

Universidad de Alcalá

Escuela Politécnica Superior

GRADO EN INGENIERIA ELECTRONICA DE COMUNICACIONES



Trabajo Fin de Grado

Estudio de las diferentes alternativas tecnológicas utilizadas en cámaras de tiempo de vuelo (ToF)

ESCUELA POLITECNICA

Autor: Francisco Javier Martín Ramos

Tutor/es: <Manuel R. Mazo Quintas

2019

UNIVERSIDAD DE ALCALÁ
Escuela Politécnica Superior

**GRADO EN INGENIERÍA
ELECTRÓNICA DE COMUNICACIONES**

Trabajo Fin de Grado

**Estudio de las diferentes alternativas
tecnológicas utilizadas en cámaras de tiempo
de vuelo (ToF)**

Autor: Francisco Javier Martin Ramos

Tutor/es: Manuel R. Mazo Quintas

TRIBUNAL:

Presidente: Álvaro Hernandez Alonso

Vocal 1º: Ignacio Fernandez Lorenzo

Vocal 2º: Manuel Mazo Quintas

FECHA: 06 septiembre 2019

Agradecimientos

*“Nuestra recompensa se encuentra en el esfuerzo y no en el resultado.
Un esfuerzo total es una victoria completa”*

Mahatma Gandhi

Después de mucho esfuerzo y años de dedicación por fin cierro una de las etapas más importantes de mi vida, en la que, gracias al apoyo incondicional de mis padres, Feli y Ciriaco, así como de mi abuela Sole y mi hermano Miguel Angel, que nunca se rindieron y lucharon cada día.

También quería dedicarle unas palabras a mi pareja Eva, ya que desde el primer momento confió en mí y apoyo todas las decisiones que he ido tomando durante estos años.

Agradecer también el esfuerzo y la ayuda recibida por parte de mi tutor Manuel Mazo, quien siempre ha atendido todas mis cuestiones rápidamente.

Por último, me gustaría dedicar este trabajo a Alejandro, quien siempre confió en mí.

Resumen

El objetivo de este proyecto es el estudio de las diferentes alternativas tecnológicas utilizadas para el diseño de cámaras de tiempo de vuelo (ToF), incluyendo sus principales características, ventajas e inconvenientes y las fuentes de error.

Con más detalle, los objetivos a alcanzar son:

- Principio de funcionamiento, el cual se basa en un análisis en profundidad de todos los elementos que integran estas cámaras.
- Diferentes alternativas tecnológicas basadas en los diferentes métodos de modulación:
 - Cámaras de modulación continua
 - Cámaras de modulación pulsada
 - Cámaras de modulación de pseudo-ruido
- Ventajas e inconvenientes de cada una de ellas.
- Fuentes de error y soluciones para mitigar su incidencia en la obtención de mapas de profundidad.

Abstract

The aim of this project is to study the different technological alternatives used for the design of time-of-flight (ToF) cameras, including their main characteristics, advantages and disadvantages and error sources.

In detail, the objectives to reach, are:

- Works principle, which is based on an in-depth analysis of all the elements that make up these cameras.
- Different technological alternatives based on different modulation methods:
 - Continuous modulation cameras
 - Pulsed modulation cameras
 - Pseudo-noise modulation cameras
- Advantages and disadvantages of each one of them.
- Sources of error and solutions to mitigate their incidence in obtaining depth maps.

Índice general

Agradecimientos	4
Resumen	6
Abstract.....	8
Índice general	10
Índice de Figuras	12
Lista de acrónimos.....	15
1. INTRODUCCION	17
1.1. Sensores basados en Visión estéreo	19
1.2. Sensores basados en la proyección de un patrón de luz estructurada	24
1.3. Cámaras de profundidad (Time of Flight – ToF)	27
2. ALTERNATIVAS TECNOLOGICAS Y CARACTERÍSTICAS DE CADA UNA DE ELLAS	40
2.1. Modulación Continua	40
2.2. Modulación pulsada.....	44
2.3. Modulación PN (Pseudo-Noise).....	48
3. FUENTES DE ERROR: ERRORES SISTEMÁTICOS Y ALEATORIOS	54
3.1. Aspectos generales.....	54
3.2. Errores sistemáticos	55
3.2.1. Error circular	55
3.2.2. Errores relacionados con la amplitud	55
3.2.3. Errores relacionados con la temperatura	56
3.2.4. Ruido de patrón fijo (FPN – Fixed Pattern Noise)	56
3.3. Errores aleatorios	57
3.3.1. Dispersión de la luz dentro de la lente y del sensor.....	57
3.3.2. Interferencias por multicamino	57
3.3.3. Movimiento.....	57
3.3.4. Ambigüedad de profundidad	58
4. CONCLUSIONES	60
Bibliografía.....	62

Índice de Figuras

Figura 1. Experimento de Galileo (1).....	17
Figura 2 Visión Estéreo.....	19
Figura 3 Geometría epipolar (4).....	20
Figura 4 Cámara Intel® RealSense™ serie ZR300	20
Figura 5 Localización de Componentes Intel ZR300	21
Figura 6 Cámara Intel® RealSense™ serie D400.....	21
Figura 7 Localización de componentes Intel D400 series.....	22
Figura 8 Cámara Intel® RealSense™ SR300	22
Figura 9 Localización de componentes Intel SR300	23
Figura 10 Proyección de luz estructurada.....	24
Figura 11 Proceso de escaneado 3D. Patrón de luz emitido sobre el busto (7).....	24
Figura 12 Astra Series.....	25
Figura 13 Localización de componentes Orbbec Astra.....	25
Figura 14 PMD Sensor-Chips based on Photogate Technology (10).....	27
Figura 15 PMD Sensor-Chips based on Photogate Technology_2 (10).....	28
Figura 16 Diagrama de bloques de los sistemas 3D TOF basados en una matriz de detección PMD	28
Figura 17 Detección de objetos (12)	29
Figura 18 Smart pixels (12)	30
Figura 19 Primera mitad del primer periodo de la señal emitida Umt tras td (tiempo necesario para recibir la señal). Diferencia de fase de 0° entre la señal emitida y la recibida, lo que acumula carga en el pozo Qa excitando la señal Ca (12).....	30
Figura 20 Segunda mitad del primer periodo de la señal emitida $Um(t)$. Diferencia de fase de 180° entre la señal emitida y la recibida, lo que acumula carga en el pozo Qb excitando la señal Cb (12)	31
Figura 21 Segunda mitad del primer periodo de la señal Umt . $Umt = 0$ lo que provoca que los electrones se repartan simétricamente hacia las puertas de lectura a y b (12).....	31
Figura 22 Segundo periodo de la señal emitida Umt . Se repite el proceso detallado del primer periodo. Se puede observar como a medida que aumenta el número del periodo de la señal Umt , la carga acumulada dentro del pozo Qa es mayor que la del pozo Qb (12).....	32
Figura 23 Conmutador asociado a cada píxel (12)	33
Figura 24 Proceso de correlación entre el eco óptico recibido $Popt$ y la tensión de um para un solo período de modulación (13).....	33
Figura 25 Desplazamientos de fase	35
Figura 26 Principio de medida de una cámara de tiempo de vuelo (16).....	36
Figura 27 Distribución de medidas de distancia (12)	37
Figura 28 Diagrama de una cámara ToF (12).....	38
Figura 29 Esta figura muestra el principio general del método de cuatro cubos que estima la señal óptica demodulada en cuatro muestras igualmente espaciadas en un período de modulación. Un circuito CCD/CMOS logra la detección de luz, demodulación y almacenamiento de carga. A partir de estos cuatro valores, es posible estimar la fase y amplitud de la señal recibida, así como la cantidad de luz de fondo (offset) (19).....	42
Figura 30 Pulso de luz emitida. Modulación pulsada.....	44
Figura 31 Pulso de luz recibido en función del pulso de luz emitido. Modulación pulsada	45
Figura 32 Cargas eléctricas acumuladas en el sensor. Modulación pulsada	45
Figura 33 Tercera ventana de obturación. Minimizar iluminación ambiental. Modulación pulsada	47
Figura 34 Secuencia PN de longitud máxima $n=7$ chips (14).....	49
Figura 35 Función de autocorrelación periódica entre señales ópticas moduladas por una secuencia PN unipolar y cuya señal de control de interruptor es de naturaleza bipolar. La duración entre dos picos es dada por el producto entre la secuencia PN de longitud n y la duración de un chip	

Tc. La forma triangular del pico permite medir la distancia mediante la adquisición de dos muestras C_0 y C_1 en $t_0 = -\tau$ y $t_0 = -\tau + Tc$ respectivamente. (14)	49
Figura 36 Error circular, medido con la cámara SR300. Los puntos azules muestran las medidas, y la línea roja es un modelo polinómico ajustado al conjunto de datos para su corrección.	55
Figura 37 Error relacionado con la amplitud debido a la luz no homogénea. El objeto es una pared plana a 0,43m.	56
Figura 38 Reflexiones multitrayecto en una esquina. Debido a la interferencia de las ondas de luz, la esquina	57

Lista de acrónimos

TOF	Time of Flight
CWM	Continuous wave modulation
PM	Pulse modulation
PN	Pseudo-Noise
RGB	Red-Green-Blue
CMOS	Complementary Metal-Oxide-Semiconductor
IR	Infrared Rays
PMD	Photonic Mixer Device
A/D	Analógico – Digital
SNR	Signal to Noise Ratio
LED	Light Emitting Diode
FPN	Fixed Pattern Noise

1. INTRODUCCION

En la naturaleza, se pueden encontrar ejemplos de medida de tiempo de vuelo, tales como, los delfines y murciélagos, que utilizan esta alternativa tanto para la navegación como para la caza. Además, los humanos también han ideado soluciones a la medida del tiempo de vuelo (ToF) desde hace mucho tiempo, por ejemplo, dejando caer rocas por agujeros profundos y esperando el eco o tratando de estimar la distancia de la descarga de un rayo evaluando el tiempo entre el relámpago y el trueno asociado. Estos métodos están basados en el tiempo de propagación del sonido.

El verdadero pionero en la medida óptica del tiempo de vuelo fue Galileo Galilei [\(1\)](#), quien fue probablemente el primero en imaginar la luz como una especie de evento activo, algo que viaja a través del espacio libre en lugar de existir instantáneamente como trató de demostrar en uno de sus experimentos, donde en la oscuridad de la noche puso a dos sujetos (A y B) en dos colinas vecinas, separadas un kilómetro, dotando de una linterna a cada sujeto, tal como se muestra en la figura 1.

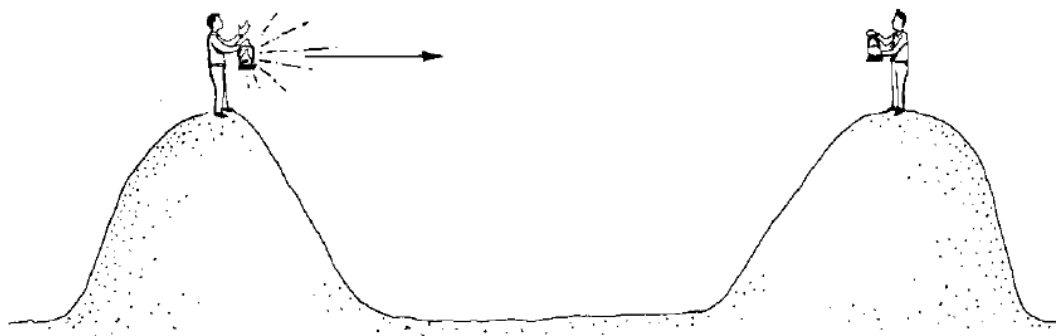


Figura 1. Experimento de Galileo [\(1\)](#)

El experimento comenzó con el sujeto A mostrando su linterna y esperando a que el sujeto B mostrara la suya al ver la luz del sujeto A. De esa manera Galilei esperaba medir el tiempo de propagación que la luz necesitaba para viajar de la persona A a la persona B y viceversa. Su conclusión fue que la velocidad medida era de la misma magnitud que los tiempos de reacción de los sujetos.

Años más tarde, a mediados del siglo XIX, Fizeau [\(1\)](#) diseñó un nuevo experimento basado en el uso de una rueda dentada giratoria como obturador mecánico para un haz de luz enfocado obteniendo un resultado asombrosamente preciso de 10.51 ppm.

En las últimas décadas se han desarrollado varias tecnologías para la obtención de imágenes de profundidad, algunas de las cuales se basan en la triangulación y otras en los tiempos de propagación de la radiación electromagnética. Las imágenes 3D en tiempo real se han vuelto muy populares en los últimos años. La información sobre la tercera dimensión es muy útil para simplificar o incluso permitir muchas tareas en los ámbitos de la navegación autónoma, la vigilancia y la seguridad. La importancia de la visión tridimensional en la naturaleza se evidencia en las numerosas especies de animales con dos o más ojos, capaces de percibir el mundo en tres dimensiones a través de la visión estereoscópica.

Aprendiendo de la naturaleza, las soluciones técnicas pueden beneficiarse enormemente de la incorporación de potentes herramientas de percepción 3D en tiempo real. La disponibilidad de información 3D en tiempo real permite la realización de nuevas aplicaciones en los campos de sistemas de seguridad, medidas biométricas, visión artificial, interfaz máquina-usuario, robótica y juegos.

Las alternativas tecnológicas más importantes para la obtención de información de profundidad de una escena son:

1. Sensores basados en **visión estéreo**
2. Sensores basados en la proyección de un patrón de **luz estructurada** (generalmente luz infrarroja)
3. Cámaras de profundidad (**Time of Flight - ToF**)

Estos métodos a su vez pueden ser clasificados como:

- **Activos:** Se caracterizan porque hay que realizar la proyección de un haz de luz sobre la escena para obtener información de profundidad.
 - Time of Flight
 - Luz estructurada
- **Pasivos:** Se basan en buscar correspondencias entre un par de imágenes, de las cuales a priori no se conoce nada. Si bien es verdad que en algunos casos también se realiza la proyección de puntos de luz para facilitar la detección de correspondencias.
 - Visión estéreo

En todos los casos la obtención de profundidad se realiza a partir de la energía luminosa reflejada. En aplicaciones de detección de personas, el uso de cámaras como sensor presenta ventajas al ser transparente al usuario y la precisión alcanzada en la posición es del orden de centímetros, sin embargo, estas soluciones basadas en cámaras presentan algunos inconvenientes, como son el tratamiento de una cantidad elevada de información, necesidades hardware elevadas y un uso “óptimo” de la información proporcionada para tratamientos en tiempo real.

1.1. Sensores basados en Visión estéreo

Este tipo de sensores obtienen la información de profundidad a partir de la correspondencia de puntos de la escena de dos cámaras RGB o de nivel de gris situadas físicamente en un único soporte, separadas entre sí una cierta distancia con la que se obtienen las correspondientes imágenes de par estéreo imitando así la visión estereoscópica de los humanos.

