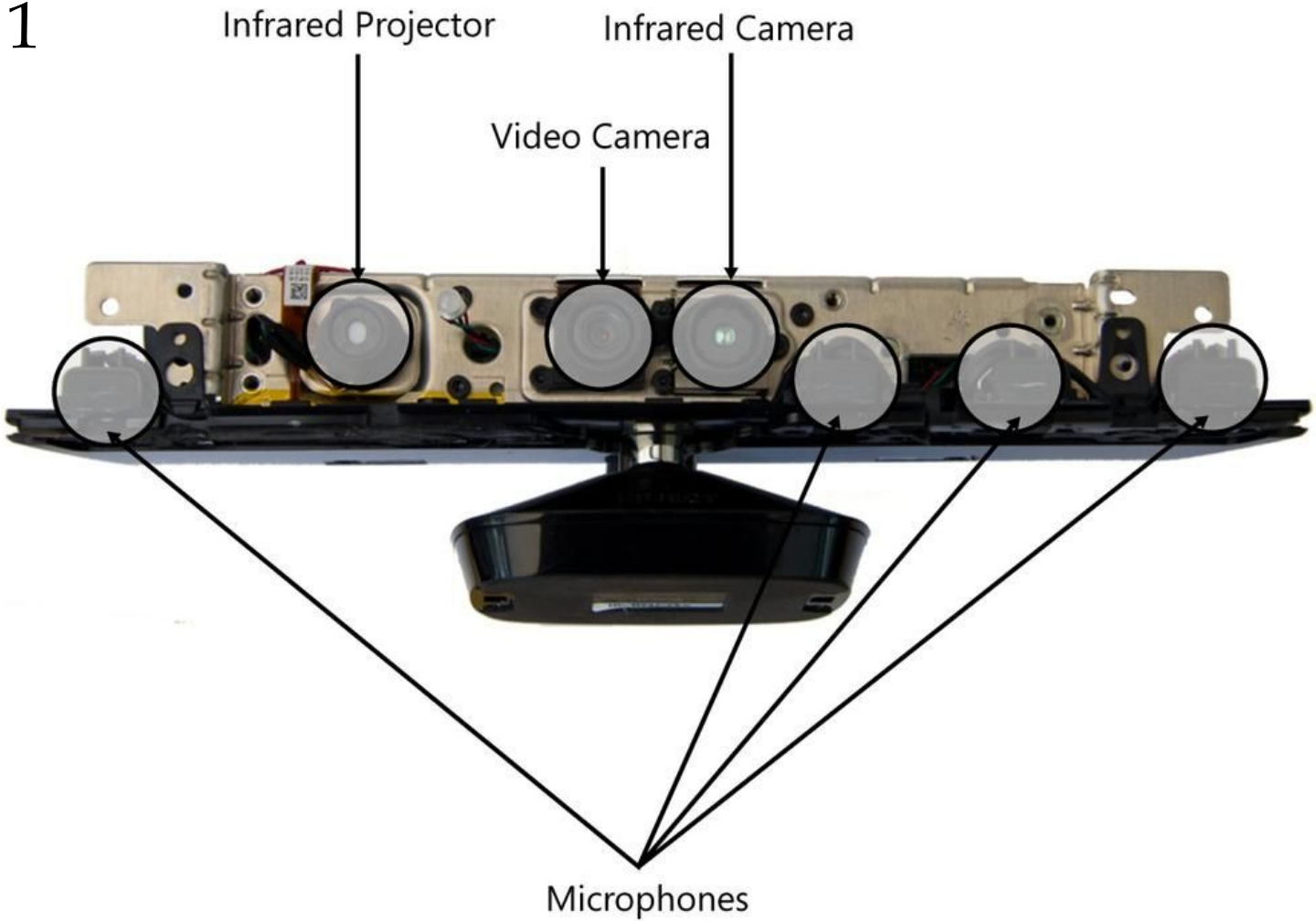


Proyectando múltiples líneas

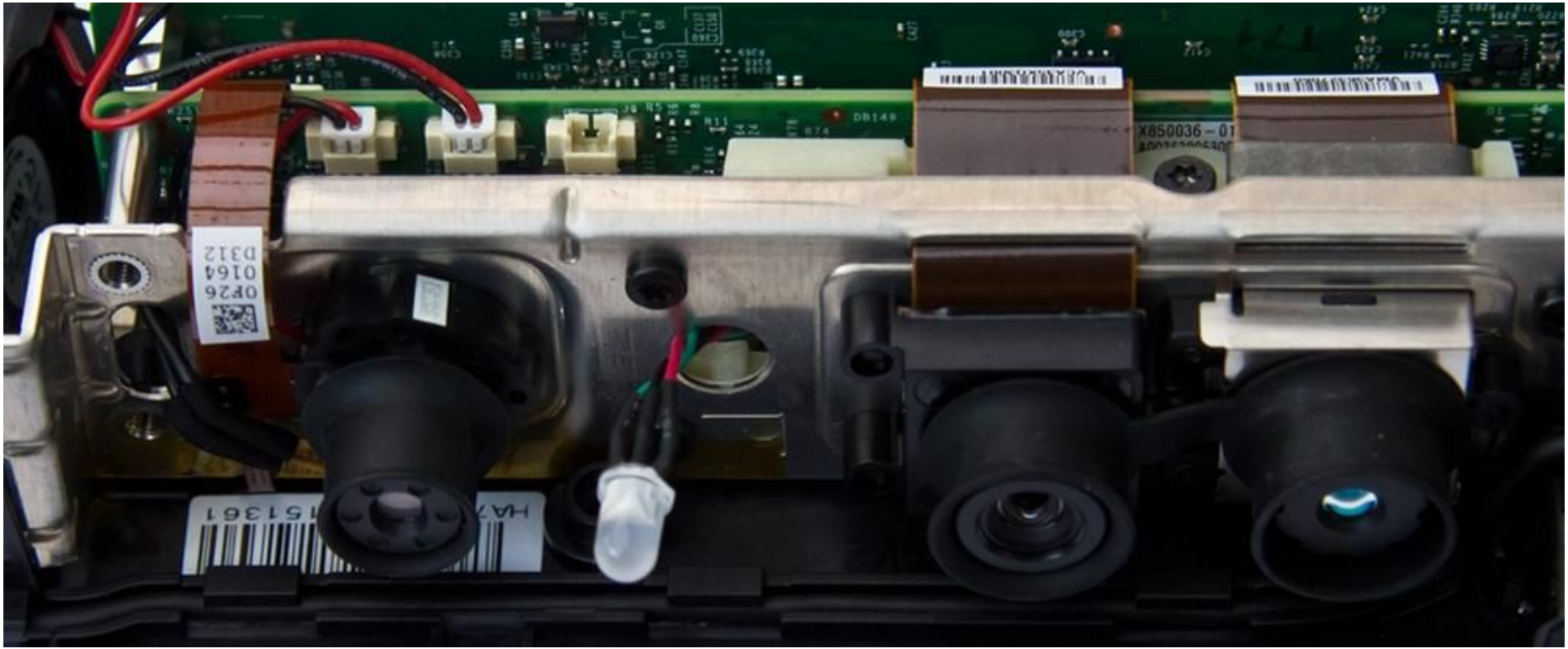


Kinect 1:
Released on 2010
with Xbox 360

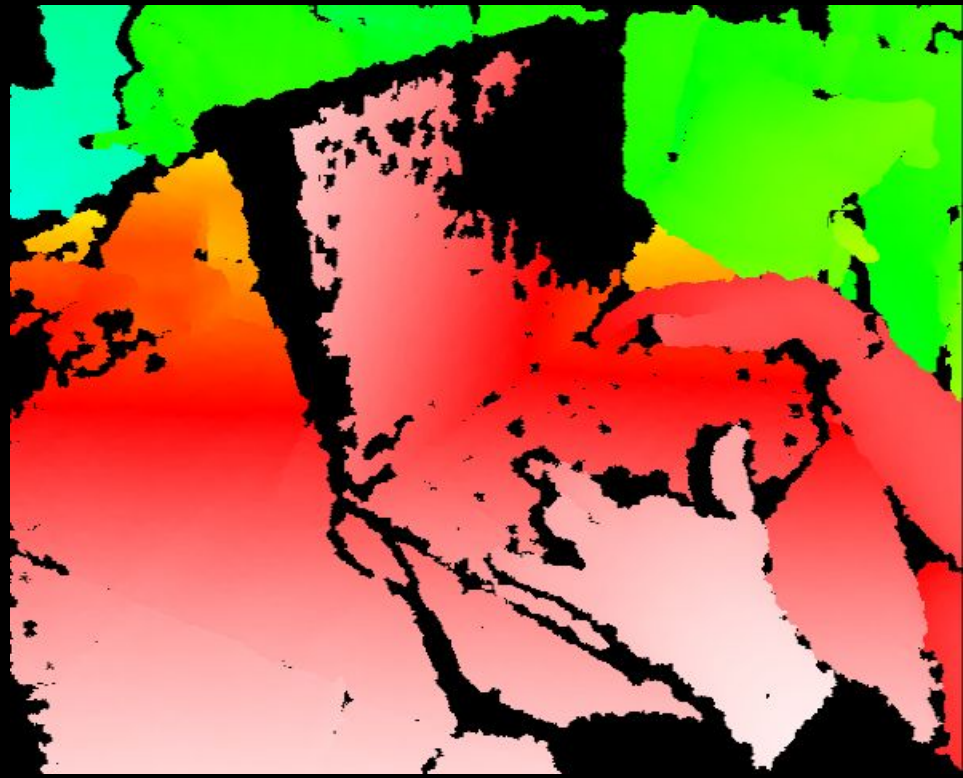
Kinect 1


Infrared Projector
Infrared Camera
Video Camera
Microphones

Kinect 1



Patrón de proyección (IR) y Mapa de Profundidad

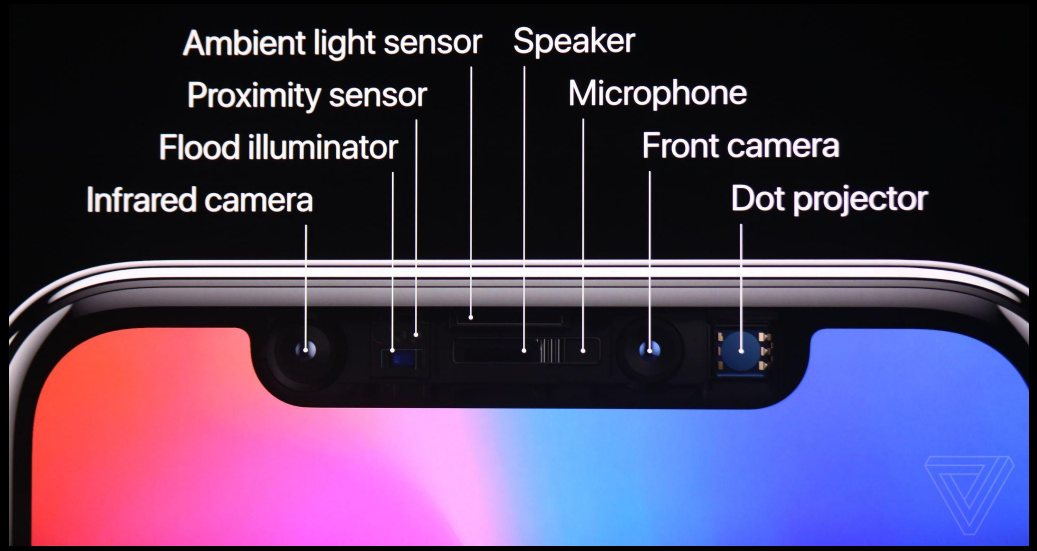


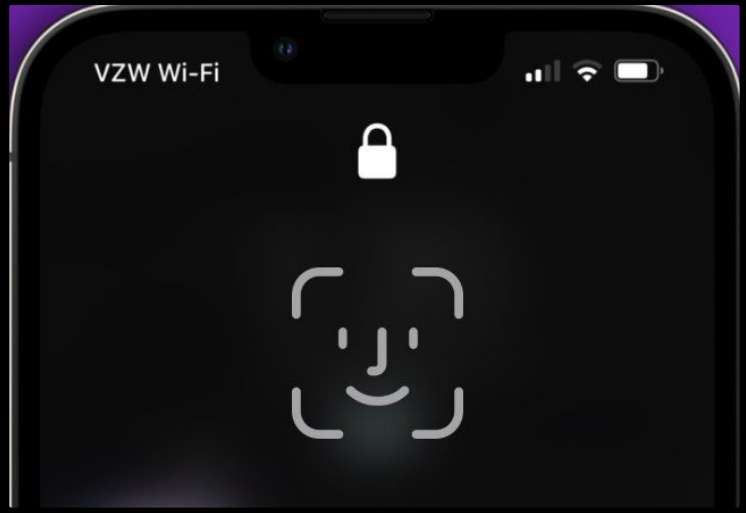

Kinect under IR camera


TECH
INSIDER
INSIDER

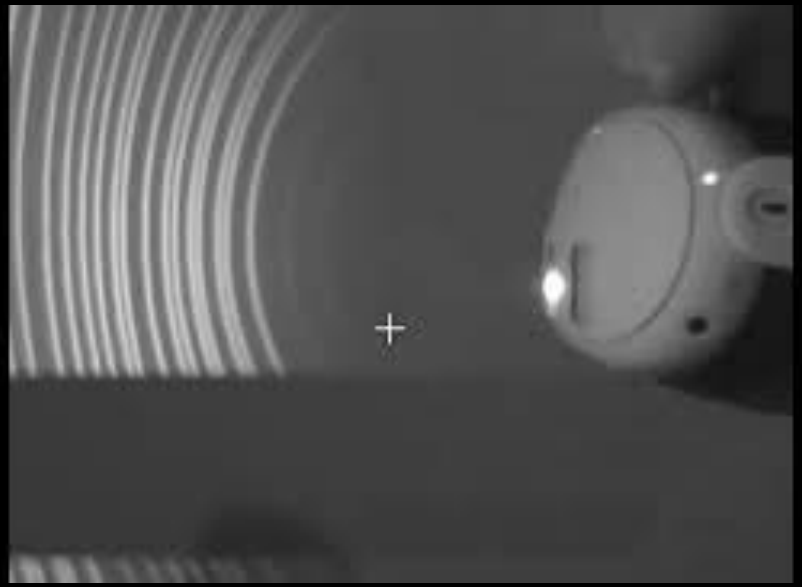
Dot projector iPhone XR

Face ID


