

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN
FACULTAD DE INGENIERÍA MECÁNICA Y ELÉCTRICA
SUBDIRECCIÓN DE POSGRADO



SISTEMA DIGITALIZADOR TRIDIMENSIONAL PARA MAPEO DE ENTORNOS

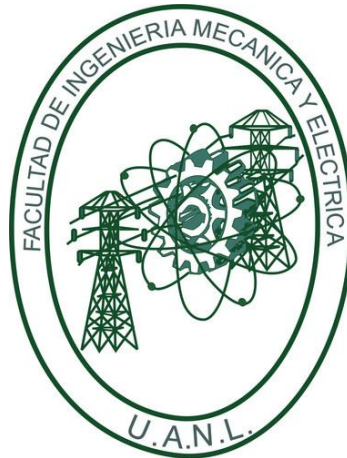
Por

ING. DAVID EDEL PÉREZ GARZA

COMO REQUISITO PARA OBTENER EL GRADO DE MAESTRÍA EN CIENCIAS DE LA INGENIERÍA ELÉCTRICA

[JUNIO, 2020]

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN
FACULTAD DE INGENIERÍA MECÁNICA Y ELÉCTRICA
SUBDIRECCIÓN DE POSGRADO



SISTEMA DIGITALIZADOR TRIDIMENSIONAL PARA MAPEO DE ENTORNOS

Por

ING. DAVID EDEL PÉREZ GARZA

COMO REQUISITO PARA OBTENER EL GRADO DE MAESTRÍA EN CIENCIAS DE LA INGENIERÍA ELÉCTRICA

[JUNIO, 2020]


UNIVERSIDAD AUTONOMA DE NUEVO LEON
FACULTAD DE INGENIERÍA MECÁNICA Y ELÉCTRICA
SUBDIRECCIÓN DE ESTUDIOS DE POSGRADO

Los miembros del comité de tesis recomendamos que la tesis "Sistema Digitalizador Tridimensional para Mapeo de Entornos" realizada por el alumno Ing. David Edel Pérez Garza, con número de matrícula 1252662, sea aceptada para su defensa como opción al grado de Maestría en Ciencias de la Ingeniería Eléctrica.


El Comité de Tesis




Dr. Miguel Angel Platas Garza
Director



Dr. Johnny Rodriguez Maldonado
Co-director

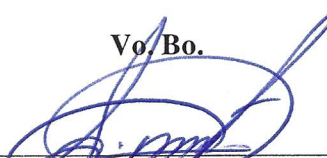


Dra. Norma Patricia Puente Ramirez
Revisor Interno



Dr. Jesus Emmanuel Gomez Correa
Revisor Externo

Vo. Bo.



Dr. Simon Martinez Martinez
Subdirector de Estudios de Posgrado



San Nicolás de los Garza, N.L., 24 de Junio de 2020

DEDICATORIA

para mí gordita ♠

AGRADECIMIENTOS

Agradezco al Consejo Nacional de Ciencia y Tecnología (CONACYT) por el apoyo económico otorgado para la realización de este trabajo a través de la beca de maestría con número de registro 476545.

Resumen

En este trabajo se propone un sensor y un algoritmo para la detección y reconstrucción de objetos y entornos de manera que el sistema no requiera un punto fijo en el entorno. Para poder lograr este objetivo se construyó un sensor basado en un proyector láser y una cámara de alta definición. El sensor envía imágenes del entorno a nuestro sistema para ser procesadas, el cual las procesa para extraer la información de cada imagen de la cual se adquiere una vista parcial en tres dimensiones del entorno. Cada vista parcial se transforma para ser anexada a la anterior para generar una vista completa del entorno. Debido a que el sensor es capaz de proveer a nuestro sistema la posibilidad de extraer suficiente información en cada imagen, los parámetros de transformación, traslación y rotación entre muestras son calculados de tal manera que el sistema no requiere de un punto de referencia fijo en el entorno por lo que el sensor se puede mover por el entorno libremente.

Índice General

Índice General	i
Índice de Tablas	iii
Índice de Figuras	iv
1. Introducción	1
1.1 Justificación	2
1.2 Antecedentes.	3
1.3 Definición del problema	4
1.4 Objetivo	5
1.5 Hipótesis	5
1.6 Metodología	5
1.7 Descripción de la Tesis.	6
2. Sistemas de Visión 3D	7
2.1 Estereopsis	9
2.1.1 Ventajas.	11
2.1.2 Desventajas	11
2.2 Tiempo de vuelo	11
2.2.1 Ventajas.	12

2.2.2	Desventajas	12
2.3	Triangulación	13
2.3.1	Ventajas	14
2.3.2	Desventajas	14
2.4	Flying Triangulation	14
2.4.1	Ventajas	15
2.4.2	Desventajas	15
3.	Flying Triangulation	16
3.1	Principio	19
3.2	Registro entre imágenes	20
4.	Sistema de visión construido	22
4.1	Patrón de luz estructurada.	22
4.2	Extracción de los puntos de interés	24
4.3	Transformación y registro.	28
5.	Resultados	32
5.1	Casos de estudio	33
6.	Conclusiones y trabajo futuro	36
6.1	Conclusiones	36
6.2	Trabajo futuro	37
	Bibliografía	38
	Anexo A. Diagrama de flujo del proceso del algoritmo Flying Triangulation	
	Desarrollado para el sistema de visión construido	40