Al igual que ocurre con la visión humana, el procedimiento de esta tecnología se basa en la captura de dos imágenes de una misma escena desde una posición de las cámaras ligeramente desplazadas entre sí de forma que solo exista un desplazamiento horizontal entre ellas (véase figura 2). Este es el principio básico que permite la obtención de distancia a la que se encuentra un objeto dentro de una escena. El principal problema de este tipo de tecnología es encontrar la correspondencia (*matching*) de un mismo punto en el espacio desde las imágenes obtenidas de cada cámara del par estéreo. Para solventar este problema, se hace uso de un tipo de geometría, denominada geometría epipolar. Como consecuencia de la geometría del sistema, se obtiene la denominada restricción epipolar, la cual determina que, para todo punto de una imagen de un par estereoscópico, existe una línea epipolar en la otra imagen sobre la que está su homólogo, lo que limita la búsqueda de correspondencia de manera que, en vez de buscar el punto objetivo por cada imagen del par, buscamos las líneas epipolares correspondientes [_2](#)) (véase figura 3). Además, otro de los métodos utilizados para facilitar el “*matching*” en situaciones de baja textura, como puede ser una pared blanca y plana, es el uso de la emisión de patrones de luz generando textura a la escena. [_3](#))

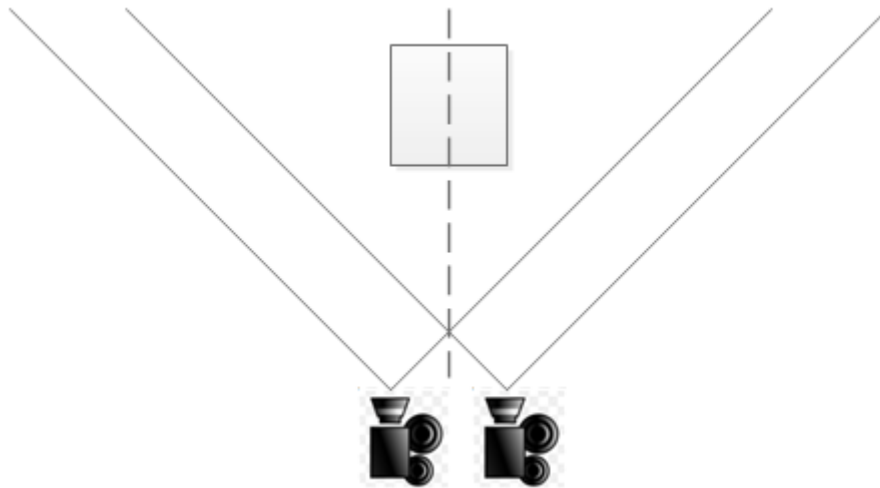


Figura 2 Visión Estéreo

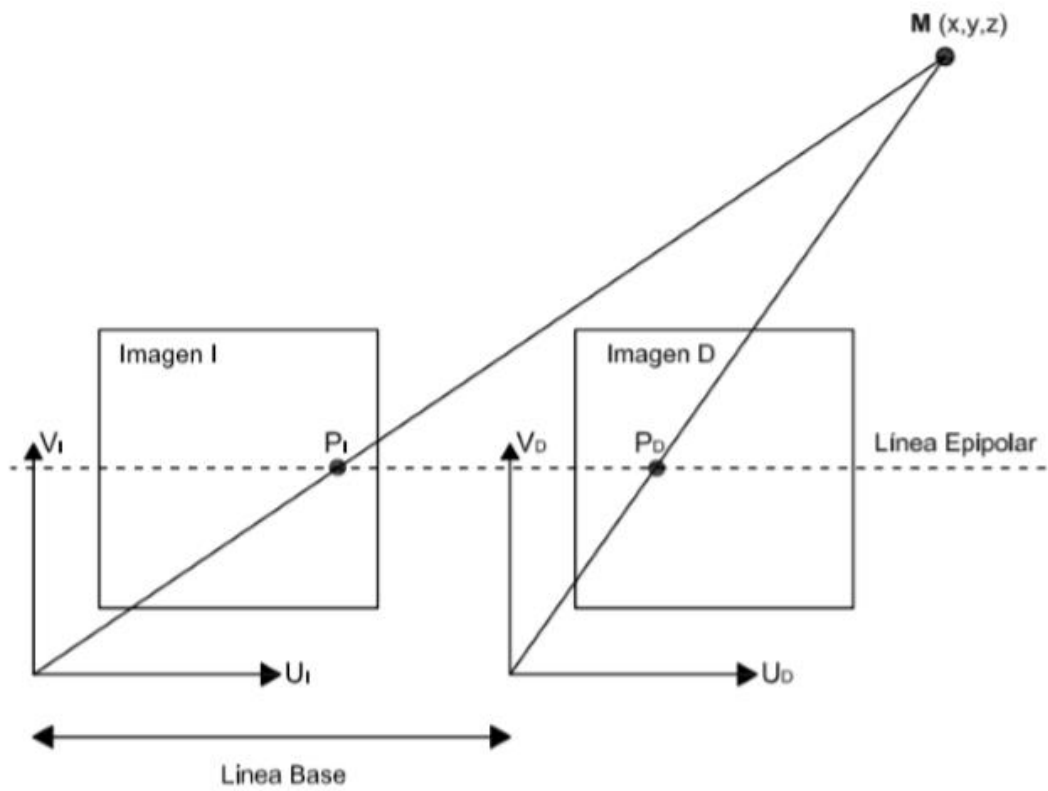


Figura 3 Geometría epipolar_(4)

Existen varios tipos de soluciones comerciales que se basan en este principio (5):

- ZR300



Figura 4 Cámara Intel® RealSense™ serie ZR300

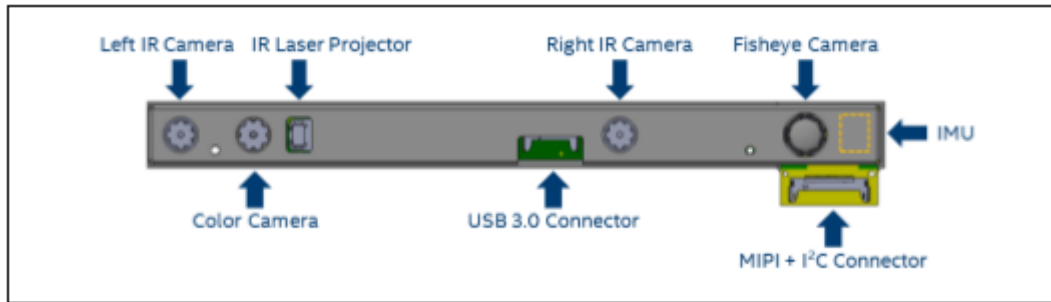


Figura 5 Localización de Componentes Intel ZR300

Intel ZR300	CÁMARAS			PROFUNDIDAD	
	Tipo	Resolución	fps	Rango	Campo de visión (máx.)
	Profundidad	628x468 480x360 320x240	30, 60	0,55m a 2,80m	73,5° (D) x 48,3° (V) x 62° (H)
RGB	1920x1080 640x480 320x240	30 15, 30, 60 30, 60			
Fisheye	640x480	5, 30, 60			

- D400 Series



Figura 6 Cámara Intel® RealSense™ serie D400

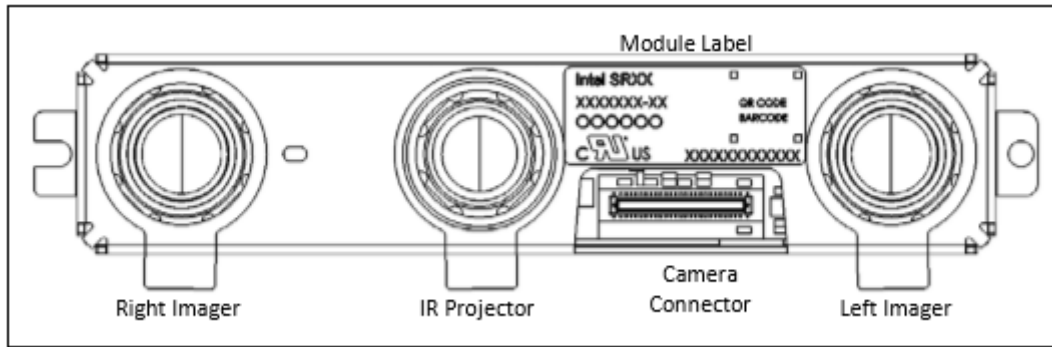


Figura 7 Localización de componentes Intel D400 series

Intel D415	CÁMARAS			PROFUNDIDAD	
	Tipo	Resolución	fps	Rango	Campo de visión (máx.)
	Profundidad	1280x720	90	0,16m a 10m	41,6° (V) x 65,3° (H)
RGB	1920x1080	30			

Intel D435	CÁMARAS			PROFUNDIDAD	
	Tipo	Resolución	fps	Rango	Campo de visión (máx.)
	Profundidad	1280x720	90	0,11m a 10m	59,7° (V) x 87,8° (H)
RGB	1920x1080	30			

- SR300



Figura 8 Cámara Intel® RealSense™ SR300

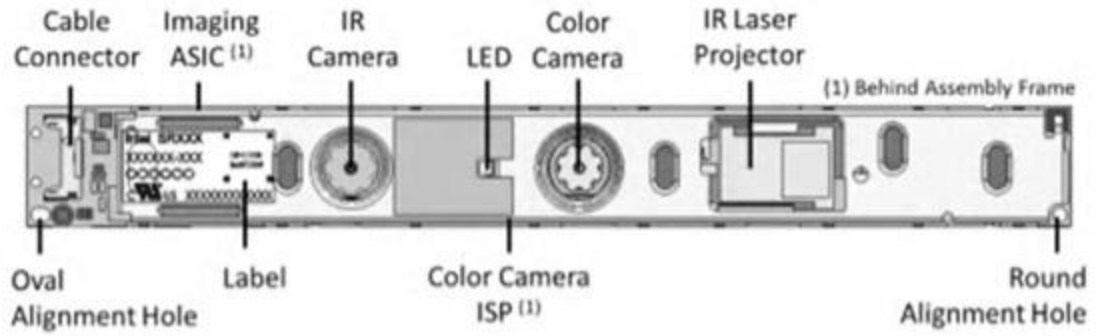


Figura 9 Localización de componentes Intel SR300

Intel SR300	CÁMARAS			PROFUNDIDAD	
	Tipo	Resolución	fps	Rango	Campo de visión (máx.)
	<i>Profundidad</i>	<i>640x480</i>	<i>60,30,10</i>	<i>0,2m a 1,5m</i>	<i>90,4° (D) x 56,1° (V) x 72,9° (H)</i>
<i>RGB</i>	<i>FHD</i>	<i>30, 10</i>			
	<i>HD</i>	<i>60,30,10</i>			
	<i>VGA</i>	<i>60,30,10</i>			
	<i>QVGA</i>	<i>60,30,10</i>			

1.2. Sensores basados en la proyección de un patrón de luz estructurada

El proceso de la obtención de profundidad se basa en estudiar la deformación que sufre un patrón de luz conocido (ocasionalmente rejillas o barras verticales como se puede observar en la figura 10) al ser interseccionado por cualquier objeto dentro de una escena. Estos sensores de profundidad están compuestos de un emisor/receptor (sensor CMOS) de luz infrarroja para la proyección del patrón y una cámara RGB que recoja todas las imágenes de la deformación del plano laser evitando que existan zonas oscuras. (6)

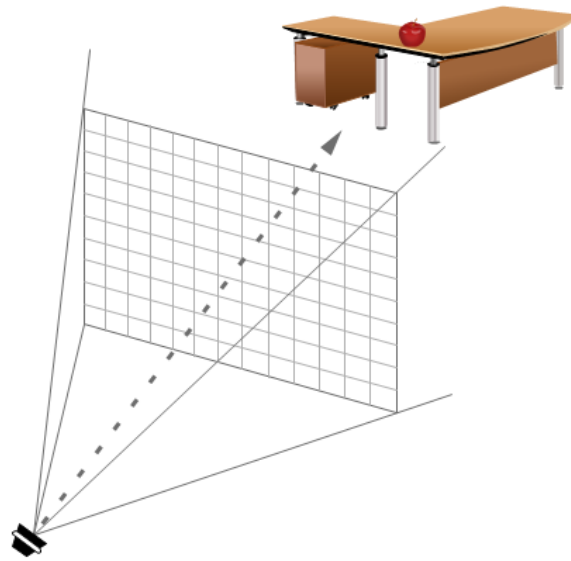


Figura 10 Proyección de luz estructurada



Figura 11 Proceso de escaneo 3D. Patrón de luz emitido sobre el busto (7)

La correspondencia entre la imagen captada y la imagen que dicho patrón forma al impactar sobre la escena (véase figura 11), permite triangular la posición de cada píxel y determinar su profundidad con respecto al plano perpendicular de la cámara.

Dependiendo del tipo de patrón y del análisis posterior realizado existen diferentes técnicas que definen las diferentes soluciones comerciales. Un ejemplo es el que se observa en la figura 12_(8)



Figura 12 Astra Series



Figura 13 Localización de componentes Orbbec Astra

Orbbec Astra PRO	CÁMARAS			PROFUNDIDAD	
	Tipo	Resolución	fps	Rango	Campo de visión (máx.)
	Profundidad	640x480	30	0,6m a 8m	73° (D) x 49,5° (V) x 60° (H)
RGB	1280x720	30			

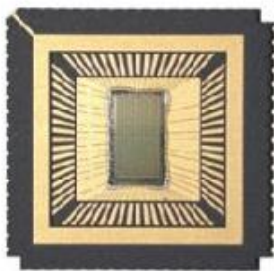
Orbbec Astra	CÁMARAS			PROFUNDIDAD	
	Tipo	Resolución	fps	Rango	Campo de visión (máx.)
	Profundidad	640x480	30	0,6m a 8m	73° (D) x 49,5° (V) x 60° (H)
RGB	640x480	30			

Orbbec Astra S	CÁMARAS			PROFUNDIDAD	
	Tipo	Resolución	fps	Rango	Campo de visión (máx.)
	<i>Profundidad</i>	<i>640x480</i>	<i>30</i>	<i>0,4m a 2m</i>	<i>73° (D) x 49,5° (V) x 60° (H)</i>
<i>RGB</i>	<i>640x480</i>	<i>30</i>			

1.3. Cámaras de profundidad (Time of Flight – ToF)

Las cámaras de tiempo de vuelo (Time of Flight – cameras, ToF cameras) son sensores que proporcionan por cada píxel la distancia entre la cámara y un punto de la escena 3D. Como consecuencia, la cámara genera una imagen de distancias de la escena. Su funcionamiento se basa en la emisión de una fuente de luz modulada, normalmente en el espectro infrarrojo (IR), que ilumina la escena. La obtención de la distancia en las cámaras ToF se consigue midiendo el tiempo de vuelo de la señal que viaja desde la fuente de luz, colisiona en la escena y retorna al elemento sensor (en la práctica el elemento sensor y la fuente de luz están muy próximos).