Ambient light sensor
Speaker
Proximity sensor
Microphone
Flood illuminator
Front camera
Infrared camera
Dot projector


VZW Wi-Fi

Meta Quest 3 (Line Projector)

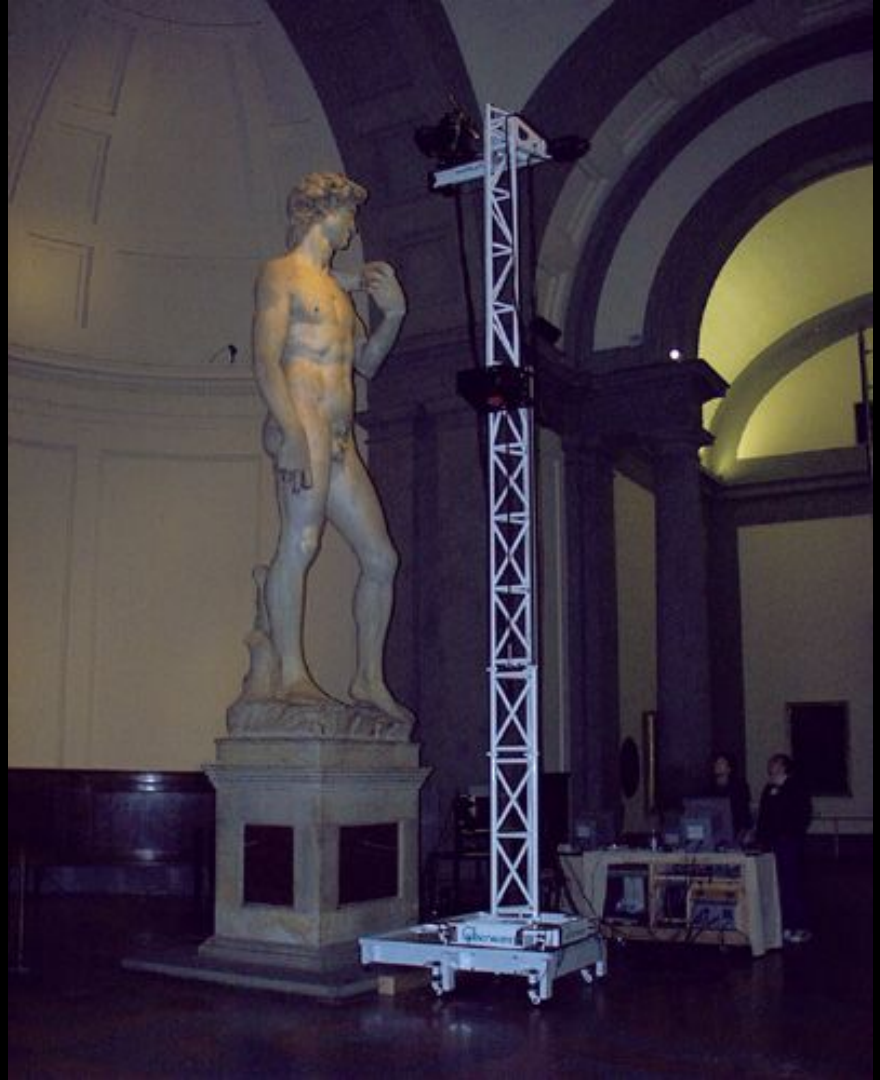


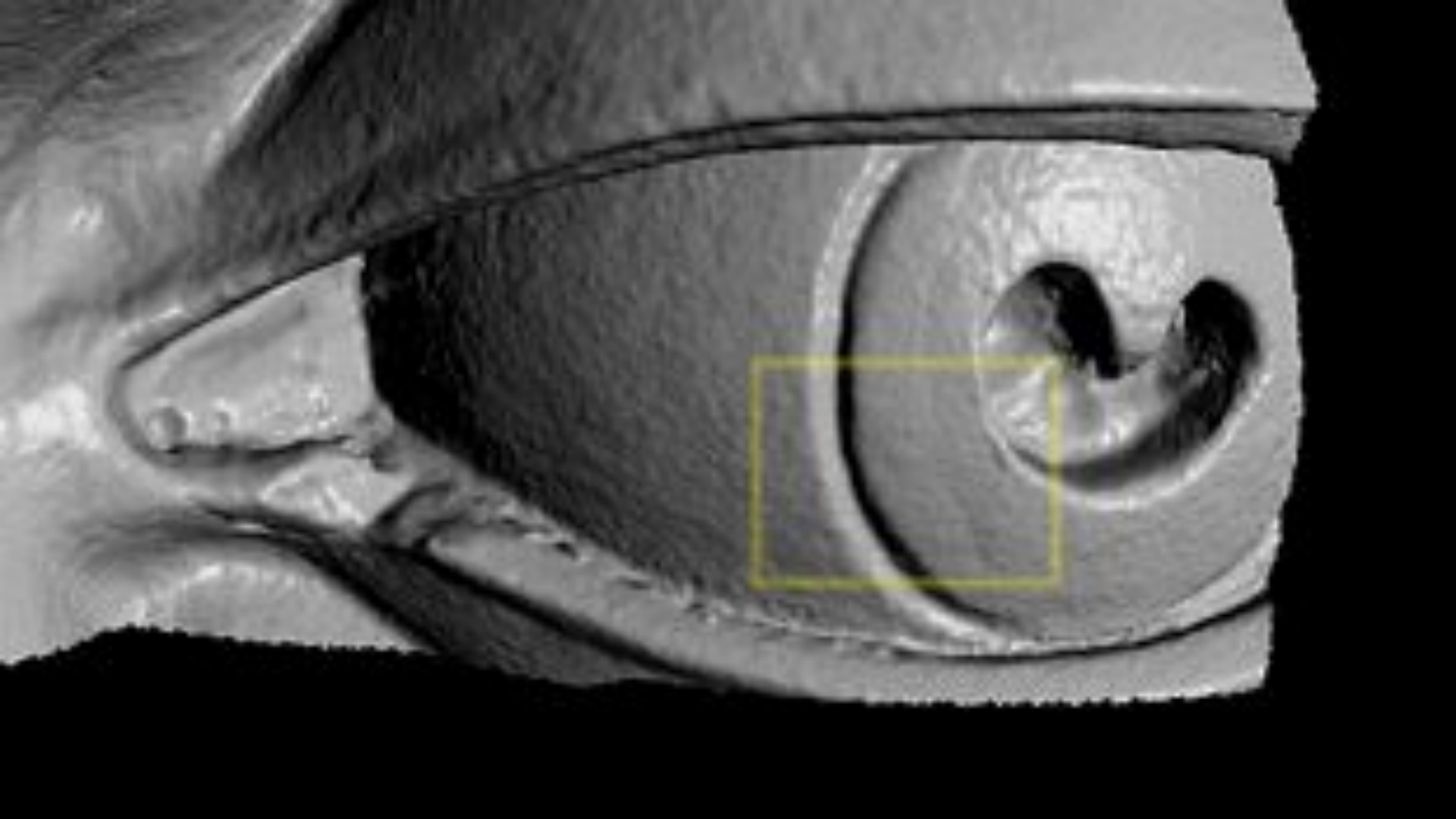
Projection Mapping

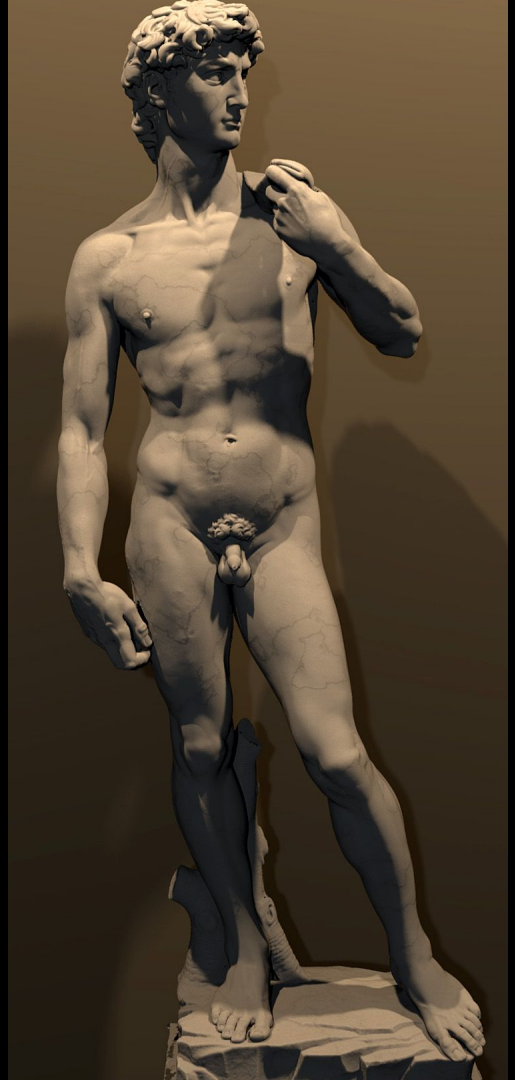

BEACONS OF MAGIC
CINDERELLA CASTLE
DISNEY
WORLD 50TH
ANNIVERSARY

Digital Michelangelo Project

https://accademia.stanford.edu/mich/

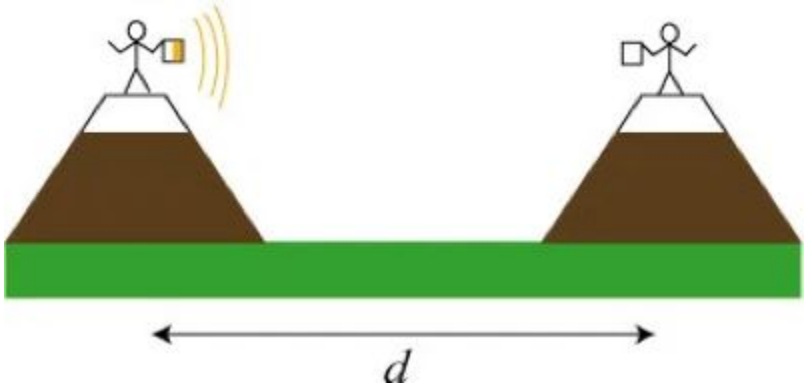




3. Time of Flight

Velocidad de la luz



¿ Recuerdas?

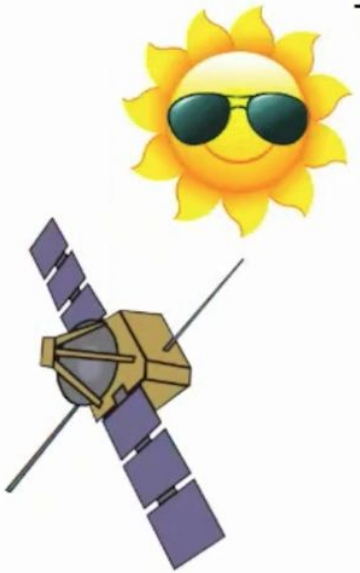
x=vt

Circunferencia ~ 40.000 km -> Δt ~ 0.13 s

Velocidad de percepción humana > 0.1 s

Velocidad de la luz

Photons travel at the speed of light (the fastest you can go): **300000km/s!**

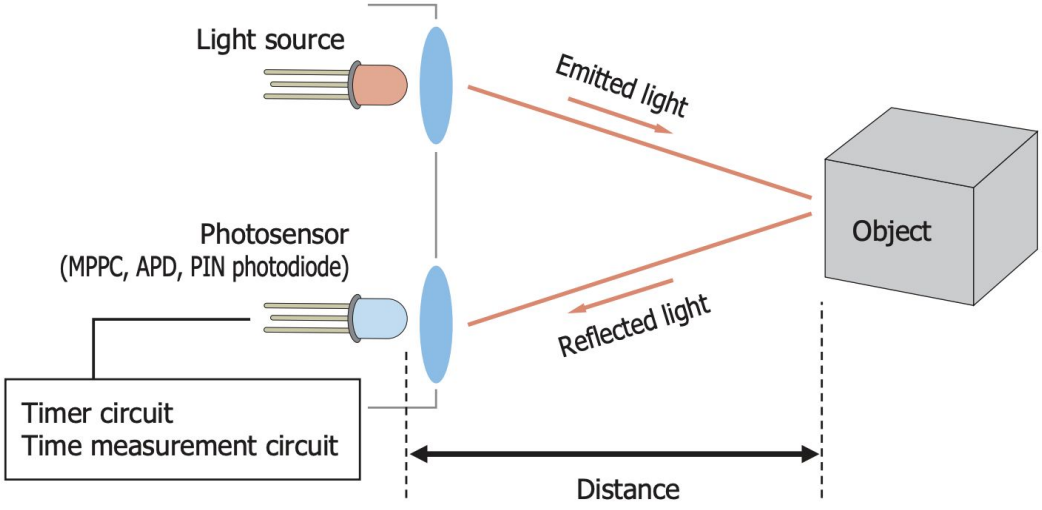


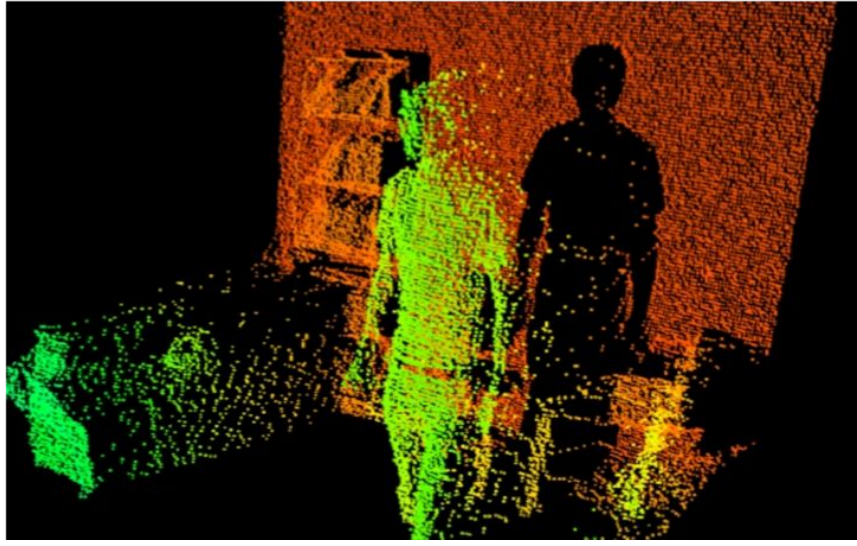
To get an idea, here the time light needs to travel:

From sun to earth: **8mins** = 480s
From earth to moon: **1.28s**
To a geostationary satellite and back: **0.24s**
From the Mont Blanc to us (80km): **0.00026s**
From here to prod building (10m): **0.000000033s**
Between you (1m): 0.0000000033s = **3.3ns**
1mm: 0.3ps = $0.3*10^{-12}$s = **0.000000000003s**

Rule of thumb: 30cm/ns

2024: Relojes veloces

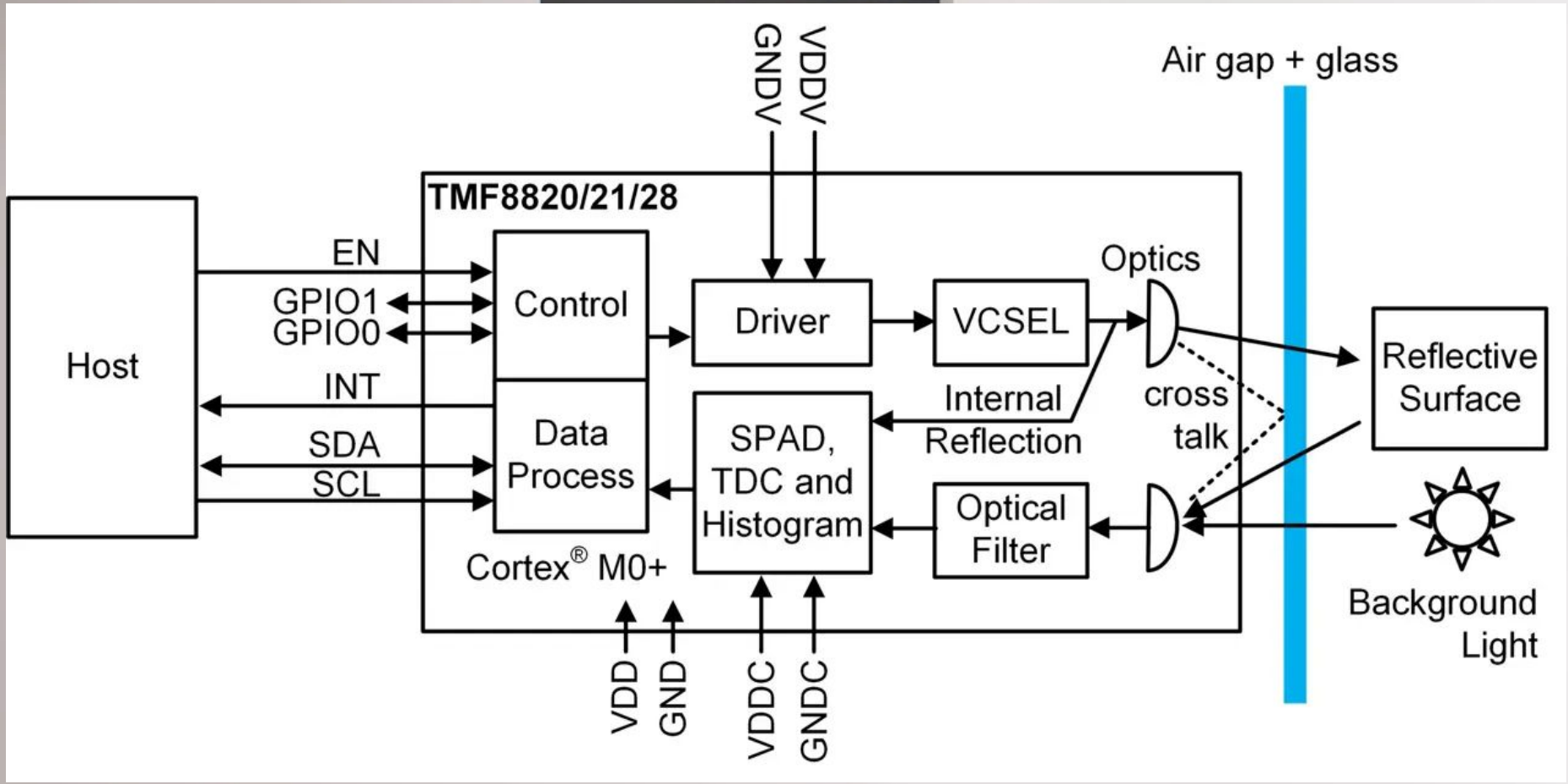




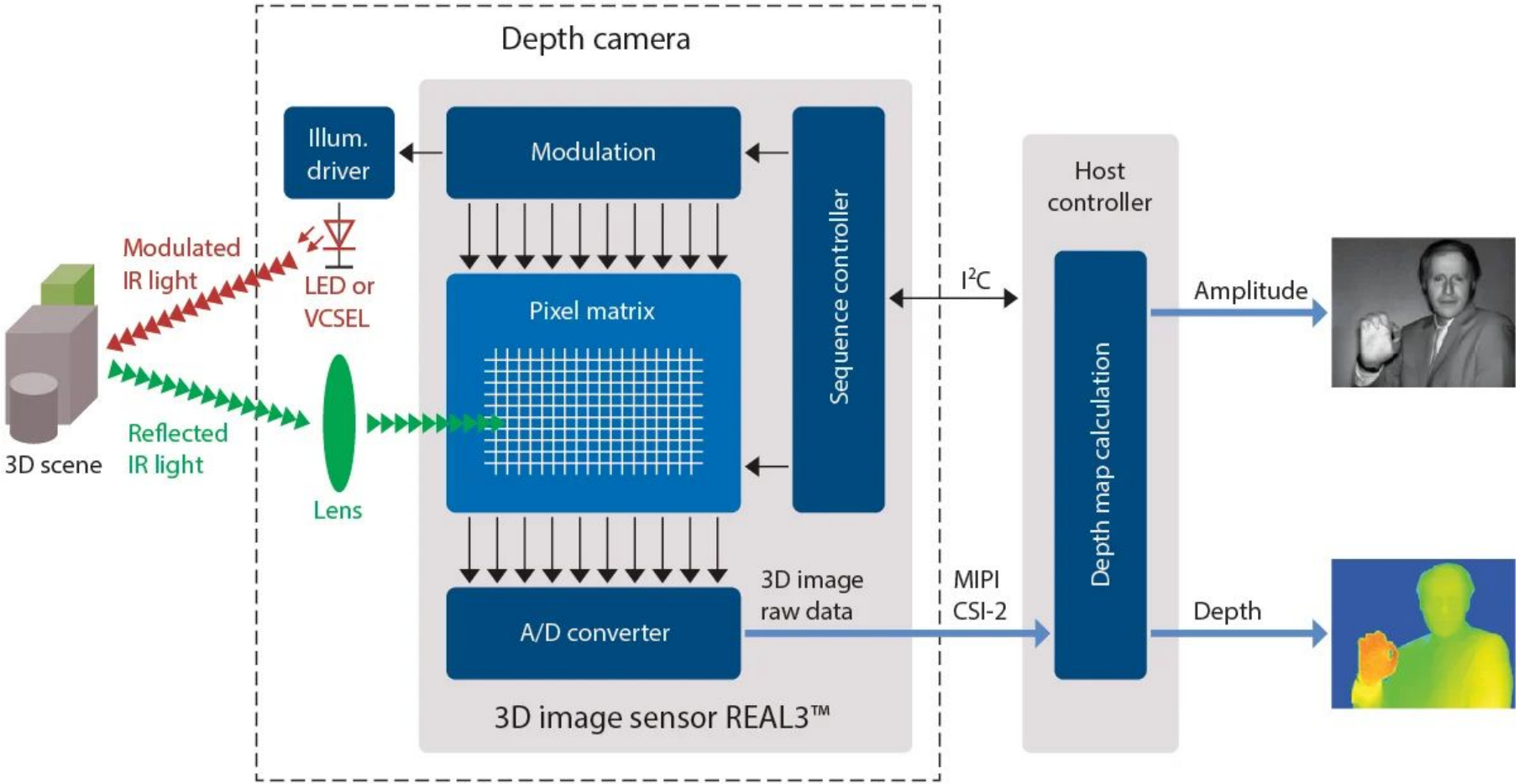
- Distancia recorrida: $x = v.t$
- Tiempo de vuelo: $t = 2x / c$
- Distancia al objeto: $x = c.t/2$