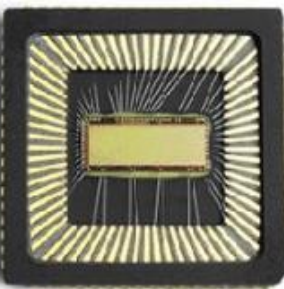
Uno de los elementos principales de las cámaras de tiempo de vuelo (ToF) para la adquisición de las imágenes de profundidad son los dispositivos mezcladores fotónicos o PMD de sus siglas en inglés Photonic Mixer Device. Estos tipos de sensores permiten captar a la vez la intensidad y la distancia de cada píxel (9). Algunos ejemplos de estos chips son los mostrados en la figura 14 y figura 15.



PhotonICs® PMD 768-S

3D Video Sensor Array with Active SBI

- Array format 24(H) x 32(V)
- Optical format 1/2", 16:9 image ratio
- Effective image area 6.75mm (H) x 3.73mm (V)
- Package Standard CLCC68
- Pixel size 210.8 μ m (H) x 155.3 μ m (V)
- Fill factor (FF) 42%
- 78dB dynamic range for active illumination
- 50dB dynamic range increase by SBI circuitry

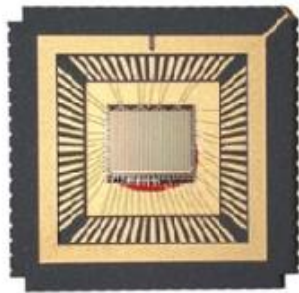


PhotonICs® PMD 1k-S 2

3D Video Sensor Array with Active SBI

- Array format 64(H) x 16(V)
- Optical format 2/3", 3:1 image ratio
- Effective image area 9.94mm (H) x 3.37mm (V)
- Package Standard CLCC68
- Pixel size 155.3 μ m (H) x 210.8 μ m (V)
- Fill factor (FF) 42%
- 78dB dynamic range for active illumination
- 50dB dynamic range increase by SBI circuitry

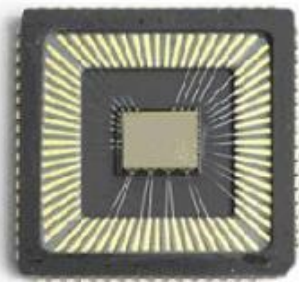
Figura 14 PMD Sensor-Chips based on Photogate Technology (10)



PhotonICs® PMD 3k-S

3D Video Sensor Array with Active SBI

- Array format 64(H) x 48(V)
- Optical format 1/2", 4:3 image ratio
- Effective image area 6.4mm (H) x 4.8mm (V)
- Package Standard CLCC68
- Pixel size 100µm (H) x 100µm (V)
- Fill factor (FF) 32%
- 74dB dynamic range for active illumination
- 50dB dynamic range increase by SBI circuitry



PhotonICs® PMD 19k

High Resolution 3D Video Sensor Array

- Array format 160(H) x 120(V)
- Optical format 1/2", 4:3 image ratio
- Effective image area 6.4mm (H) x 4.8mm (V)
- Package Standard CLCC68
- Pixel size 40µm (H) x 40µm (V)
- Fill factor (FF) 31%
- 65dB dynamic range for active illumination

Figura 15 PMD Sensor-Chips based on Photogate Technology_2 (10)

En la Figura 16 se representa mediante un diagrama de bloques, los elementos que componen una cámara ToF.

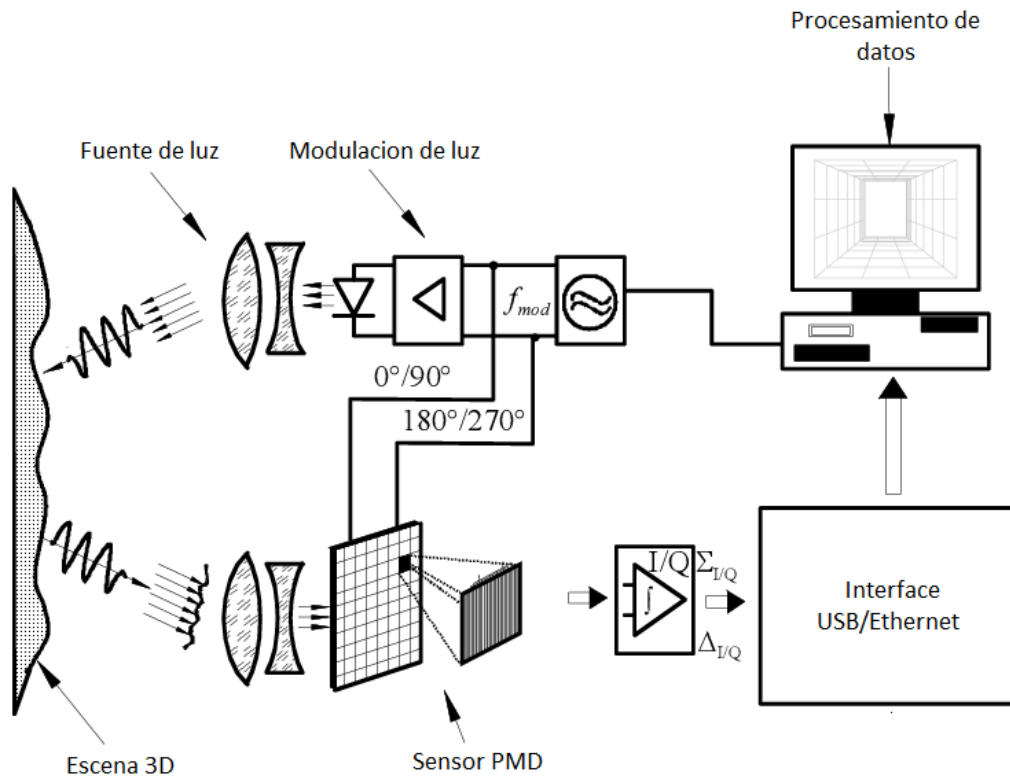


Figura 16 Diagrama de bloques de los sistemas 3D TOF basados en una matriz de detección PMD

Este sistema se compone fundamentalmente de un chip PMD compuesto de una matriz de píxeles que proporcionan la información de profundidad, una fuente de luz modulada que ilumina la escena y algunos componentes externos como convertidores A/D o electrónica del controlador que son integrados en el chip para simplificar el sistema. Como se muestra en la Figura 17, una señal óptica $P(t)$ es emitida hacia la escena y detectada por el sensor tras un tiempo t_d obteniendo en el sensor la señal $P(t - t_d)$. Esta señal detectada por el sensor contiene la información de amplitud y fase que permite el cálculo de distancias mediante la correlación entre la señal reflejada y la señal emitida. Este sensor de imagen está diseñado para que, al recibir la luz de la onda reflejada, esta energía fotónica sea convertida en corriente eléctrica. [\(11\)](#)

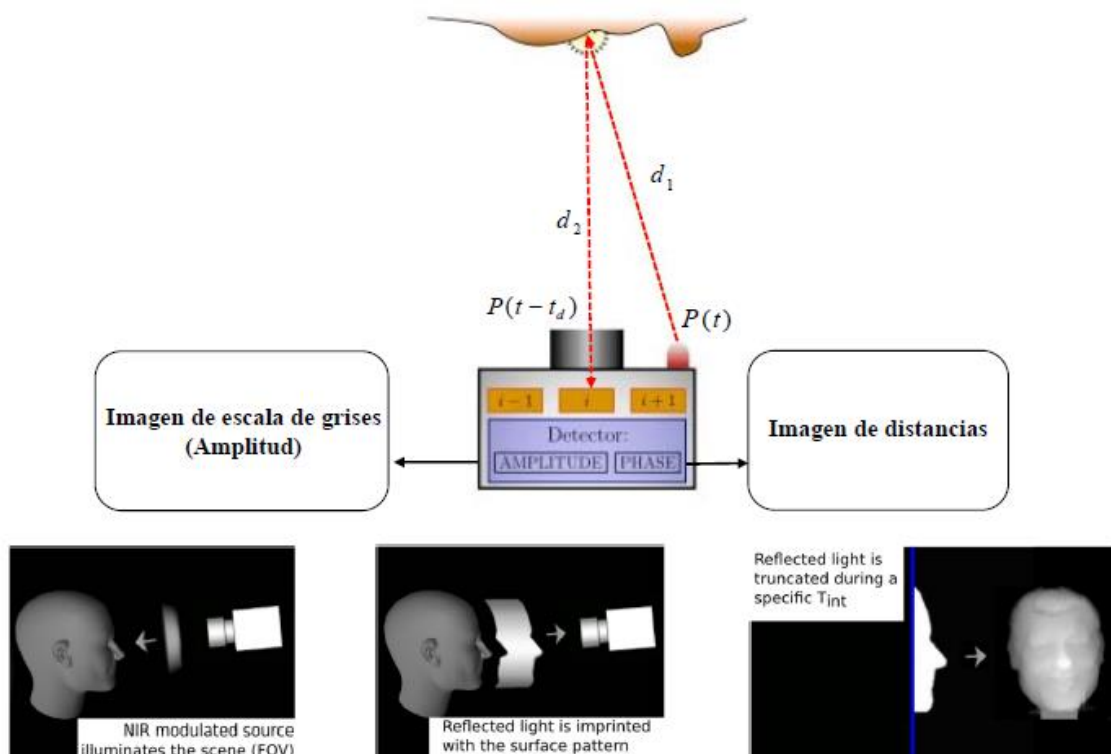


Figura 17 Detección de objetos [\(12\)](#)

Como el proceso completo de conversión de la señal óptica tiene lugar dentro de cada píxel, llamamos a los elementos PMD "Smart pixels". Estos elementos son los encargados de hacer la correlación entre la señal emitida y la señal recibida para obtener la información de distancias. La figura 18 muestra la sección transversal de un píxel PMD compuesto por dos pozos simétricos. Los fotones del eco óptico recibido P_{opt} inciden generando portadores de carga (pares de electrones y huecos). Las puertas están aisladas mediante una capa de SiO_2 o Si_3N_4 [\(13\)](#). Las puertas fotoeléctricas están controladas por la tensión de modulación u_m y la tensión de offset U_0 . Si no se recibe luz modulada, las cargas generadas se desplazan simétricamente hacia las puertas de lectura a y b . Si se recibe luz modulada, las cargas generadas se desplazan hacia la puerta de lectura b cuando la luz modulada y la tensión de modulación tienen una diferencia de fase de 180° . Si la diferencia de fase es de 0° , las cargas generadas serán desplazadas hasta la lectura de la puerta a .

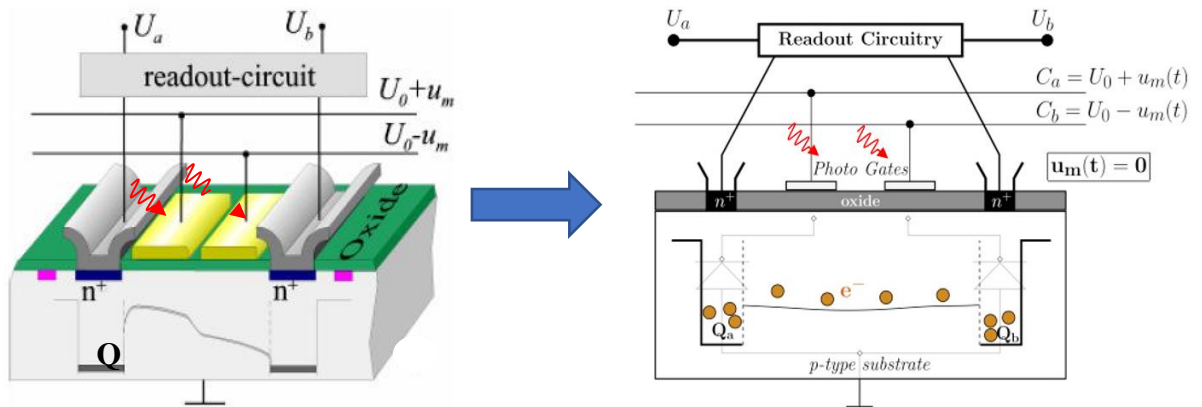


Figura 18 Smart pixels (12)

En la Figura 19 podemos observar como la señal emitida $U_m(t)$ (señal cuadrada) es mayor que 0 durante la primera mitad del periodo. Hasta que la señal no es recibida por el receptor tras el tiempo t_d las cargas generadas se desplazan simétricamente hacia las puertas de lectura a y b . Una vez la señal es recibida por el receptor mientras el pulso de la señal emitida sigue siendo mayor que 0, es decir, la diferencia de fase es 0° , los electrones se empiezan a almacenar en el pozo Q_a excitando la señal C_a , mientras Q_b se mantiene constante.

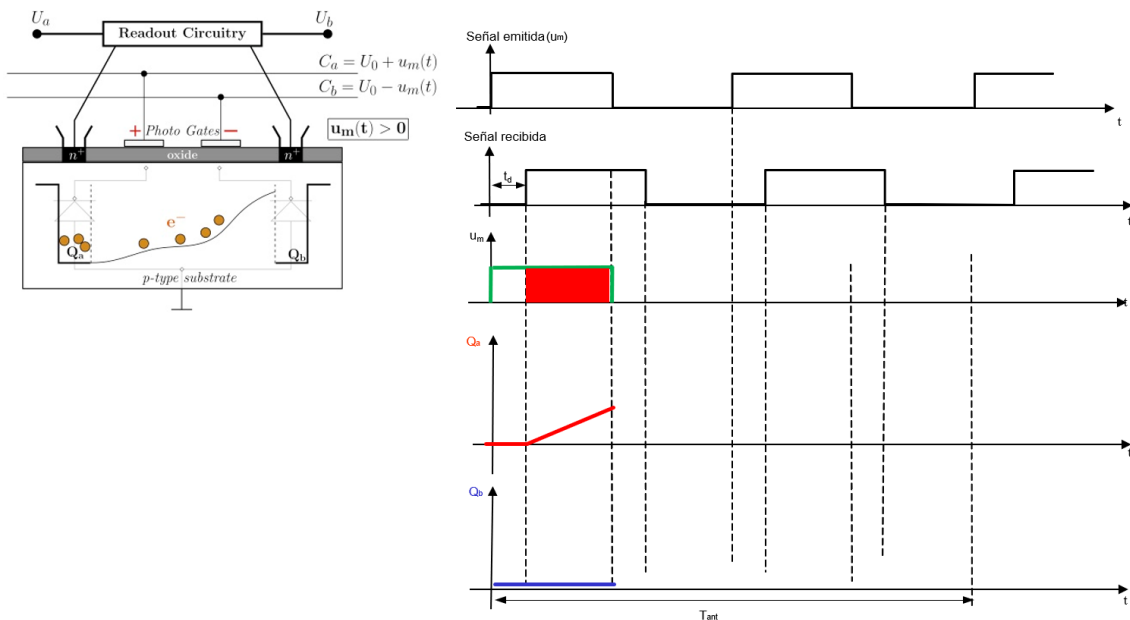


Figura 19 Primera mitad del primer periodo de la señal emitida $U_m(t)$ tras t_d (tiempo necesario para recibir la señal). Diferencia de fase de 0° entre la señal emitida y la recibida, lo que acumula carga en el pozo Q_a excitando la señal C_a (12)