Principio de operación

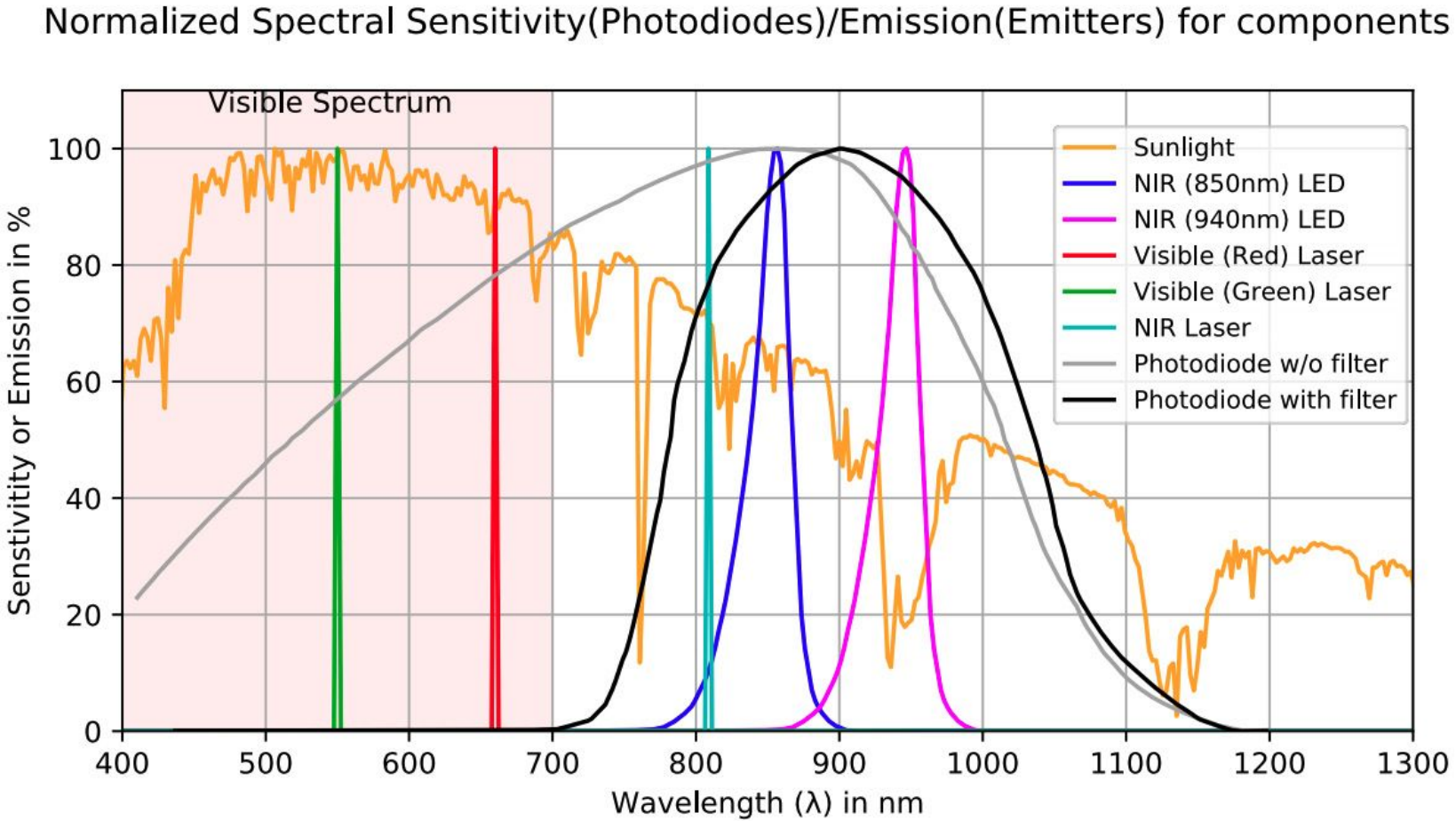
Looking inside



Time-of-Flight principle and block diagram



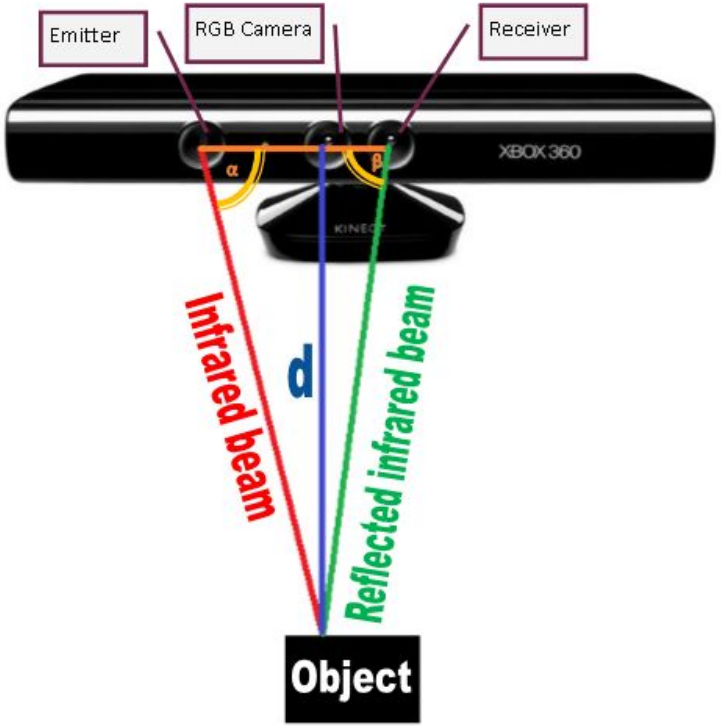
https://www.allaboutcircuits.com/technical-articles/how-do-time-of-flight-sensors-work-pmdtechnologies-tof-3D-camera/

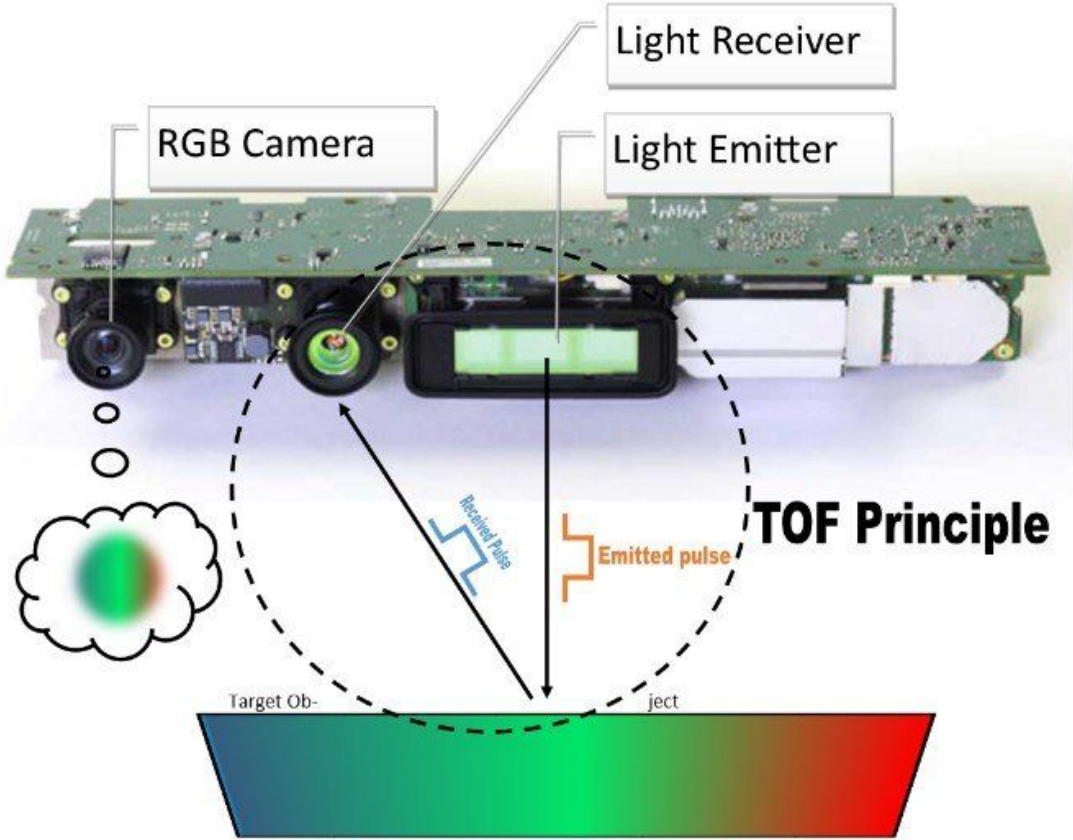


Suelen operar en el espectro infrarrojo. Esto permite minimizar la interferencia de la luz ambiente mediante el uso de un filtro pasa banda infrarrojo en el receptor, y que el sistema parezca “invisible” para los usuarios.

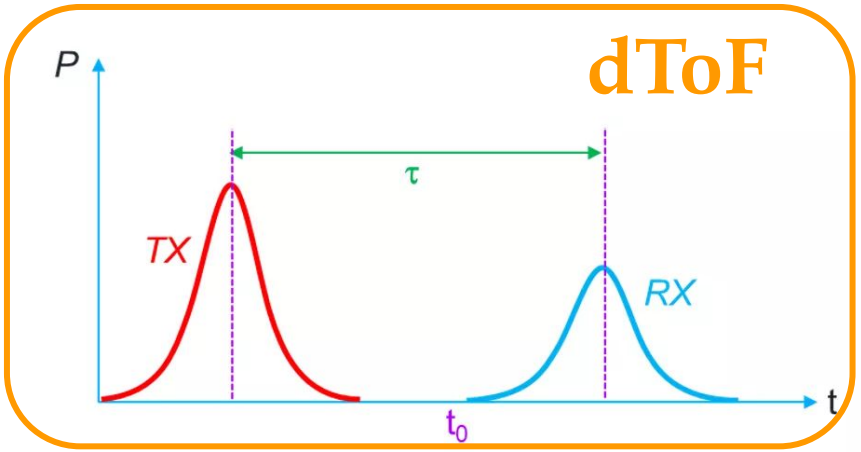

WHAT IS ToF?

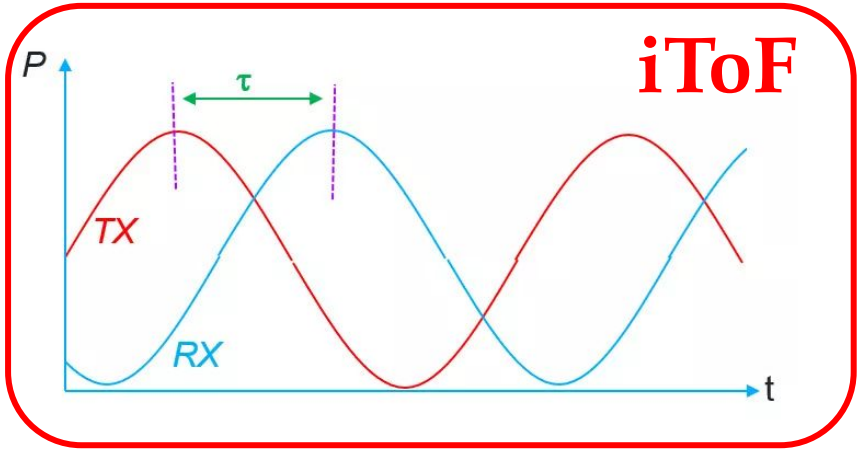
Kinect 1 (Luz estructurada) vs Kinect 2 (ToF)

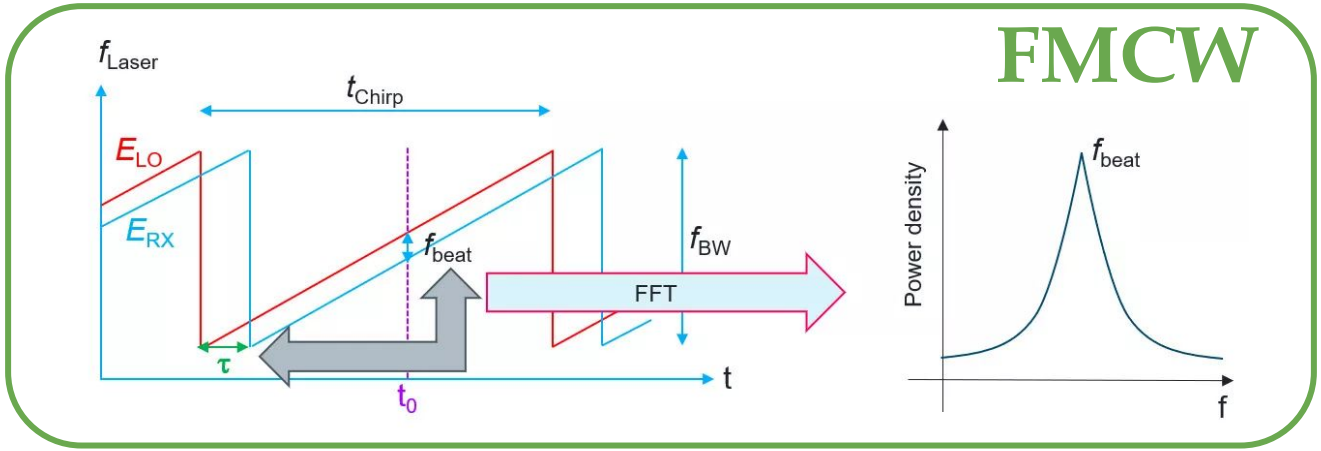




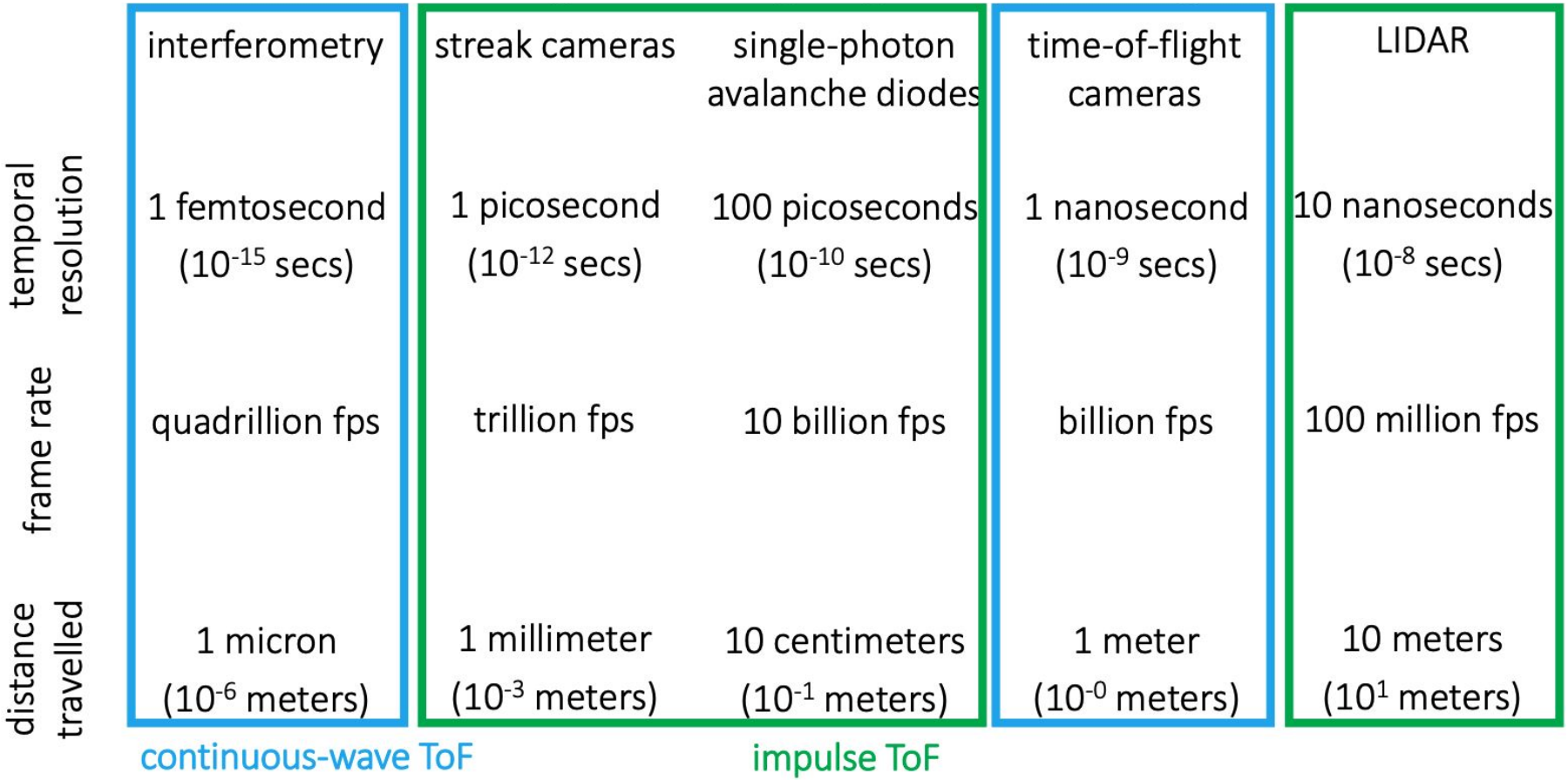
Técnicas para medir el tiempo de vuelo







Tecnología para ToF



Name		iToF sensor	dToF sensor	Flash LIDAR	TSS LIDAR	Scanning LIDAR	FMCW LIDAR
System type	Measurement principle	iToF	dToF	dToF	dToF	dToF	FMCW
	Scanning architecture	Single emitter + detector array	Single emitter + detector array	Single emitter + detector array	Emitter array + detector array	Scanning mirror(s)	Scanning mirror(s)
	Optical aperture / power	Compact: short range, low power	Compact: short range, low power	Longer range: bulk optics & discretes	Longer range: bulk optics & discretes	Longer range: bulk optics & discretes	Longer range: bulk optics & discretes
Typical performance*	Range	<5m	<10m	<100m	<100m	~200m	~300m
	Resolution	Image sensor	<100 points	<100 points	<100 points	Scanning mirror dependent	Scanning mirror dependent
	Robustness	Crosstalk & multipath, ambient interference	Ambient interference	Ambient interference	Ambient interference	Ambient interference	Ambient immune
Example applications		3D depth camera	Auto-focus Touchless controls	Industrial robotics & automation	Industrial robotics & automation	Autonomous vehicles	Specialized high value applications
ams-OSRAM products		VCSELs	dToF modules VCSELs & VCSEL modules	VCSELs Edge Emitting Lasers	VCSELs Edge Emitting Lasers	VCSELs Edge Emitting Lasers	(Custom emitters)

Note: System types and parameters given for purposes of high-level comparison only.

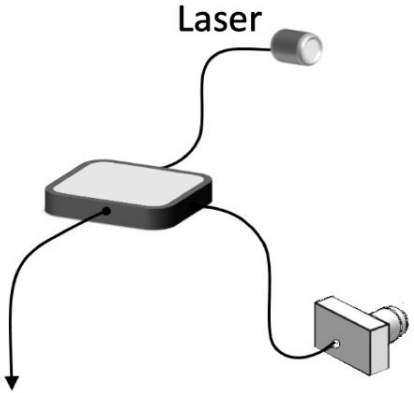
Many range measurement systems exist in practice.

*Key:

Low	Medium	High

dToF = Direct Time of Flight

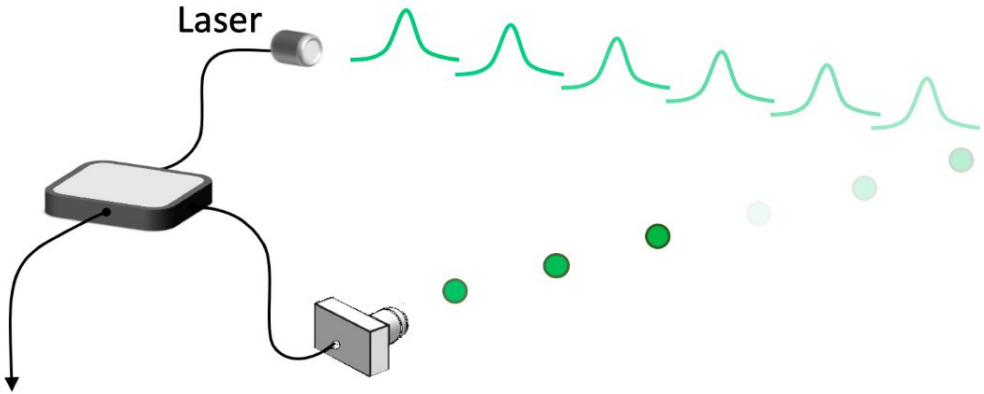
dToF = Direct Time of Flight



Single-photon avalanche diode (SPAD)