Dentro del primer periodo, cuando la señal emitida $U_m(t)$ es igual a 0, la señal recibida se mantiene constante durante el tiempo t_d , produciendo una diferencia de fase de 180° entre la señal recibida y la señal emitida, lo que provoca que los electrones se almacenen dentro del pozo Q_b , excitando la señal C_b como se muestra en la Figura 20:

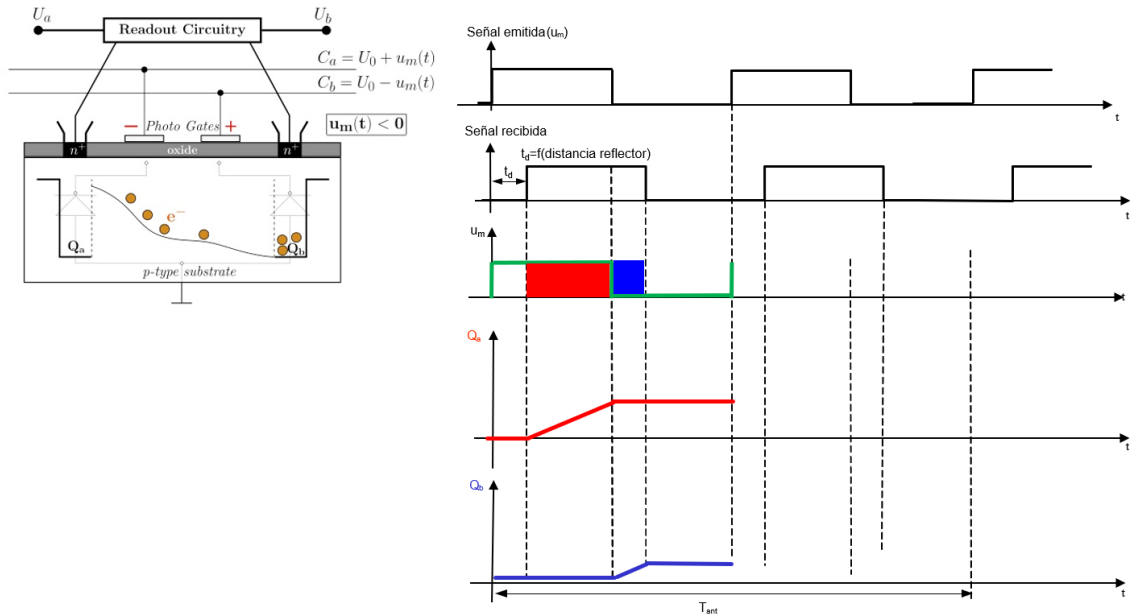


Figura 20 Segunda mitad del primer periodo de la señal emitida $U_m(t)$. Diferencia de fase de 180° entre la señal emitida y la recibida, lo que acumula carga en el pozo Q_b excitando la señal C_b (12)

Dentro de la segunda mitad del periodo de la señal emitida tras el tiempo t_d , no se recibe luz modulada ($U_m(t) = 0$), por lo que los electrones se desplazan simétricamente hacia las puertas de lectura a y b como se muestra en la figura 21:

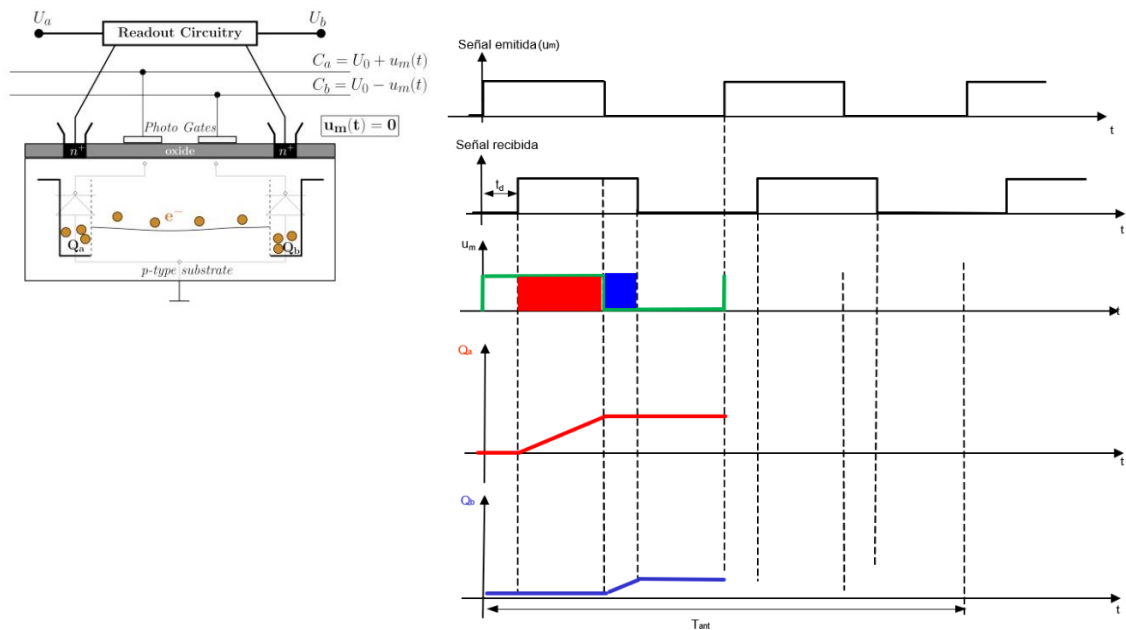


Figura 21 Segunda mitad del primer periodo de la señal $U_m(t)$. $U_m(t) = 0$ lo que provoca que los electrones se repartan simétricamente hacia las puertas de lectura a y b (12)

En la Figura 22 observamos un segundo periodo de la señal emitida donde se ve reflejado el mismo resultado.

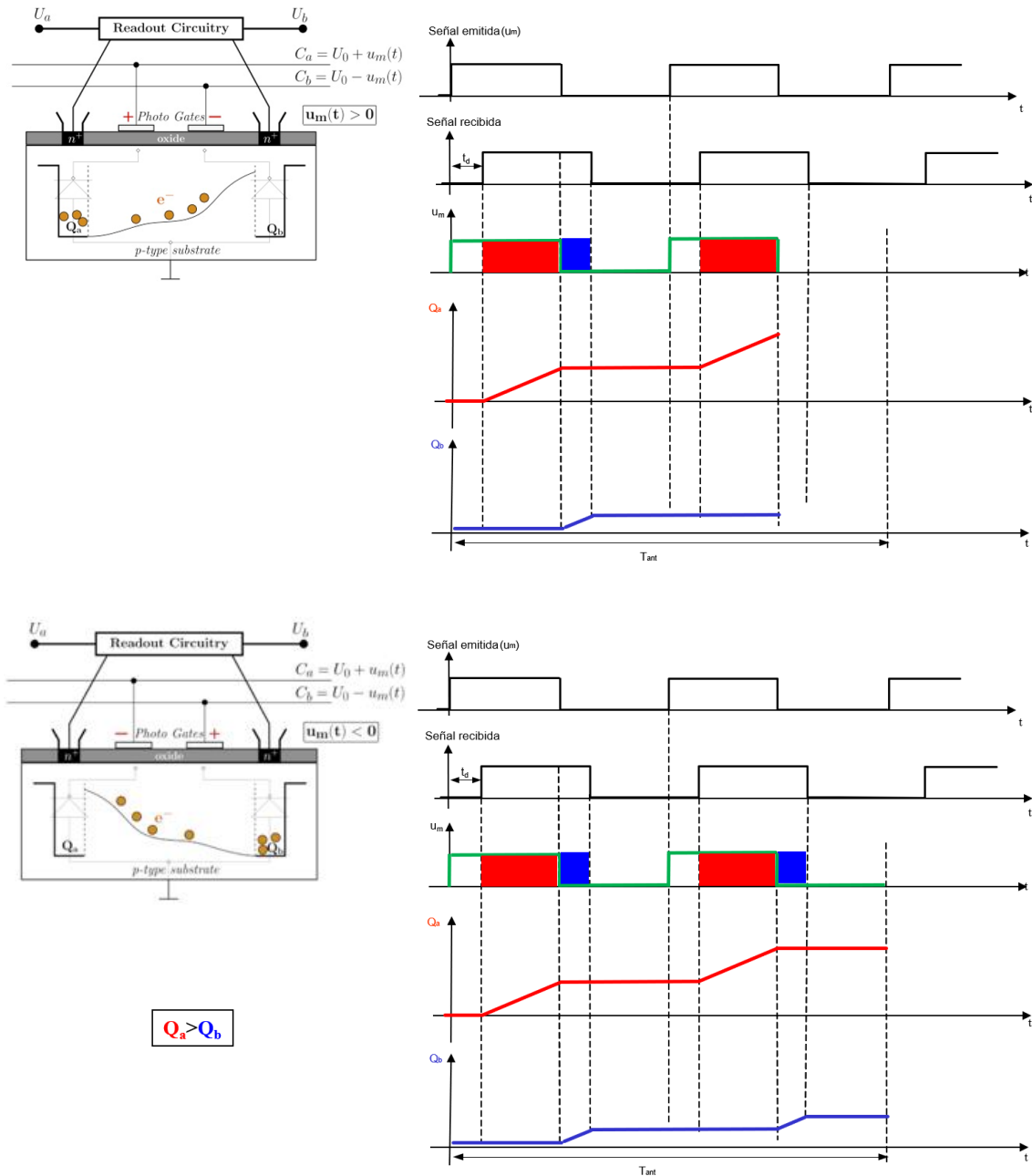


Figura 22 Segundo periodo de la señal emitida $U_m(t)$. Se repite el proceso detallado del primer periodo. Se puede observar como a medida que aumenta el número del periodo de la señal $U_m(t)$, la carga acumulada dentro del pozo Q_a es mayor que la del pozo Q_b (12)

El conmutador asociado a cada píxel distribuye la carga fotogenerada en dos capacidades de almacenamiento idénticas. Dependiendo de la distancia, la relación entre Q_a y Q_b cambia.

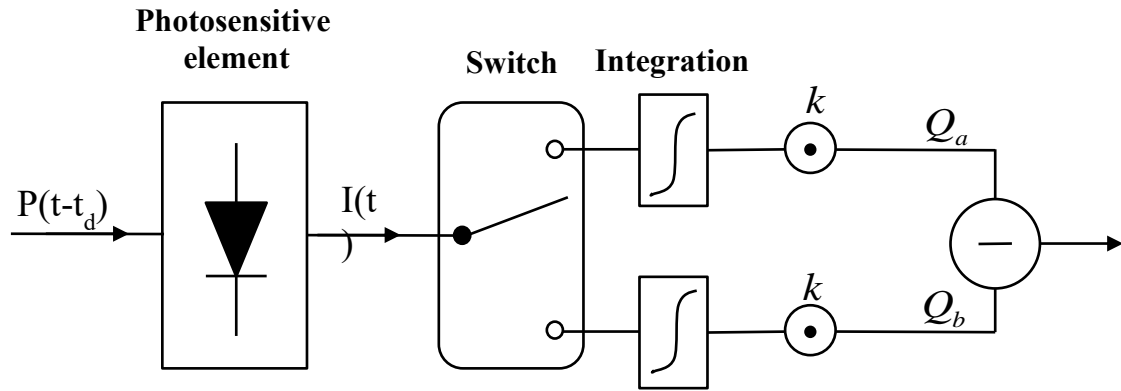


Figura 23 Conmutador asociado a cada píxel (12)

Un amplificador diferencial genera la diferencia entre los dos valores de integración (véase figura 23). Por consiguiente, la expresión matemática de la salida viene dada por:

$$C(-t_d) = Q_a - Q_b = K \int_{t_0}^{t_0+T} P(t - t_d) \cdot S(t) dt \quad 1.1$$

Donde Q_a, Q_b representan los valores de integración individuales. $P(t)$ describe las características temporales de la potencia óptica de la señal en un píxel, con un retardo de t_d y $S(t)$ es la secuencia binaria bipolar que controla el switch. (14)

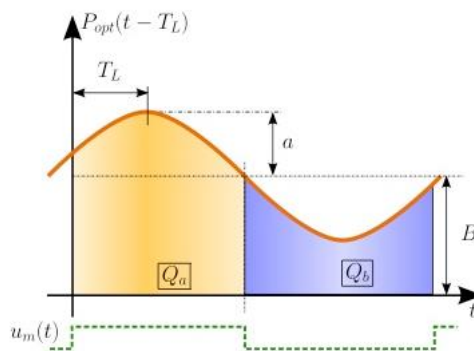


Figura 24 Proceso de correlación entre el eco óptico recibido P_{opt} y la tensión de u_m para un solo período de modulación (13)

Los pozos Q_a y Q_b están conectados a un condensador de integración. Por tanto, las tensiones correspondientes a U_a y U_b pueden expresarse como una función de correlación entre el eco óptico $(t - T_L)$ y la tensión de modulación $u_m(t)$ durante el tiempo de integración T_{int} . El tiempo de integración determina el periodo de tiempo durante el cual el sensor recoge la luz IR (15). En la práctica $T_{int} = nT$, con "n" en el orden de los miles (≈ 20.000). La figura 24 ilustra el proceso de correlación para un único periodo T de dos señales $P_{opt}(t - T_L)$ y $u_m(t)$. La tensión de modulación $u_m(t)$ y el eco óptico $P_{opt}(t - T_L)$ se definen como siguen (13):

$$u_m(t) = \begin{cases} 1, & \text{para } 0 \leq t - NT \leq \frac{T}{2} \\ 0, & \text{para } \frac{T}{2} < t - NT \leq T \end{cases} \quad N = 0,1,2 \dots \quad 1.2$$

$$P_{opt}(t - T_L) = a_0 \cos \omega(t - T_L) + B \quad 1.3$$

Donde T_L es el tiempo de vuelo (cámara – escena – cámara), B es la luz media recibida (luz de fondo y componente DC de la fuente de luz modulada) y a_0 es la amplitud de la luz modulada recibida. Por tanto U_a y U_b son proporcionales a las áreas Q_a y Q_b como se muestra en la figura 24. Si tenemos en cuenta el tiempo de integración T_{int} (que corresponde a varios cientos/miles de periodos) necesario para que los pozos tengan cargas suficientes, U_a y U_b se puede escribir como (13):

$$\begin{aligned} U_a(T_L) &= K \int_0^{T_{int}} P_{opt}(t - T_L) u_m(t) dt = K \frac{T_{int}}{T} \left[\frac{T a_0}{\pi} \sin(T_L) + \frac{T B}{2} \right] \\ &= K \cdot \frac{T_{int}}{T} Q_a(T_L) \end{aligned} \quad 1.4$$

$$\begin{aligned} U_b(T_L) &= K \int_0^{T_{int}} P_{opt}(t - T_L) u_m\left(t - \frac{T}{2}\right) dt = K \frac{T_{int}}{T} \left[-\frac{T a_0}{\pi} \sin(T_L) + \frac{T B}{2} \right] \\ &= K \cdot \frac{T_{int}}{T} Q_b(T_L) \end{aligned} \quad 1.5$$