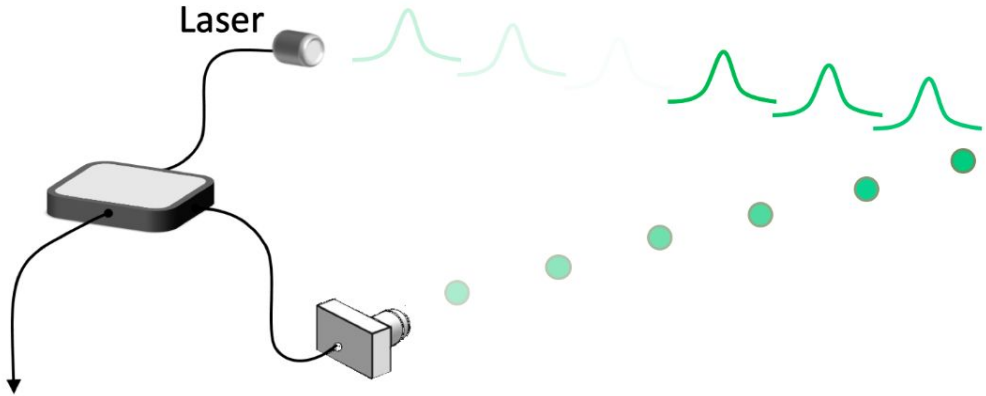
Scene

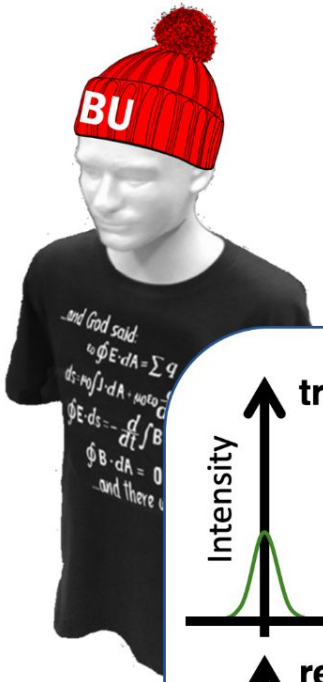
dToF = Direct Time of Flight

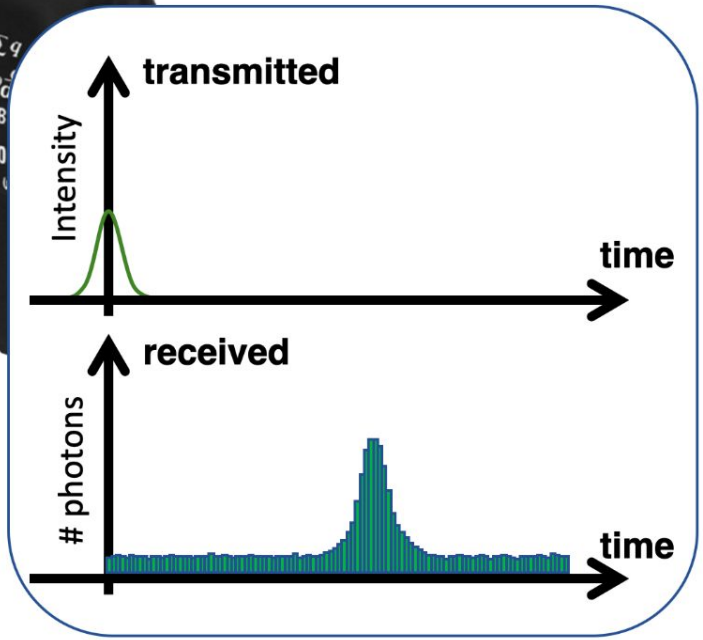


Scene

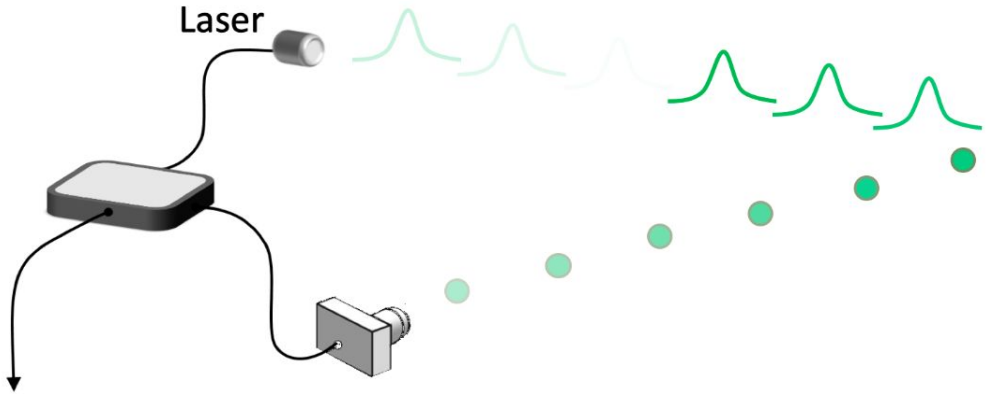
dToF = Direct Time of Flight





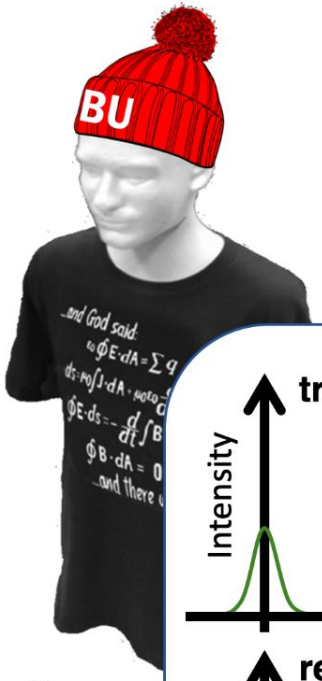


dToF = Direct Time of Flight

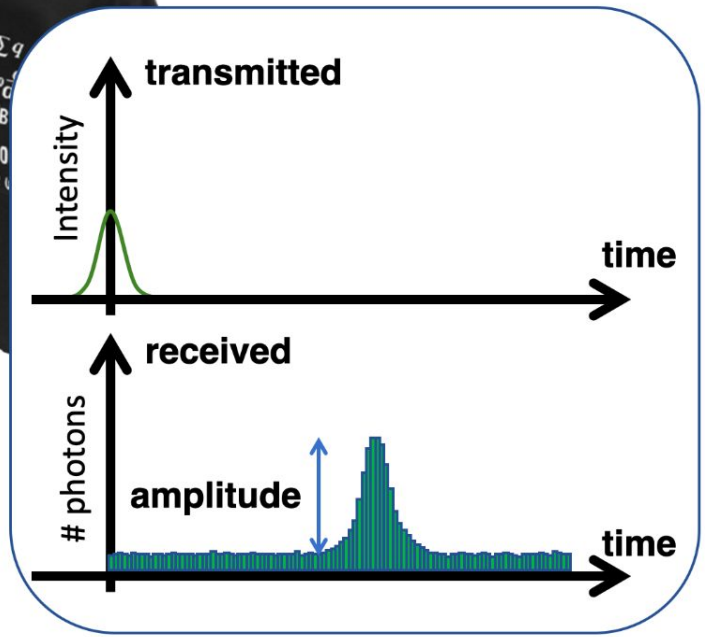


Single-photon avalanche diode (SPAD)

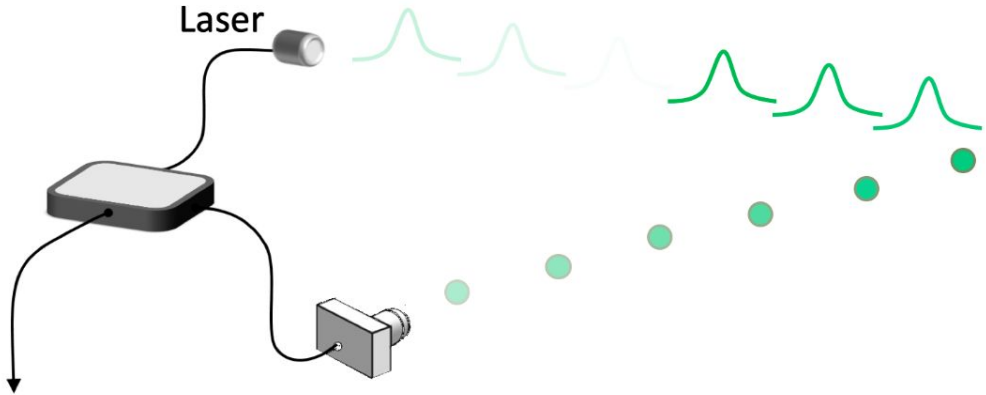
reflectivity image



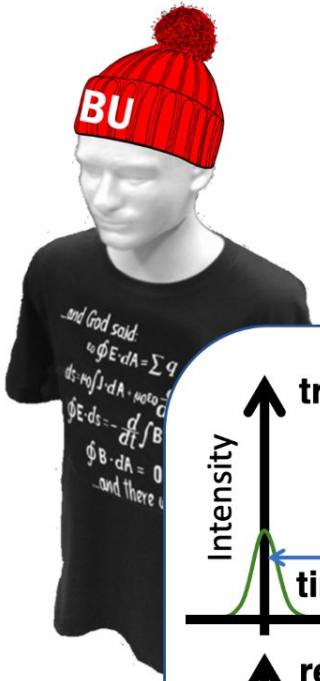
Scene



dToF = Direct Time of Flight

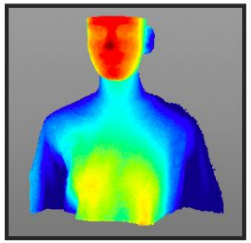


Single-photon avalanche diode (SPAD)

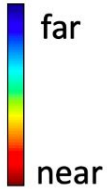


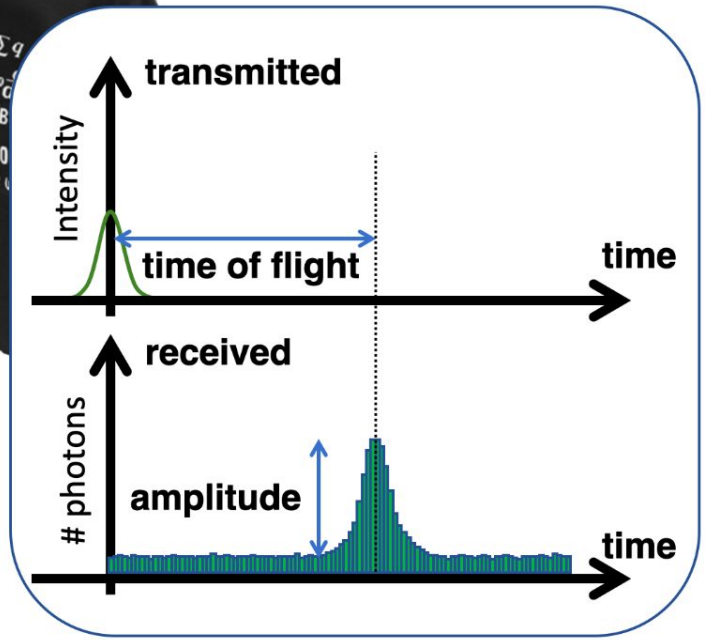
reflectivity image

depth image

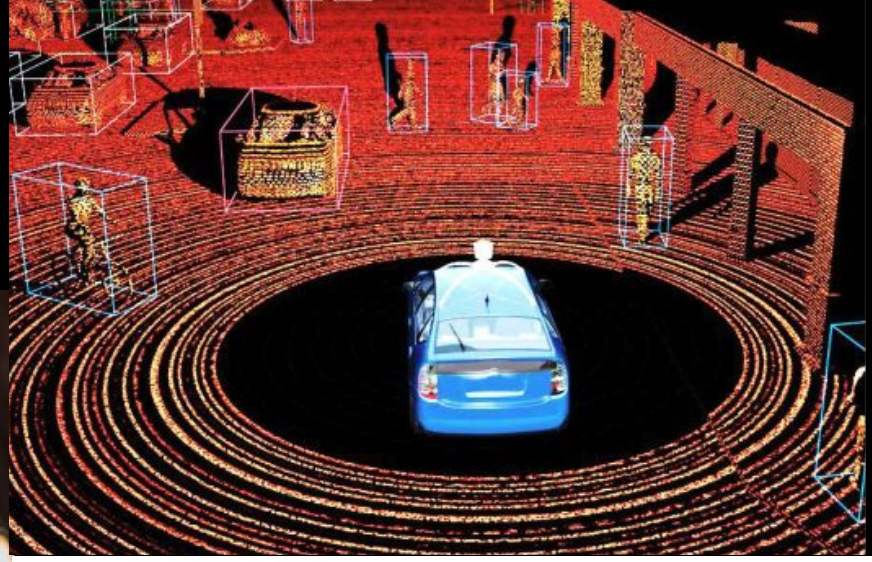


Scene

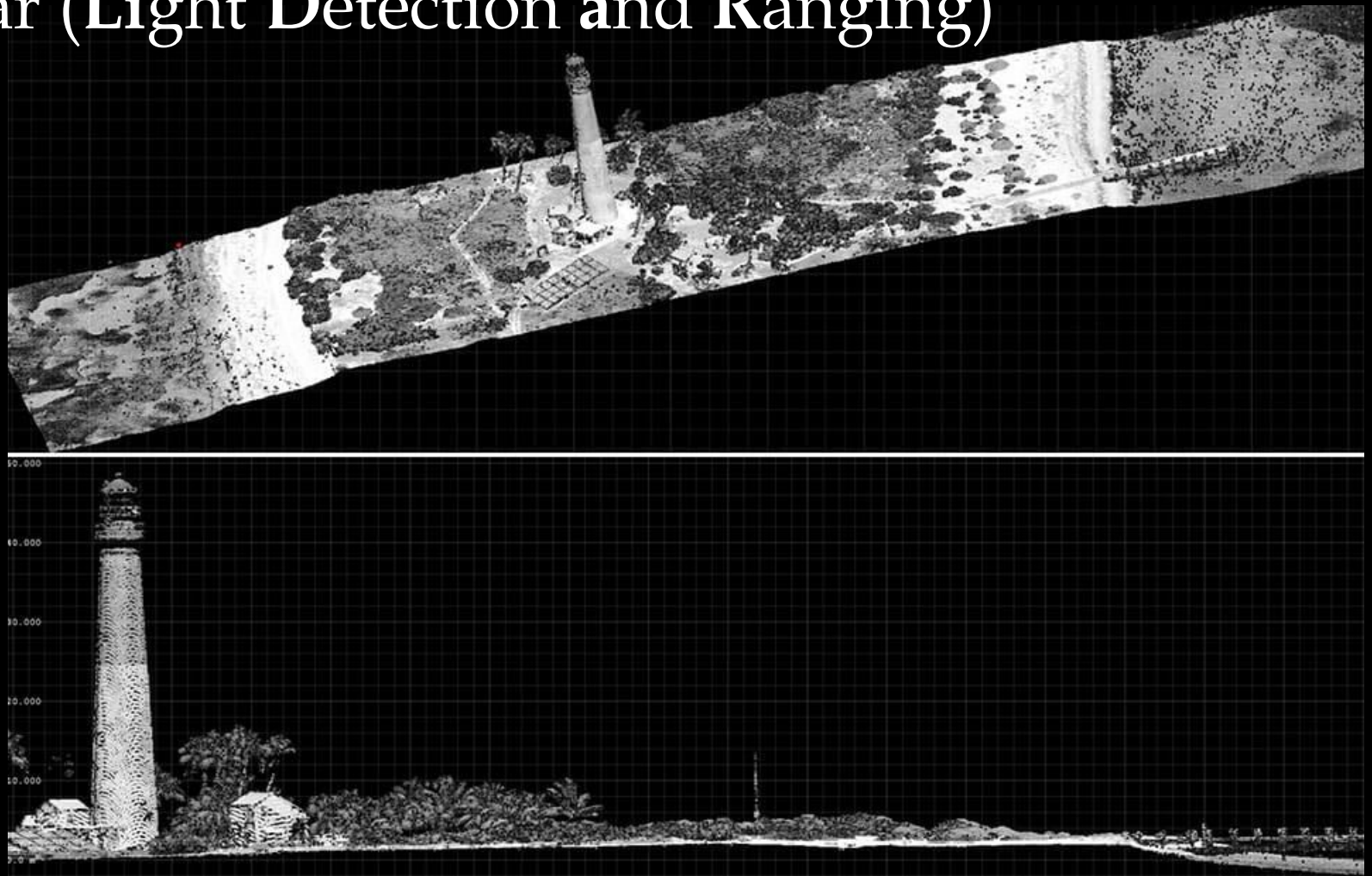




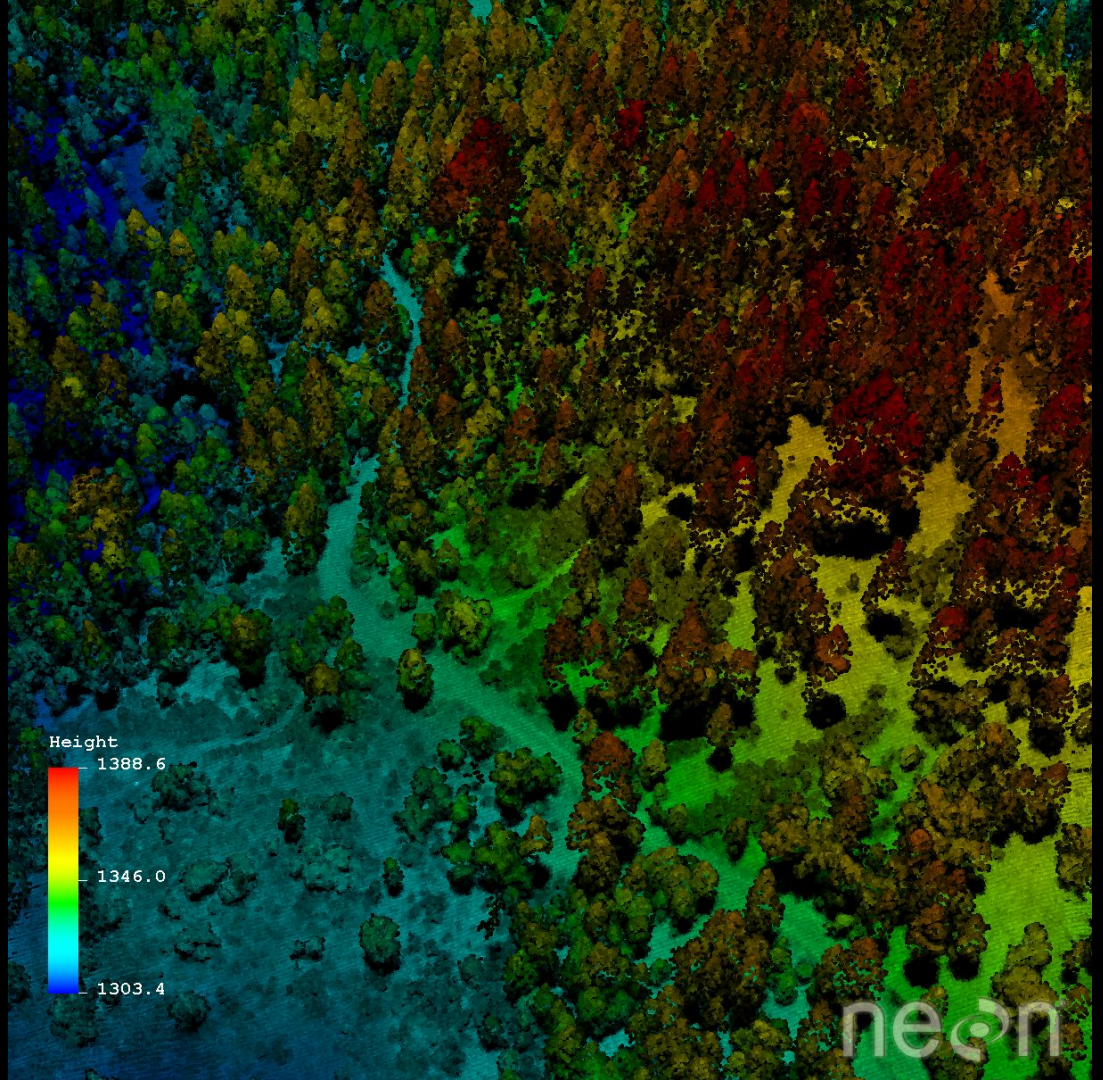
Vehículos autónomos



Lidar (**Li**ght **D**etection **a**nd **R**anging)