La ganancia K representa la conversión de energía óptica recibida en tensión. Mirando la figura 23 se puede observar que U_a y U_b son siempre voltajes positivos. Para eliminar la influencia de la luz de fondo B, se debe determinar la diferencia de U_{ab} (13):

$$\Delta U_{ab}(T_L) = U_a - U_b = K \cdot \frac{T_{int}}{T} \cdot (Q_a(T_L) - Q_b(T_L)) \quad 1.6$$

La función de autocorrelación U_{ab} corresponde al valor de distancia de un píxel PMD. La suma de U_a y U_b corresponde a todos los fotones recibidos y convertidos. Por tanto, esta suma es equivalente al nivel de gris (amplitud) (13):

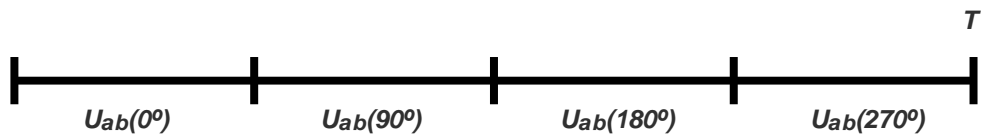
$$\sum U_{ab} = U_a + U_b = K \cdot \frac{T_{int}}{T} (Q_a + Q_b) = K \int_0^{T_{int}} P_{opt}(t - T_L) dt = B \quad 1.7$$

La amplitud del eco óptico recibido P_{opt} varía en función del coeficiente de reflexión del objeto (escena), la luz ambiental y la distancia. Esto implica por tanto que

la amplitud de la tensión de salida U_{ab} también se vea afectada por estos cambios. Para superar los problemas de la amplitud de la tensión de salida U_{ab} , la técnica a utilizar se basa en un algoritmo de desplazamiento de 4 fases (13):

$$\phi_0 = \arctan\left(\frac{\Delta U_{ab}(270^\circ) - \Delta U_{ab}(90^\circ)}{\Delta U_{ab}(0^\circ) - \Delta U_{ab}(180^\circ)}\right) \quad 1.8$$

Esta técnica empleará una ráfaga de luz durante un tiempo T , donde se distinguirán los cuatro desplazamientos de fase:



A partir de la figura 25, podemos ver el funcionamiento de la medida de diferencia de fase entre el rayo emitido, el rayo reflejado y la distancia d , las cuales se calcularán de la siguiente forma (16):

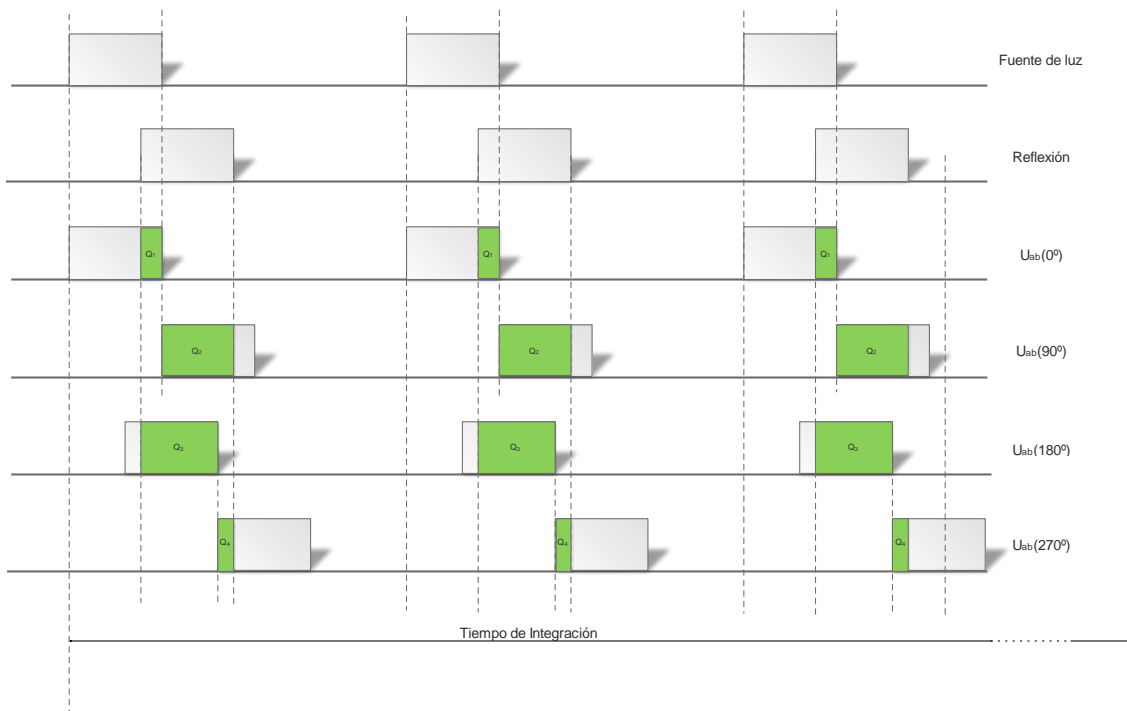


Figura 25 Desplazamientos de fase

$$\beta = \arctan\left(\frac{Q_2 - Q_3}{Q_0 - Q_1}\right) \quad 1.9$$

Donde $Q_0 \dots Q_3$ son las cuatro muestras mostradas en la figura 26.

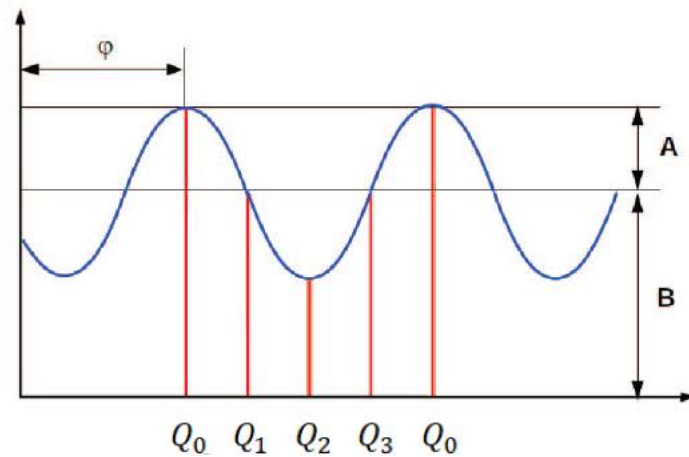


Figura 26 Principio de medida de una cámara de tiempo de vuelo (16)

Físicamente estas muestras miden la diferencia de carga de los pixeles vecinos (16).

$$d = \frac{c}{4\pi f_{mod}} \beta \quad 1.10$$

Siendo:

- $c = 3 \cdot 10^8 \text{ m/s} \Rightarrow$ Velocidad de la luz en el vacío
- $f_{mod} \Rightarrow$ frecuencia de modulación

La distancia máxima (d_{max}) viene definida por la frecuencia de modulación del emisor (f_{mod}) y la velocidad de la luz (c) y está limitada por el rango de no ambigüedad. Este rango viene determinado por que el retardo de fase máximo que puede ser medido entre la señal transmitida y la señal reflejada (tiempo de vuelo) está limitado a un ciclo completo del periodo de modulación, es decir, el ángulo máximo es 2π , por tanto, de la ecuación 1.3, podemos determinar que:

$$\omega(t - T_L) \leq 2\pi$$

Teniendo en cuenta que el tiempo de vuelo viene definido por:

$$t - T_L = \frac{2d}{c}$$

Entonces:

$$\omega \cdot \frac{2d}{c} = 2\pi$$

1.11

$$d = \frac{2\pi c}{2 \cdot 2\pi f}$$

$$d_{max} = \frac{c}{2f_{mod}}$$

Por ejemplo, para una $f_{mod} = 20MHz$ la distancia máxima sería de 7,5m. Si un objeto se encontrase a una distancia de 9m, la distancia obtenida sería de 1.5m.

La Figura 27 muestra la distribución de medidas en función de la distancia:

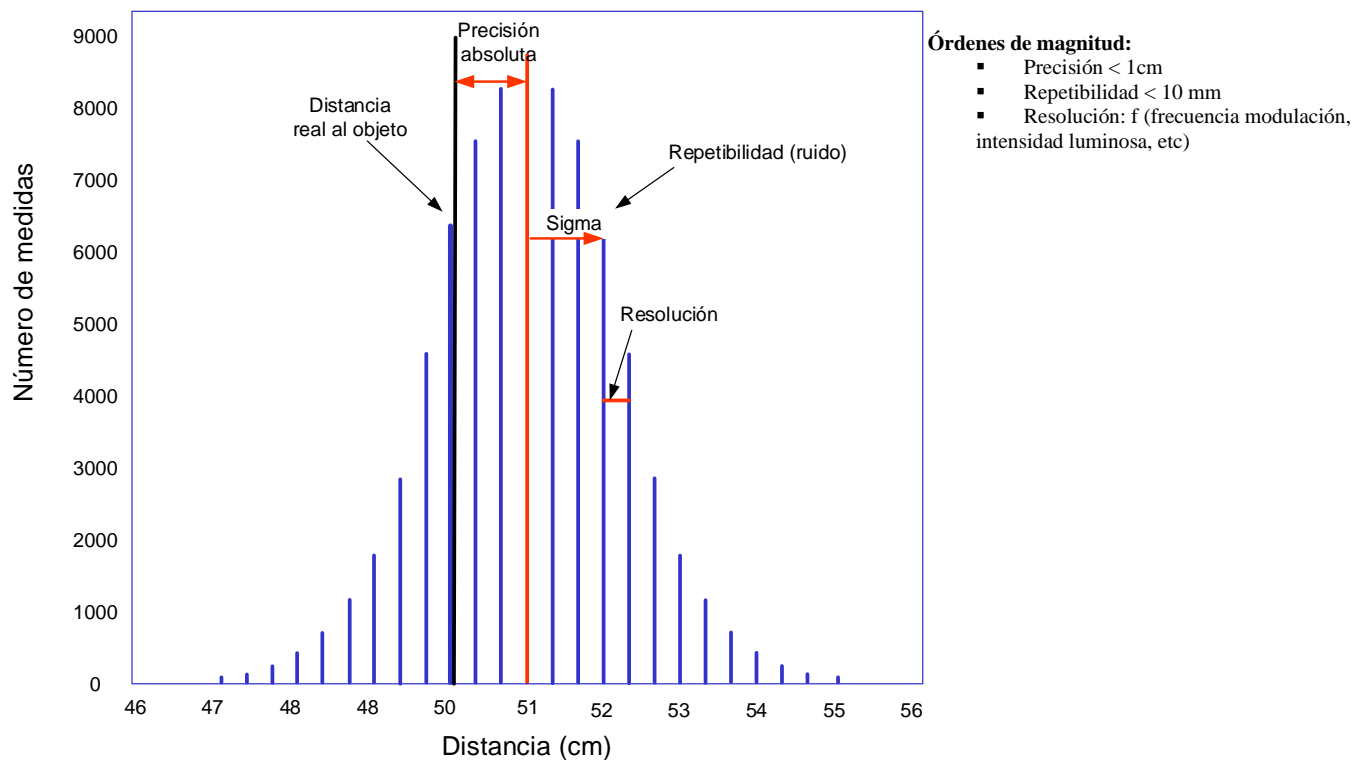


Figura 27 Distribución de medidas de distancia (12)

La precisión absoluta es la media de la diferencia entre la distancia medida y la distancia real. La repetibilidad indica la variación de los valores medidos alrededor del valor medio. Este valor indica la cantidad de ruido existente en una medición y depende de la reflectividad de los objetos existentes en la escena.

En la figura 28 podemos observar un diagrama completo de una cámara ToF, y se puede distinguir en función de la señal moduladora las diferentes alternativas tecnológicas para la obtención de las distancias: modulación de onda continua (CWM), modulación pulsada (PM) y modulación por secuencia pseudo aleatoria o de pseudo ruido (pseudonoise – PN).

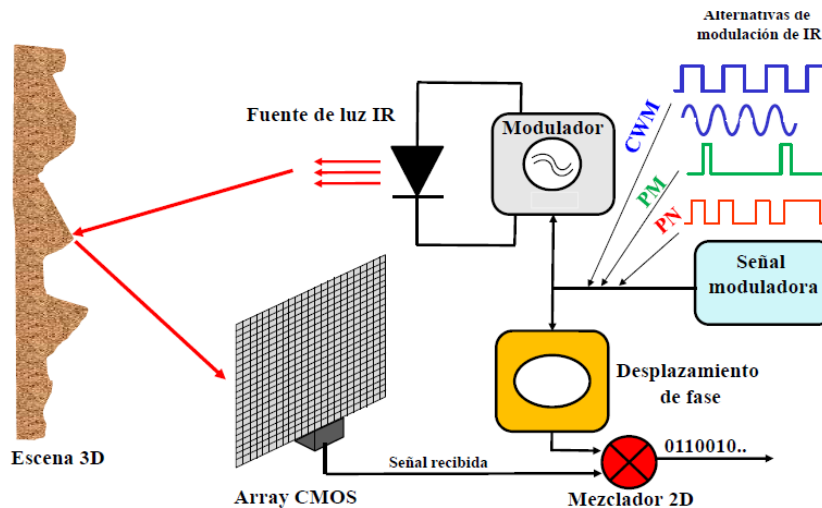


Figura 28 Diagrama de una cámara ToF (12)

- **Modulación de onda continua (CWM):** Con esta tecnología la distancia se obtiene a partir del desplazamiento de una fase entre dos señales, una señal periódica (normalmente señales cuadradas o sinusoidales) utilizada para modular la fuente de luz infrarroja y luz reflejada desde la escena y que los píxeles de la cámara convierten en una señal eléctrica. Este tipo de modulación permite el uso de varias cámaras al mismo tiempo usando frecuencias de modulación diferentes para cada una de las cámaras_(17). Se pueden encontrar cámaras con esta tecnología en firmas comerciales como Mesa y PMD.
- **Modulación pulsada (PM):** En este caso la medida de la distancia a un objeto 3D se obtiene midiendo el tiempo absoluto que un pulso de luz infrarroja necesita para viajar de las fuentes a la escena 3D y viceversa después de la reflexión. Algunos ejemplos comerciales que se basan en esta tecnología de modulación son Basler y Odos, los cuales utilizan el protocolo de red PTP para la sincronización de los relojes, lo que permite el uso de varias cámaras. _(18)
- **Modulación Pseudoruido (PN):** Una posible solución consiste en modular en intensidad la luz infrarroja mediante un código PN (Pseudoruido), obteniéndose el tiempo de vuelo (y en consecuencia la distancia) mediante la correlación de la intensidad de la luz incidente con dos versiones desplazadas en el tiempo de la secuencia PN (aprovechando las propiedades ortogonales de las secuencias PN para la medida de la distancia). La modulación PN facilita la utilización de varias cámaras al mismo tiempo y en un mismo entorno, ya que, basándose en la señal de modulación, cada cámara puede distinguir inequívocamente su propio código PN._(12)

2. ALTERNATIVAS TECNOLÓGICAS Y CARACTERÍSTICAS DE CADA UNA DE ELLAS

2.1. Modulación Continua

Con este tipo de modulación, la obtención de profundidad (distancia) se realiza a partir de la diferencia de fase entre las señales emitidas y recibidas.

La luz se modula alterando la corriente de entrada a la fuente de luz, generando una señal con forma de onda. Se pueden utilizar diferentes formas de señal de modulación, pero normalmente son utilizadas ondas cuadradas o sinusoidales [\(9\)](#)

Hay diferentes métodos para demodular la señal recibida y extraer la información de amplitud y fase. La función de correlación cruzada que realizan los Smart pixels explicados en el apartado anterior es el método más eficiente puesto que no necesitan de circuitos auxiliares y permite el uso de tamaños de píxeles más pequeños lo que conlleva en una mayor capacidad de píxeles en un sensor, resultando en mayor resolución. [\(1\)](#)

Utilizando como señal de modulación de luz emitida una señal sinusoidal $s(t)$:

$$s(t) = a_1 + a_2 \cos(2\pi f_{mod} t) \quad 2.1$$

Obtendríamos una señal continua en el receptor como se muestra a continuación:

$$r(t) = A \cos(2\pi f_{mod}(t - T_L)) + B \quad 2.2$$

Donde:

f_{mod} = Frecuencia de modulación

T_L = Diferencia temporal entre la señal emitida y recibida

a_1 y a_2 = Offset y amplitud de la señal emitida modulada

A = Amplitud de la señal recibida

B = Offset de la señal recibida debido a la iluminación ambiental

La función de correlación cruzada entre la señal emitida y recibida:

$$C(x) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} s(t)r(t-x)dt \quad 2.3$$

Sustituyendo $s(t)$ y $r(t)$ con sus expresiones (2.1) y (2.2) y desarrollando sus términos, obtenemos (19):

$$C(x, \tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} \left(a_2 B \cos(2\pi f_{mod} t) \right. \\ \left. + a_2 A \cos(2\pi f_{mod} t) \cos(2\pi f_{mod} t - 2\pi f_{mod} (T_L + x)) \right. \\ \left. + a_1 A \cos(2\pi f_{mod} t - 2\pi f_{mod} (T_L + x)) \right) dt + a_1 B \quad 2.4$$

Usando las identidades:

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} \cos t dt = 0 \quad 2.5$$

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} \cos t \cos(t-u) dt = \frac{1}{2} \cos u \quad 2.6$$

Obtenemos:

$$C(x, \tau) = \frac{a_2 A}{2} \cos(2\pi f_{mod} (x + T_L)) + a_1 B \quad 2.7$$

Usando las notaciones $\psi = 2\pi f_{mod} x$ y $\phi = 2\pi f_{mod} T_L$, podemos escribir:

$$C(\psi, \phi) = \frac{a_2 A}{2} \cos(\psi + \phi) + a_1 B \quad 2.8$$

Para aumentar la relación SNR, consideramos los valores de la función de correlación en cuatro muestras equidistantes dentro de un periodo de modulación (19):

$$\psi_0 = 0, \quad \psi_1 = \frac{\pi}{2}, \quad \psi_2 = \pi, \quad \psi_3 = \frac{3\pi}{2}$$

Por tanto:

$$C_0 = C(0, \phi), \quad C_1 = C\left(\frac{\pi}{2}, \phi\right), \quad C_2 = C(\pi, \phi), \quad C_3 = C\left(\frac{3\pi}{2}, \phi\right)$$

En la figura 29 se observa estos cuatro valores de muestra que son suficientes para el cálculo inequívoco del offset B , amplitud A y fase ϕ . (20)

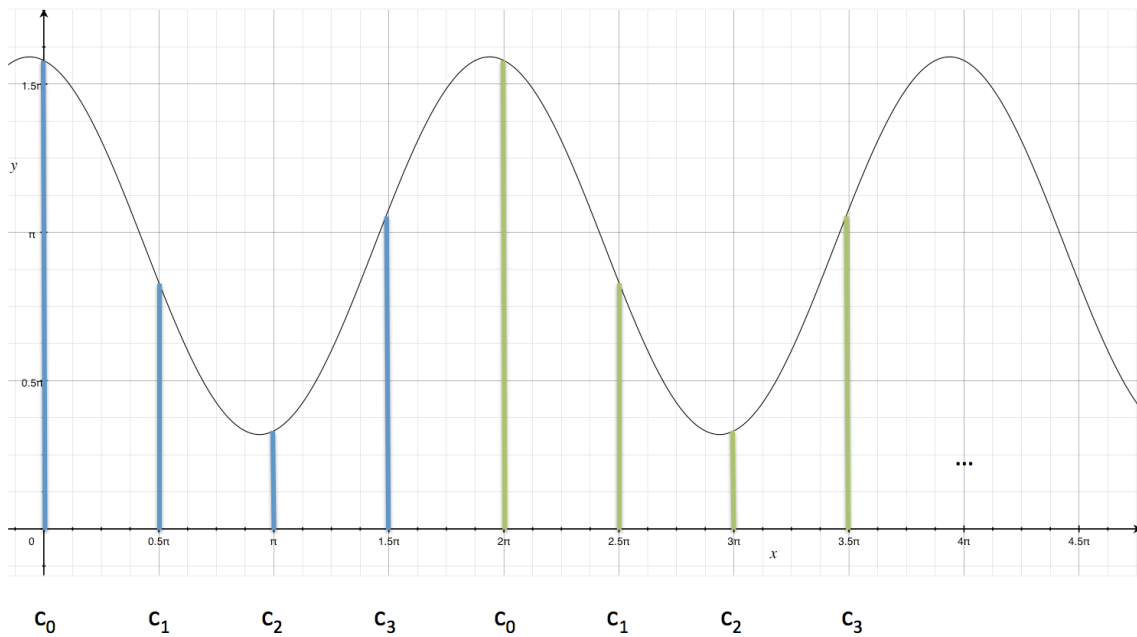


Figura 29 Esta figura muestra el principio general del método de cuatro cubos que estima la señal óptica demodulada en cuatro muestras igualmente espaciadas en un periodo de modulación. Un circuito CCD/CMOS logra la detección de luz, demodulación y almacenamiento de carga. A partir de estos cuatro valores, es posible estimar la fase y amplitud de la señal recibida, así como la cantidad de luz de fondo (offset) (19).

$$\phi = \arctan\left(\frac{C_3 - C_1}{C_0 - C_2}\right) \quad 2.9$$

$$A = \frac{1}{a_2} \sqrt{(C_3 - C_1)^2 + (C_0 - C_2)^2} \quad 2.10$$

$$B = \frac{1}{4a_1} (C_0 + C_1 + C_2 + C_3)$$

2.11

2.2. Modulación pulsada

Este tipo de modulación tiene la ventaja de poder transmitir una gran cantidad de energía en un periodo de tiempo muy corto. De este modo se puede reducir la influencia de la iluminación de fondo consiguiendo una alta relación señal/ruido manteniendo un valor medio de potencia óptica bajo [\(12\)](#).

Otro aspecto característico es que reduce la demanda de una alta sensibilidad y relación señal/ruido en el detector lo que permite realizar medidas de larga distancia [\(12\)](#).

Uno de los principales problemas de este tipo de modulación, es la detección del instante de llegada del haz de luz reflejado. Esto se justifica debido a tres razones principalmente [\(12\)](#):

- El umbral óptico no es un valor fijo, sino que varía en función del fondo y la distancia al objeto.
- La atenuación atmosférica produce dispersión del pulso de luz y aplanar la pendiente del pulso recibido.
- La producción de pulsos de luz muy cortos con tiempos de baja y subida rápidos son difíciles de lograr, los cuales se hacen necesarios para asegurar una detección precisa del pulso de luz recibido.

Para entrar en el detalle de cómo se obtiene la distancia mediante este tipo de modulación, lo primero es generar el pulso de luz emitido, el cual se puede observar en la figura 30:

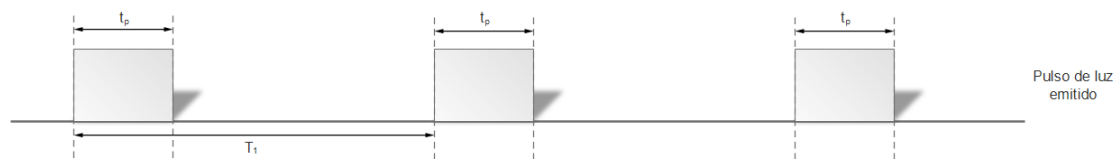


Figura 30 Pulso de luz emitida. Modulación pulsada

Donde t_p corresponde con el tiempo en el que el pulso de luz este activo, y T_1 indica el periodo completo del cada pulso.

A continuación, mostramos cual sería el pulso de luz recibido en el sensor en función del pulso de luz emitido:

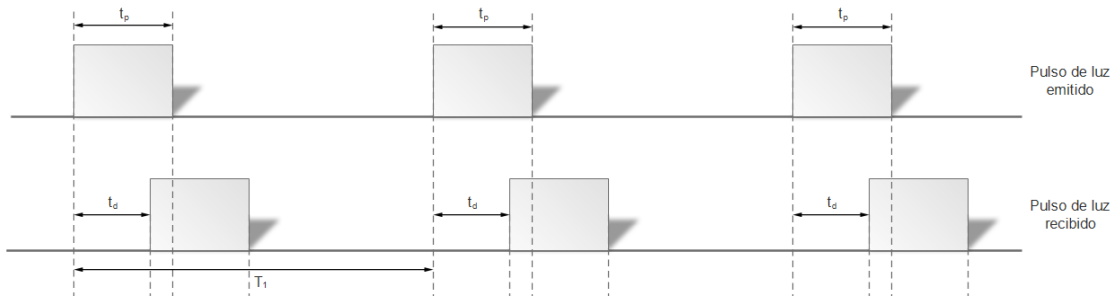


Figura 31 Pulso de luz recibido en función del pulso de luz emitido. Modulación pulsada

Observamos en la figura 31 en este caso el retraso de la recepción del pulso en función del emitido. A este retraso le denominamos t_d .

Con este valor podemos tener una primera aproximación de la distancia (d) a la que se encuentra el objeto de interés:

$$2d = c \cdot t_d \quad 2.12$$

Donde el 2 nos indica el recorrido de ida y vuelta que tiene que realizar la luz desde el emisor al objeto, y de vuelta al receptor.

Al activarse la cámara, todos los LEDs se activan simultáneamente durante el periodo de tiempo t_p y la energía reflejada desde el objeto se recoge en cada pixel del sensor utilizando dos ventanas de obturación con el mismo periodo t_p . Este proceso de emisión de pulsos de luz, apertura y cierre de ventanas de obturación se repite hasta que se completa el tiempo de exposición (21). Por último, para poder obtener la medida, observaremos las cargas eléctricas acumuladas:

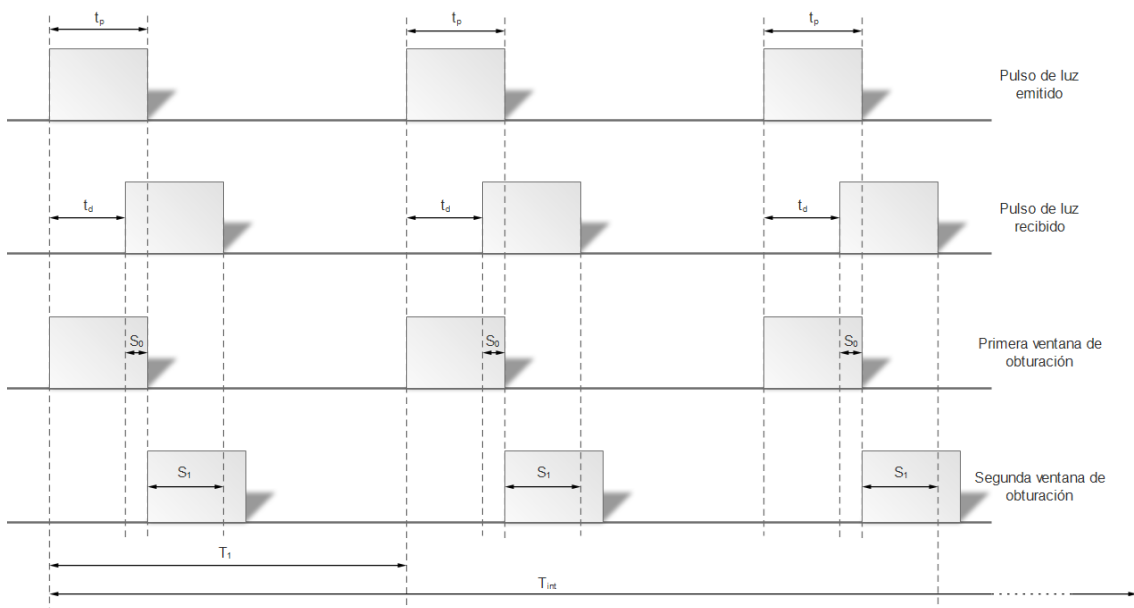


Figura 32 Cargas eléctricas acumuladas en el sensor. Modulación pulsada

Donde:

- S_0 corresponde con la carga eléctrica acumulada en el sensor durante la primera ventana de obturación. Como podemos observar, esta ventana coincide con t_p .

$$S_0 = K(t_p - t_d) \quad 2.13$$

- S_1 corresponde con la carga eléctrica acumulada en el sensor durante la segunda ventana de obturación. En este caso, esta ventana tiene una duración idéntica a t_p , pero se abre al finalizar el impulso enviado.

$$S_1 = K t_d \quad 2.14$$

Con estas variables, podemos determinar la distancia como:

$$t_d = t_p \left(\frac{S_1}{S_1 + S_0} \right) \quad 2.15$$

$$d = \frac{c \cdot t_p}{2} \left(\frac{S_1}{S_1 + S_0} \right) \quad 2.16$$

Con esto podemos determinar la máxima distancia sin ambigüedad, esta es, aquella distancia máxima que podemos medir con el pulso emitido:

$$t_d = t_p \Rightarrow S_0 = 0 \quad 2.17$$

$$d_{max} = \frac{c \cdot t_p}{2} \quad 2.18$$

Además, como hemos comentado anteriormente una de las ventajas de este tipo de cámaras es que permite reducir la influencia de la iluminación de fondo debido a que es posible transmitir una gran cantidad de energía en un tiempo muy corto. Para poder minimizar los efectos de la iluminación de fondo, se introduce una tercera ventana de obturación(12):