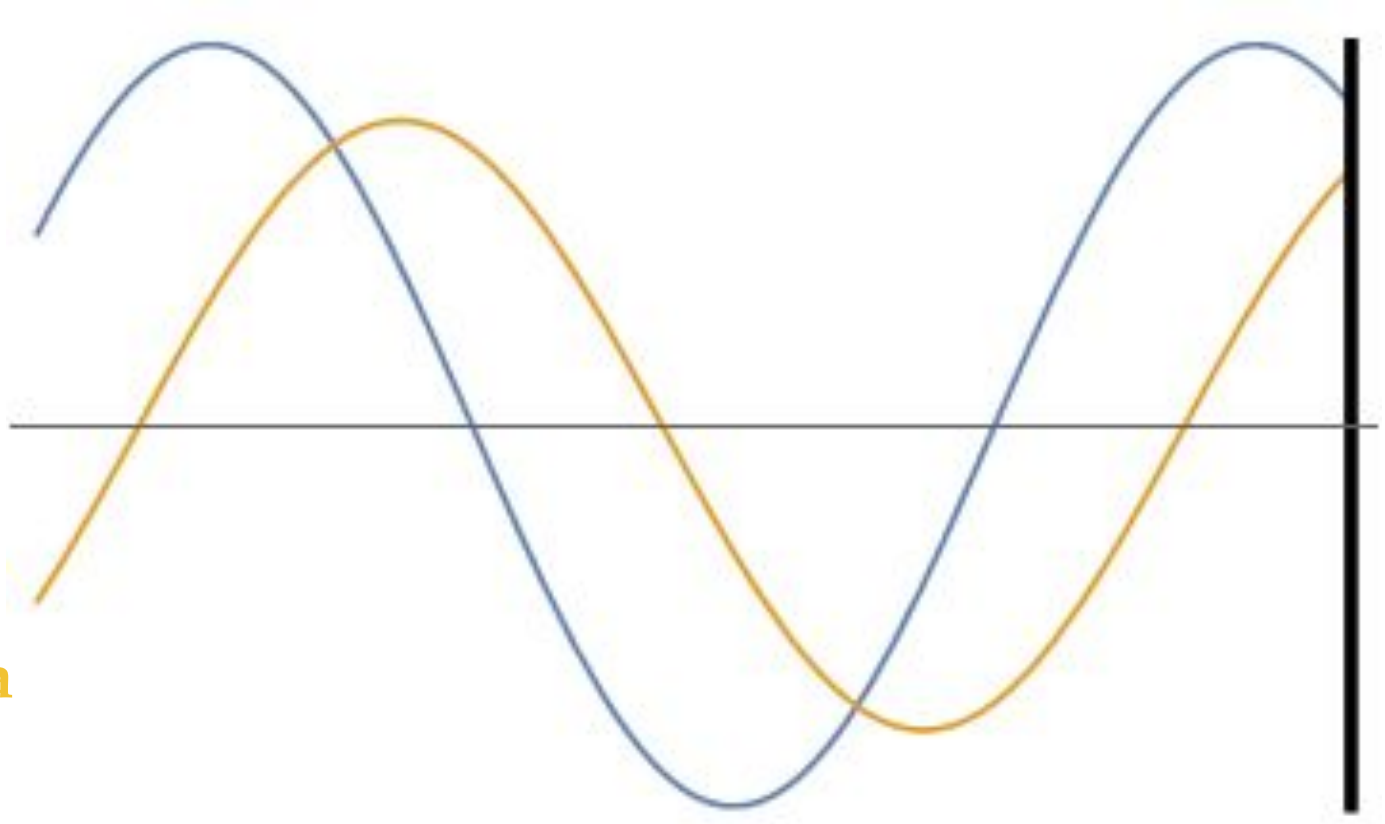
Lidar



iToF = Indirect Time of Flight

iToF = Indirect Time of Flight

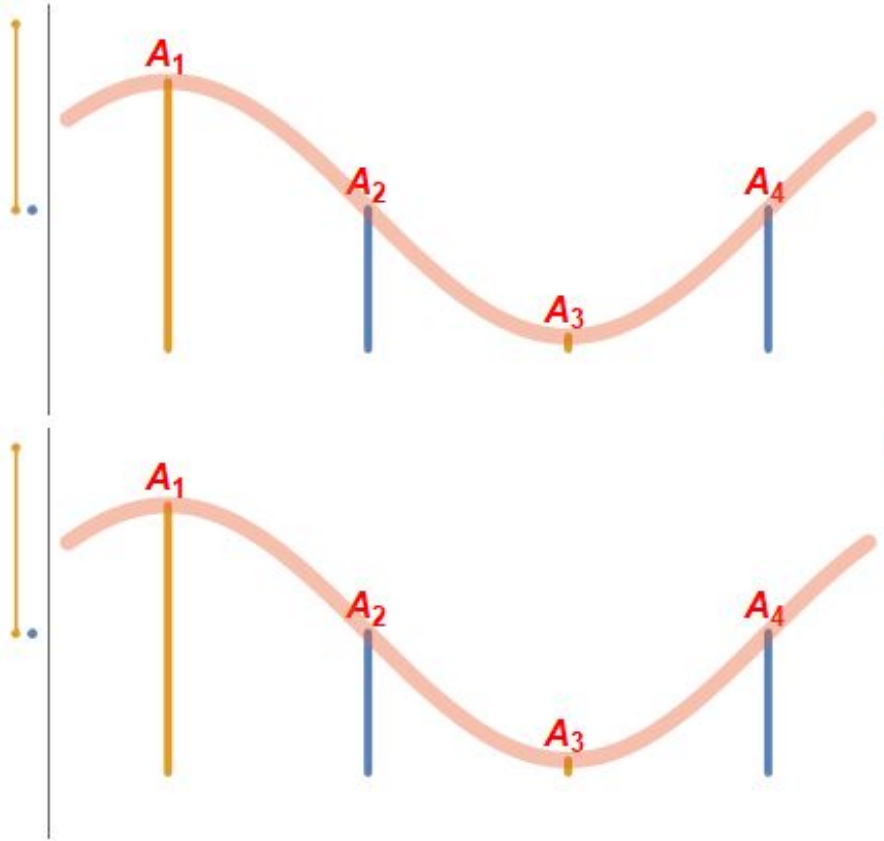
Luz emitida

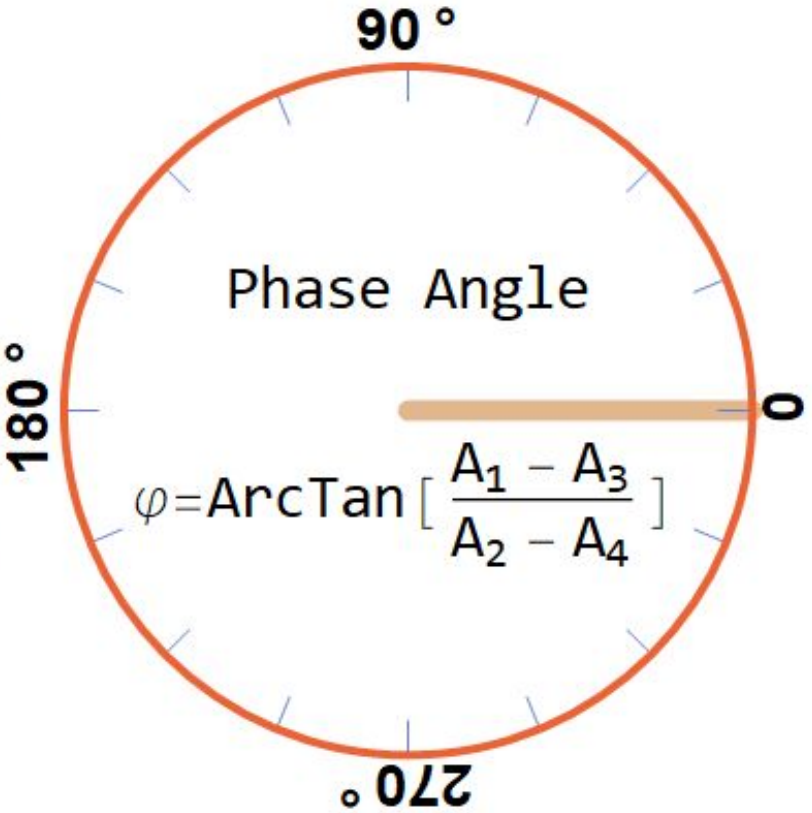


Luz reflejada

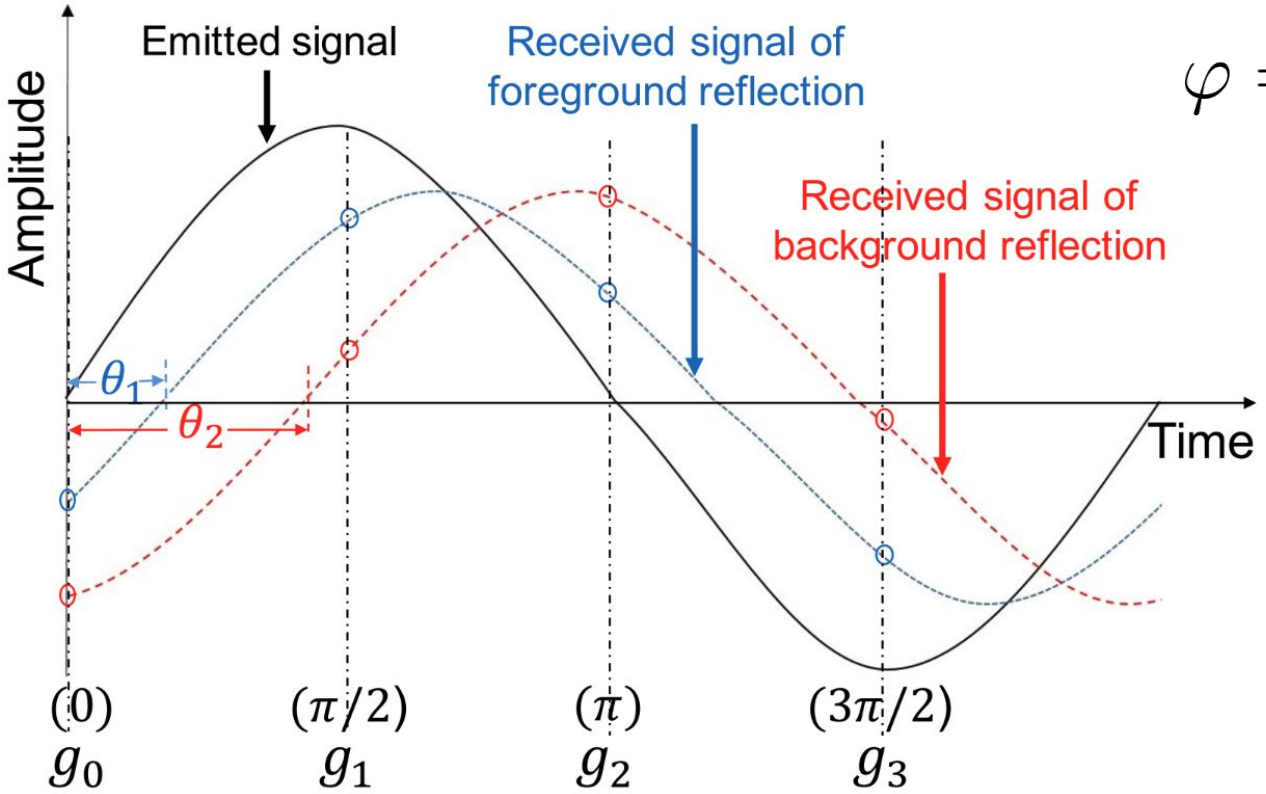
Nota: La onda cambia de dirección y disminuye su amplitud al interactuar

iToF = Indirect Time of Flight



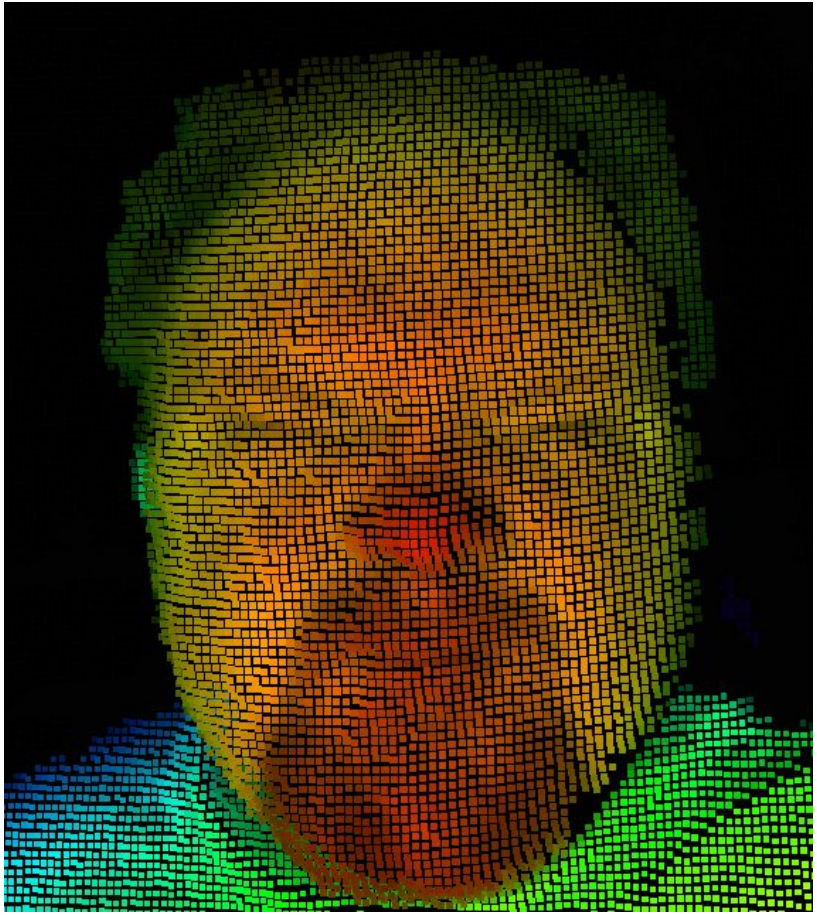


iToF = Indirect Time of Flight



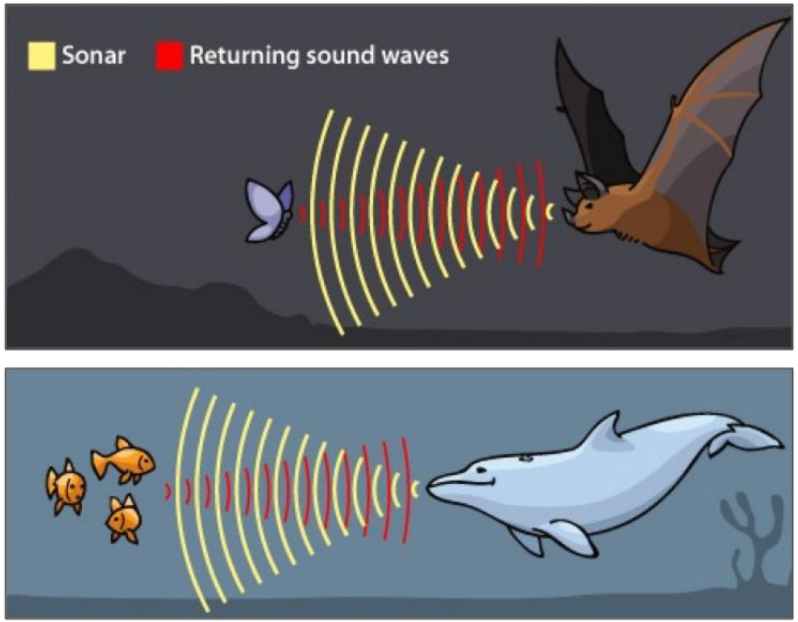
$$\varphi = \arctan\left(\frac{g_1 - g_3}{g_0 - g_2}\right)$$

$$d = \frac{c \cdot \varphi}{4\pi \cdot f}$$



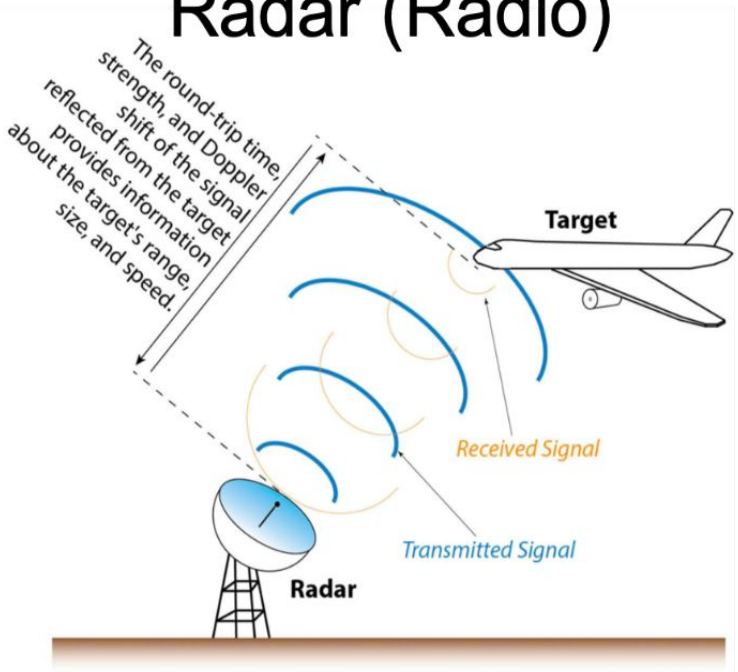
Otros tipos de ToF

Sonar (Sound)



Ultrasonic waves: 1.5-15 km
Frequency: 20kHz-200kHz.

Radar (Radio)



Radio waves: 1mm -100 km;
Frequency: 300 GHz-3 kHz.

4. Hands-on:
Active Depth Imaging