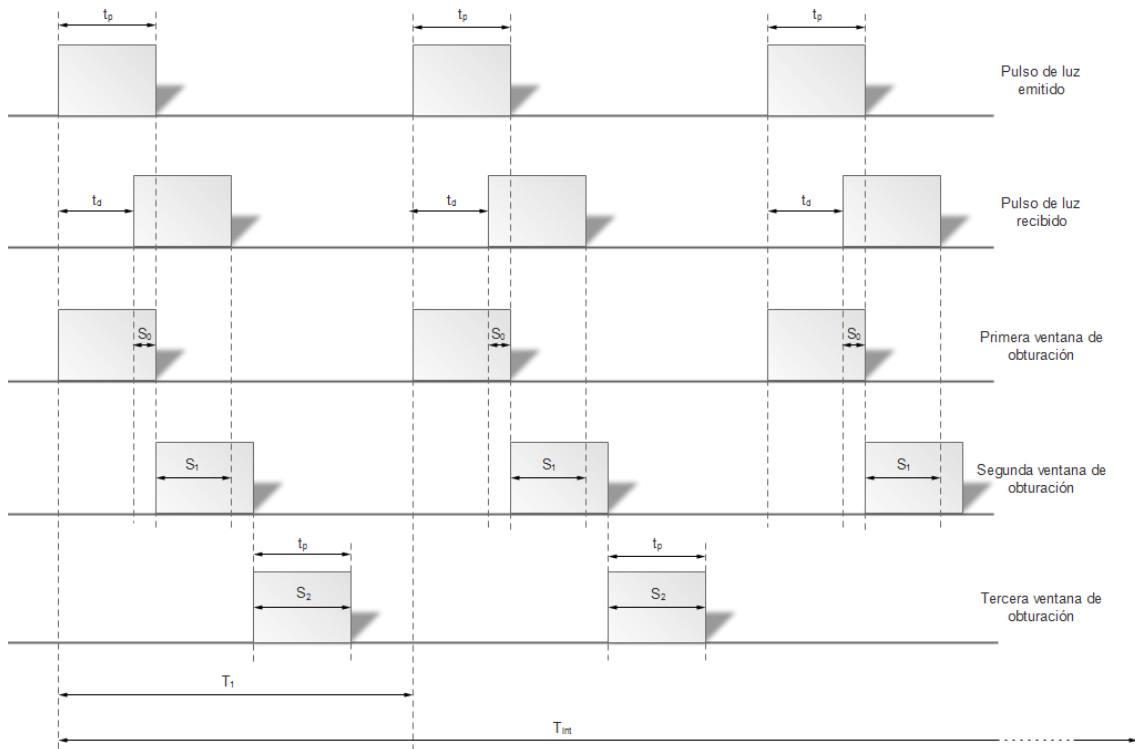


Figura 33 Tercera ventana de obturación. Minimizar iluminación ambiental. Modulación pulsada

Como observamos en el diagrama de la figura 33, la tercera ventana de obturación S_2 tiene una duración t_p idéntica a las dos ventanas anteriores y se abre cuando finaliza la ventana de S_1 . Esta ventana S_2 almacenará la carga provocada por la luz de fondo. A raíz de esta tercera ventana, podemos reescribir las ecuaciones como:

$$\left. \begin{aligned} 2d &= c \cdot t_d \\ S_0 - S_2 &= K_1(t_p - t_d) \\ S_1 - S_2 &= K_1 t_d \end{aligned} \right\} \begin{aligned} t_d &= t_p \left(\frac{S_1 - S_2}{S_0 - S_2 + S_1 - S_2} \right) \\ d &= \frac{c \cdot t_p}{2} \left(\frac{S_1 - S_2}{S_0 - S_2 + S_1 - S_2} \right) \end{aligned} \quad 2.19$$

2.3. Modulación PN (Pseudo-Noise)

En este caso la luz se modula en intensidad mediante una secuencia PN. Este tipo de secuencias se definen como un conjunto de señales binarias, periódicas y de cierta longitud de forma que, dentro de cada periodo, la señal puede aproximarse a una señal aleatoria (ruido blanco).

Existen muchos tipos diferentes de secuencias PN, cada una tiene sus propias ventajas e inconvenientes. Algunos ejemplos podrían ser secuencias *maximal-length* (m -), secuencias *Gold*, secuencias Hadamard-Walsh, secuencias Barker, etc. (14)

Las secuencias m – tienen buenas propiedades de autocorrelación, pero los valores de correlación cruzada entre dos secuencias m – no son despreciables. Las secuencias Gold tienen características opuestas, buena correlación cruzada y pobre autocorrelación. Las secuencias de Barker muestran de nuevo buenas propiedades de autocorrelación, pero su longitud máxima está limitada a 13 bits, lo que prohíbe el funcionamiento en un gran número de cámaras. Las secuencias de Walsh-Hadamard se utilizan ampliamente en las comunicaciones móviles debido a sus atractivas propiedades de ortogonalidad y simplicidad de implementación. (22)

En el caso de las cámaras 3D ToF, el uso más común son las secuencias m – porque muestran excelentes propiedades de autocorrelación que son requerimientos necesarios para la extracción de los valores de distancia y además permiten una sencilla implementación hardware por lo que serán las que definamos para este tipo de modulación en este proyecto. (22)

El tiempo de vuelo (y en consecuencia la distancia) se obtiene mediante la correlación de la intensidad de la luz incidente con dos versiones desplazadas en el tiempo de la secuencia PN (aprovechando las propiedades ortogonales de las secuencias PN) como se muestra en la figura 34. En este ejemplo se muestra una señal óptica modulada por una secuencia PN de longitud de 7 bits. Un único periodo tiene una duración de T_c , que se llama duración del chip. La palabra “chip” es típicamente usada en sistemas de comunicaciones como indicador de un bit en una secuencia binaria, por tanto, usaremos esta misma expresión en este trabajo. La modulación óptica de la señal que ha experimentado un retardo de τ y es detectada por el sensor, se describe como (14):

$$P(t - \tau) = P_{PN} \cdot M_u(t - \tau) \quad 2.20$$

Donde:

P_{PN} = Potencia pico a pico

$M_u(t)$ = Señal óptica basada en una secuencia PN unipolar (0 o +1)

$M_b(t) = S(t) \rightarrow$ Señal de control del interruptor basada en una secuencia PN bipolar (-1 o +1)

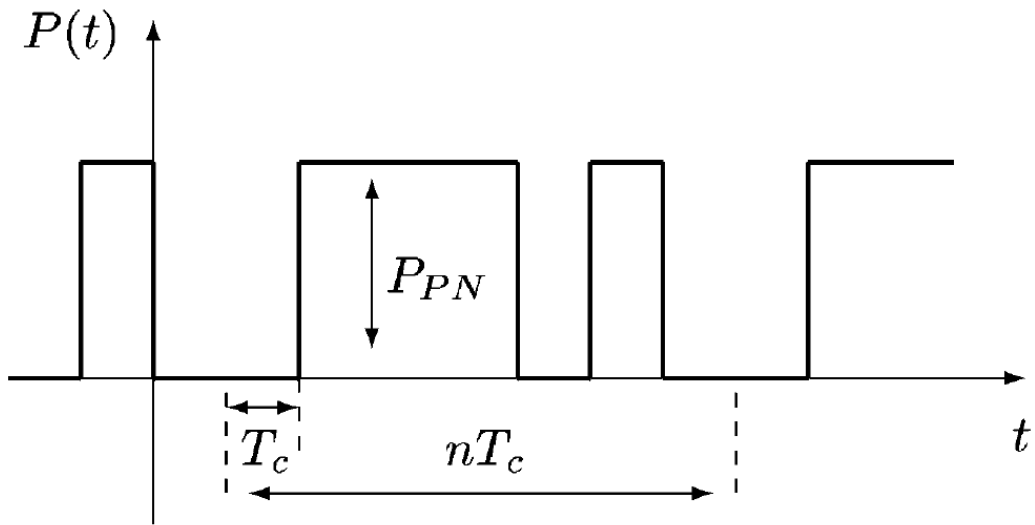


Figura 34 Secuencia PN de longitud máxima $n=7$ chips ₍₁₄₎

La función de correlación de las secuencias m – es mostrada en la figura 35. El pico tiene un ancho de 2 veces la duración del chip T_c . Este valor corresponde al número total de electrones fotogenerados E_S dividido por 2, teniendo en cuenta que E_S denota la suma de electrones generados de ambas muestras _{(14),(23)}.

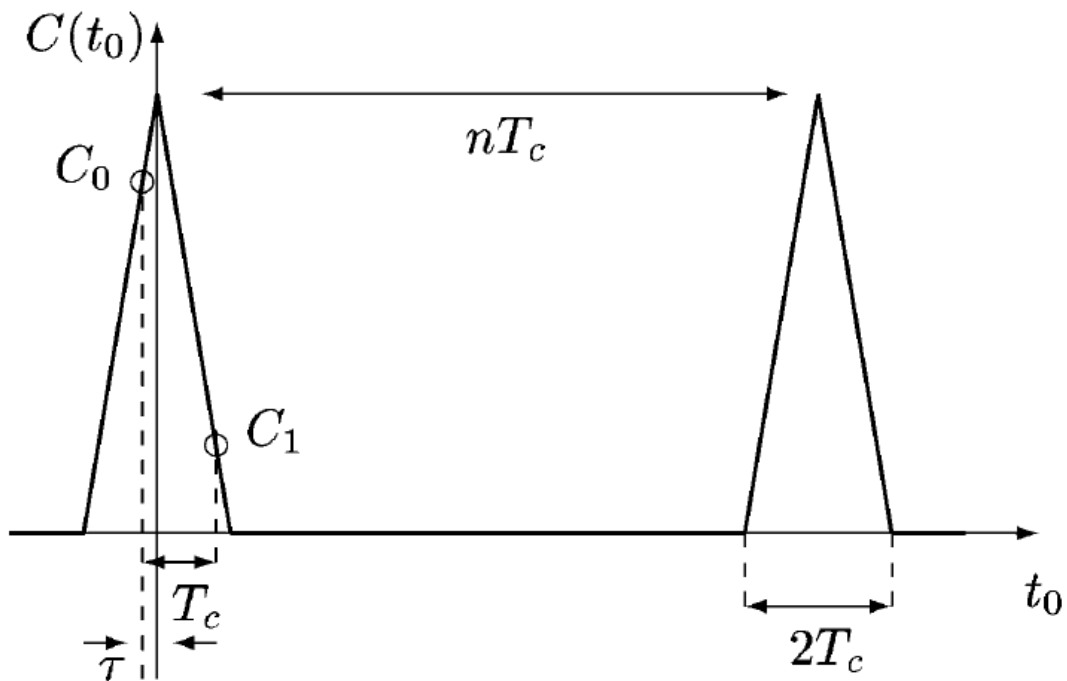


Figura 35 Función de autocorrelación periódica entre señales ópticas moduladas por una secuencia PN unipolar y cuya señal de control de interruptor es de naturaleza bipolar. La duración entre dos picos es dada por el producto entre la secuencia PN de longitud n y la duración de un chip T_c . La forma triangular del pico permite medir la distancia mediante la adquisición de dos muestras C_0 y C_1 en $t_0 = -\tau$ y $t_0 = -\tau + T_c$ respectivamente. ₍₁₄₎

Matemáticamente, la función de correlación es expresada como:

$$C(t_0) = \begin{cases} \frac{E_S}{2} \left(1 - \frac{|t_0 - nkT_c|}{T_c} \right), & \text{para } \epsilon_- < t_0 < \epsilon_+ \\ 0, & \text{resto} \end{cases} \quad 2.21$$

Cuyos límites son:

$$\left. \begin{aligned} \epsilon_- &= T_c(nk + 1) \\ \epsilon_+ &= T_c(nk - 1) \end{aligned} \right\} k \in Z$$

Donde E_S se expresa como:

$$E_S = \frac{n + 1}{n} GP_{PN} = 2GP_A \quad 2.22$$

Con:

$$G = \frac{R_{ph}T}{q} \quad 2.23$$

Donde:

P_A = Nivel de potencia óptica de señal
 G = Eficiencia de generación. Define el número de electrones generados
 R_{ph} = Valor óptico de respuesta de un píxel
 q = Carga correspondiente al número de electrones

Para las medir el tiempo de vuelo, dos valores de correlación C_0 y C_1 son tomados en $t_0 = -\tau$ y $t_0 = -\tau + T_c$ respectivamente (véase figura 35):

$$C_0 = \frac{E_S}{2} \left(1 - \frac{\tau}{T_c} \right) \quad 2.24$$

$$C_1 = \frac{E_S}{2} \frac{\tau}{T_c}$$

Mientras que el valor de correlación C_0 medido solo se ve afectado por el retardo de la señal óptica τ , la segunda medida de C_1 requiere de un desplazamiento adicional entre la señal óptica y la señal de control con duración de un chip T_c , lo que puede lograrse adelantando la señal de control en el tiempo por medio de T_c . Basado en dos muestras C_0 y C_1 el tiempo de vuelo sería (14):

$$\tau = \frac{C_1}{C_0 + C_1} T_c \quad 2.25$$

Teniendo en cuenta que $R = \frac{c\tau}{2}$, la distancia puede ser calculada como:

$$R = \frac{C_1}{C_0 + C_1} R_{max} \quad 2.26$$

Donde la duración de un chip T_c determina la distancia máxima detectable:

$$R_{max} = cT_c/2 \quad 2.27$$

Este tipo de modulación facilita el uso de varias cámaras al mismo tiempo y en un mismo entorno. Basándose en la técnica de modulación aplicada en un ambiente multicámara, un parámetro importante es el número de cámaras que pueden operar al mismo tiempo. El número potencial de cámaras se determina por el número de secuencias diferentes disponibles de modo que cada cámara pueda distinguir inequívocamente su propio código PN. Además, la ortogonalidad de las secuencias PN implica una interferencia mínima entre señales de manera que se minimizan los errores de la medida de distancia (12).

Al igual que en las modulaciones descritas anteriormente, si introducimos la luz ambiental en la ecuación, la señal óptica es expresada como:

$$P(t) = P_B + P_A M_u(t) \quad 2.28$$

La luz ambiente implica un offset positivo en la función de correlación. Solo considerando el pico principal de la función de correlación alrededor de cero, $C(t_0)$ se puede reescribir de la siguiente manera (14):

$$C(t_0) = \begin{cases} \frac{E_S}{2} \left(1 - \frac{|t_0|}{T_c}\right) + \frac{E_B}{2n}, & \text{para } -T_c < t_0 < T_c \\ \frac{E_B}{2n}, & \text{resto} \end{cases} \quad 2.29$$

Donde el parámetro E_S denota el número total de electrones generados por la señal de modulación óptica y E_B describe el número total de electrones generados por la luz ambiental (14):

$$E_B = 2GP_B \quad 2.30$$

Donde:

P_B = Nivel de potencia óptica de luz ambiental

G = Eficiencia de generación. Define el número de electrones generados.

De acuerdo con la [figura 23](#), el valor de correlación es el resultado de la sustracción de dos paquetes de cargas Q_a y Q_b . Se describen matemáticamente alrededor del pico principal mediante las siguientes expresiones_(14):

$$Q_a = \begin{cases} \frac{E_S}{2} \left(1 - \frac{|t_0|}{2T_c} \right) + \frac{E_B}{4} \cdot \frac{n+1}{n}, & \text{para } -T_c < t_0 < T_c \\ \frac{E_S}{4} + \frac{E_B}{4} \cdot \frac{n+1}{n}, & \text{resto} \end{cases} \quad 2.31$$

$$C(t_0) = \begin{cases} \frac{E_S |t_0|}{2 \cdot 2T_c} + \frac{E_B}{4} \cdot \frac{n-1}{n}, & \text{para } -T_c < t_0 < T_c \\ \frac{E_S}{4} + \frac{E_B}{4} \cdot \frac{n-1}{n}, & \text{resto} \end{cases} \quad 2.32$$

3. FUENTES DE ERROR: ERRORES SISTEMÁTICOS Y ALEATORIOS

3.1. Aspectos generales

El error en una medición es la desviación del valor verdadero al valor medido. Las medidas de profundidad con cámaras ToF se enfrentan a la aparición de errores sistemáticos y aleatorios.

Un error sistemático es aquel error que se origina esencialmente por una deficiencia en la calibración del instrumento, en este caso la cámara. Estos errores, por lo general, se pueden reducir al mínimo.

En el caso de los errores aleatorios, son aquellos sin posibilidad de ser controlados, impredecibles, ya que son generalmente por las variaciones cuya causa principal es el azar.

3.2. Errores sistemáticos

De acuerdo con el principio y la arquitectura de la detección ToF, las cámaras de profundidad sufren errores sistemáticos diferentes en función de la tecnología usada. Esto conlleva que los métodos de calibración sean exclusivos para cada tecnología. Estos errores tienen un fuerte impacto en la precisión y fiabilidad de las medidas de profundidad, teniendo su origen normalmente en los procesos de conversión de luz a señal y la fabricación del sensor. (12)

3.2.1. Error circular

Este tipo de error se origina en los cálculos del método de modulación continua, donde la señal de modulación se supone sinusoidal, pero está en realidad, tiene irregularidades, que conllevan en una desviación que depende de la distancia medida. Este efecto se muestra en la figura 36.

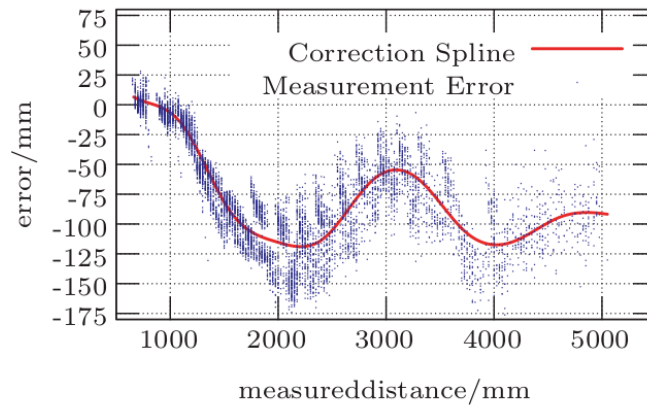


Figura 36 Error circular, medido con la cámara SR300. Los puntos azules muestran las medidas, y la línea roja es un modelo polinómico ajustado al conjunto de datos para su corrección.

Este tipo de error puede ser compensado usando mediciones múltiples que modelen el error. (11)

3.2.2. Errores relacionados con la amplitud

La amplitud de la luz recibida está estrechamente relacionada con la medida de profundidad. Algunas de las razones principales que pueden provocar estos errores son, en primer lugar, la fuente de luz no homogénea que causa diferentes niveles de iluminación en diferentes partes de la escena. Generalmente la potencia de luz va disminuyendo del centro hacia los bordes, como se muestra en la figura 37. Los valores de amplitud más bajos resultan en mediciones más distantes en los bordes, aun cuando se trate de una pared plana. (11)

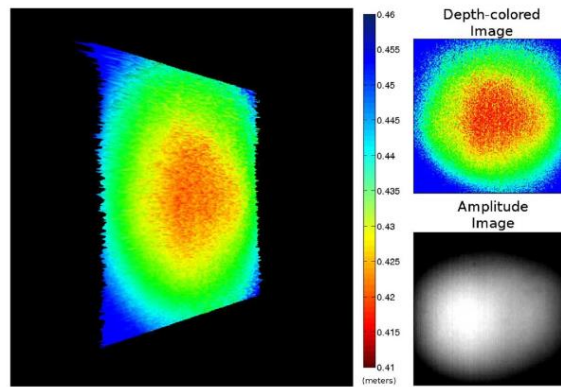


Figura 37 Error relacionado con la amplitud debido a la luz no homogénea. El objeto es una pared plana a 0,43m.

Otra causa que puede provocar estos errores es la diferente reflectividad de los materiales que componen los objetos. Por ejemplo, un papel blanco refleja aproximadamente el 80% de la luz de 850nm, mientras que un neumático de goma negra refleja solo el 2%. Normalmente los materiales de baja reflexión parecen estar más cerca de la cámara. Una escena compuesta de múltiples materiales con diferente reflexión puede provocar errores que causen medidas de profundidad no válidas, ya que hace difícil que se distingan las formas.

3.2.3. Errores relacionados con la temperatura

Este tipo de errores son debidos a la tecnología de los semiconductores, especialmente fotosensores, ya que son altamente sensibles a los cambios de temperatura (24). Los propios fabricantes consideran este error en sus informes sugiriendo que las cámaras deberían calentarse primero para obtener unas medidas estables. El tiempo mínimo que se estima para el correcto funcionamiento de esta tecnología está a partir de los 20 minutos, tiempo mínimo suficiente para poder alcanzar una temperatura estable. (25)

3.2.4. Ruido de patrón fijo (FPN – Fixed Pattern Noise)

Está relacionado con el proceso de fabricación de los sensores de imágenes y la colocación de los píxeles. Cada píxel tiene diferentes características de detección de luz, causando un desplazamiento constante. Las filas/columnas de los píxeles están conectadas en serie, lo que resulta en pequeños retardos de señal consecutivos. El FPN es siempre constante, por lo que puede calibrarse de forma bastante fiable. (26)

3.3. Errores aleatorios

Dentro de estos errores, los más importantes y estudiados son [\(12\)](#):

3.3.1. Dispersión de la luz dentro de la lente y del sensor

Estos errores son provocados por reflejos de luz entre el objetivo y el sensor de imagen debido a la baja sensibilidad del dispositivo como por ejemplo la dispersión debida a superficies brillantes situadas cercas de los LEDs emisores. [\(27\)](#)

3.3.2. Interferencias por multicamino

Los errores de reflexión multitrayecto aparecen porque la luz se refleja desde múltiples superficies antes de regresar al receptor. Múltiples rayos de luz interfieren y causan una medición incorrecta de la profundidad, resultando en esquinas redondeadas como se observa en la figura 38.

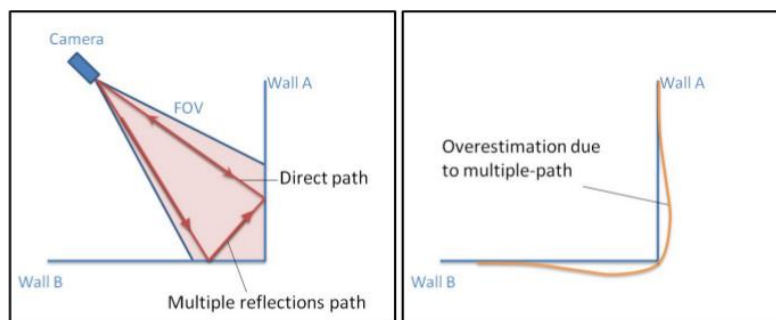


Figura 38 Reflexiones multitrayecto en una esquina. Debido a la interferencia de las ondas de luz, la esquina

Los objetos cóncavos pueden reflejar la luz de modo que nunca sea captada por el sensor, provocando discontinuidades.

Se han propuesto varios enfoques para filtrar los bordes, pero requieren modificaciones en el HW o equipos externos. Además, la mayoría de las implementaciones requieren un número tan elevado de muestras que no pueden ser utilizadas en tiempo real. [\(28\)](#)

3.3.3. Movimiento

Si durante una captura alguno de los objetos de la escena se encuentra en movimiento, provocaría errores en la medida. Esto produciría un ruido aleatorio en los bordes del objeto en cuestión.

3.3.4. Ambigüedad de profundidad

Estos errores son provocados por la ambigüedad en la distancia cuando se supera el rango de no ambigüedad.

4. CONCLUSIONES

En el presente trabajo se han estudiado las diferentes alternativas tecnológicas utilizadas para el diseño de las cámaras de tiempo de vuelo ToF. Actualmente el uso de este tipo de tecnología se está difundiendo en el desarrollo de aplicaciones para dispositivos móviles, donde cada vez es más habitual que se implemente este tipo de cámaras. Esta evolución se debe principalmente por algunas de las ventajas que aporta esta tecnología. Una de las principales, es la simplicidad comparada con otros sistemas, ya que se trata de una solución compacta sin partes móviles y con la iluminación situada junto a la lente.

Además, al ser un proceso directo para extraer la información de distancia, provoca que solo se utilice una pequeña cantidad de potencia de procesamiento. En contraste con los escáneres láser que se mueven y miden punto por punto, las cámaras de ToF miden una escena completa con una única toma con una velocidad de hasta 100 fotogramas por segundo, lo que hace que las cámaras ToF sean mucho más rápidas que sus alternativas láser.

Sin embargo, no todos son ventajas, como hemos estudiado en este trabajo, existen una serie de inconvenientes en el uso de esta tecnología. Sin ir más lejos, unos de los principales problemas a los que se enfrentan son las propiedades reflectantes de los objetos, el color, el brillo y la complejidad de la escena. Estos parámetros tienen efectos considerables en la precisión.

Otro de los factores que pueden provocar errores en la precisión de la medida, es la luz de fondo, ya sea artificial o solar. A diferencia de los sistemas de escaneo láser, las cámaras ToF iluminan toda una escena. Debido a los múltiples reflejos, la luz puede llegar a los objetos a lo largo de varios caminos, lo que provocaría que la distancia medida sea mayor a la real.

Por último, otra de los inconvenientes se presenta cuando existen varias cámaras ToF funcionando al mismo tiempo sobre una misma escena. Esta situación puede perturbar las mediciones de cada una de ellas, aunque existen varias acciones para tratar este problema, como son:

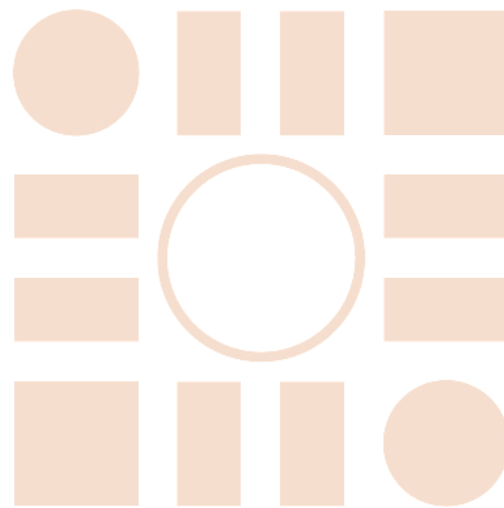
- Multiplexación en tiempo: Un sistema de control inicia la medición de cada una de las cámaras de forma consecutiva, de modo que una unidad de iluminación este activa a la vez.
- Diferentes frecuencias de modulación: Si las cámaras modulan su luz con diferentes frecuencias de modulación, su luz se recoge en los otros sistemas solo como iluminación de fondo, pero no perturba la medición de la distancia.

Bibliografía

1. Lange R. 3D Time-of-flight distance measurement with custom solid-state image sensors in CMOS/CCD-technology; September 8, 2000.
2. G. Bradski AK. Computer Vision with the OpenCV Library.: O'Reilly Media; 2008.
3. Intel. Intelrealsense. [Online]. Available from: <https://dev.intelrealsense.com/docs/depth-post-processing>.
4. López J, Fernández A, Fernández M. Conceptos y Técnicas de Estereovisión por Computador. Revista Iberoamericana de Inteligencia Artificial. 2005.
5. Intel. [Online]. Available from: <https://software.intel.com/es-es/realsense>.
6. Universidad Politecnica de Madrid. [Online]. Available from: <http://www.elai.upm.es/webantigua/spain/Investiga/GCII/areas/luz/luz%20estructurada.htm>.
7. 3D D. Digitalizacion-3D. [Online]. Available from: <https://www.digitalizacion-3d.com/impresion-3d-de-un-busto-partiendo-de-escaneado-3d/>.
8. Orbbec. Astra Series..
9. Luan X. Experimental investiagion of photonic mixel device and development of TOF 3D ranging systems based on PMD technology.; 2001.
10. Möller T, Holger K, Frey J, Alrecht M, Lange R. Robust 3D Measurement with PMD sensors [PMD Technologies GmbH].
11. Laukkanen M. Performance Evaluation of Time-of-Flight Depth Cameras; 2015.
12. Mazo M, Luna C, Losada C. Camaras de Tiempo de Vuelo (Camaras ToF). Alcalá de Henares;; 2018.
13. Hussmann S, Edeler T, Hermanski A. Real-Time Processing of 3D-TOF Data in Machine Vision Applications Germany; February 2012.
14. Büttgen B, El Mechat MA, Lustenberger F, Seitz P. Pseudonoise Optical Modulation for Real-Time 3-D Imaging with minimum interference. IEEE Transactions on circuits and systems. 2007.
15. Bagnuoli G. Interempresas. [Online].; 2019. Available from: <https://www.interempresas.net/Electronica/Articulos/246424-monitorizacion-automovil-mediante-imagenes-3D-ahora-es-realidad-gracias-tecnologia-ToF.html>.
16. Rüter M, Pirker K, Bischof H, Schweighofer G, Mayer H. Int. Conf. Control, Automation, Robotics and Vision; 2010; Singapore.
17. Hansard M, Lee S, Choi O, Horaud R. Time of Flight Cameras - Principles, Methods and Applications: Springer; 2012.
18. Basler. Baslerweb. [Online]. Available from: https://www.baslerweb.com/fp-1565613925/media/downloads/documents/brochure/BAS1906_Camera_Brochure_blaze_A4_SAP0463_No01_EN_web~1.pdf.

19. Horaud R, Hansard M, Evangelidis G, Clément M. An Overview of Depth Cameras and Range Scanners Based on Time-of-Flight Technologies. Machine Vision and Application Journal. 2016.
20. Lange R, Seitz P. Solid-state time-of-flight range camera. IEEE Journal of Quantum Electronics. 2001.
21. Basler. Basler Users Manual. Available from: https://www.baslerweb.com/fp-1523277387/media/downloads/documents/users_manuals/AW00133810000_Tof_GigE_UM.pdf.
22. Zepernick HJ, Finger A. Pseudo Random Signal Processing: New York; 2005.
23. Büttgen , Seitz P. Robust Optical Time of Flight Range imaging base on smart pixel structures. TRANSACTIONS ON CIRCUITS AND SYSTEMS. 2008; 55(6).
24. Timo Kahlmann FRHI. Calibration for Increased Accuracy of the Range Imaging Camera Swinssranger; 2006.
25. Dario Piatti FR. ToF cameras: Tests and comparison in remote sensing; 2012.
26. Stefan MAy DDDHSFEMANaJH. Three dimensional mapping with time of flight cameras. Journal of Field Robotics. 2009.
27. BASLER. Basler ToF Camera, User's manual; 2016.
28. A. A. Dorrington JPGMJCADPaLVS. Separating true range measurements from multi-path and scattering interference in commercial range cameras.; 2011.

Universidad de Alcalá
Escuela Politécnica Superior



ESCUELA POLITECNICA
SUPERIOR



Universidad
de Alcalá