Índice de Tablas

4.1	Comparación entre un proyector LED y uno láser	23
-----	--	----

Índice de Figuras

1.1	Geometría de la técnica de triangulación	3
1.2	Configuración del sensor	4
2.1	Diagrama de los sistemas de digitalización	8
2.2	Diagrama de los sistemas ópticos	9
2.3	Diagrama de un sistema basado en estereopsis	10
2.4	Diagrama de un digitalizador de tiempo de vuelo.	12
2.5	Diagrama de un digitalizador de triangulación	14
3.1	Diagrama de flujo del proceso del digitalizador del sistema Flying Triangulation	18
3.2	Ejemplo del proceso del digitalizador del sistema Flying Triangulation en varias iteraciones	18
3.3	Configuración del sensor	19
3.4	Registro entre imágenes	21
4.1	CAD del proyecto construido	24
4.2	Gráfica de la intensidad del haz de luz generado por el láser . . .	25
4.3	Pseudocódigo del algoritmo para detectar máximos locales . . .	27
4.4	Ejemplo del uso del escáner.	28
4.5	Diagrama de representación para mostrar dos barridos del sistema de Visión construido de la misma superficie en dos	

	diferentes posiciones	30
4.6	Diagrama para ejemplificar el uso del escáner construido y el resultado obtenido del algoritmo ICP	31
5.1	Ejemplo de láseres proyectados en el objeto a escanear	32
5.2	Sistema de visión construido	33
5.3	Objeto de prueba A para ser escaneado	34
5.4	Resultados del escaneo de objeto de prueba A	34
5.5	Objeto de prueba B para ser escaneado.	35
5.6	Resultados obtenidos del escaneo de objeto de prueba B, vista de 90° y 45°	35
A.1	Diagrama de flujo del algoritmo para el ciclo principal del sistema para el procesamiento de las imágenes del sensor construido	41
A.2	Diagrama de flujo del algoritmo para detectar máximos locales.	42
A.3	Diagrama de flujo del algoritmo ICP desarrollado.	43
A.4	Diagrama de flujo de la función match utilizada en el algoritmo ICP desarrollado	44
A.5	Diagrama de flujo de la función samepoint utilizada en el algoritmo ICP desarrollado	45

Capítulo 1

Introducción

Los humanos poseen sentidos mediante los cuales perciben estímulos, de estos sentidos del cual una persona tiene mayor dependencia es la vista, ya que no solo mediante la visión se recibe información, sino que también se interpreta el entorno físico [1]. Los sistemas de visión computacional 3D tienen el objetivo de extraer información útil de imágenes para con esto interpretar las imágenes tal cual lo haría el ojo humano. Y hoy en día el procesamiento de imágenes nos da la posibilidad de extraer características de un entorno mediante imágenes. Utilizando técnicas como la estereoscopía y la triangulación láser, es posible obtener imágenes de alta precisión a alta velocidad y a su vez mediante el uso de sensores permiten medir distancia, desplazamiento, profundidad y posición de objetos; los cuales mediante estas técnicas y el uso de métodos matemáticos permiten el procesamiento de nubes de puntos para la reconstrucción de entornos en tres dimensiones. Y aunque en la última década se ha tenido un progreso impresionante en el campo de la visión computacional, en donde se observa que las cámaras digitales cuentan con algoritmos tanto como de detección de rostro o de un cuerpo u objeto en tiempo real, así como su profundidad. Aun se presentan algunos problemas entre los cuales está la complejidad visual, en donde en una escena o entorno pudieran existir diversos objetos, a su vez la capacidad computacional para la interpretación de la imagen se ve limitada

debido a la velocidad de procesamiento de los algoritmos para diferenciar en donde comienza y termina un objeto, entre otros [2].

Los sistemas de visión se utilizan en diversas áreas de la industria e investigación. Entre las aplicaciones de los sistemas de visión 3D se encuentran la detección de objetos, identificación facial o biometría, automatización de robots guiados, control de calidad en productos mediante la identificación de defectos, análisis de características y errores, transporte guiado de vehículos, aplicaciones de tráfico, el mapeo 3D en tiempo real, entre otros. La aplicación de interés en esta investigación es en el mapeo de entornos.

1.1 Justificación

En la actualidad existe incremento en demanda de sensores tridimensionales robustos al movimiento, con una tasa del 16.4 anual, puesto que simplifican la tarea de medir superficies complejas al permitir el movimiento del sensor [3,4]. Los sistemas de escaneo en 3D más eficientes utilizan el principio de principio de tiempo de vuelo o el de triangulación con láser (TOF) [5]. En los sistemas basados en tiempo de vuelo se tiene poca precisión, pero un alto rango y variación de profundidad. Por otra parte, lo opuesto ocurre con los sistemas basados en triangulación, los cuales tienen una gran precisión, pero la variación y profundidad son limitados [5]. Los escáneres más nuevos intentan resolver el problema del rango al intentar reducir la necesidad de la manipulación del objeto para tener un punto de referencia, como se hace en el principio de triangulación, esto mediante el uso de sensores que permitan calcular por software la orientación y distancia del escáner con respecto al objeto, sensores tales como acelerómetros,

giroscopios y sensores infrarrojos de distancia. Sin embargo, en este trabajo se presenta un sensor a base de láser, el cual sea capaz de eliminar la necesidad de tener un punto de referencia fijo en el espacio, o el uso de sensores adicionales de manera que se aumente el volumen de los objetos o escenas que puedan ser digitalizadas y reconstruidas en tres dimensiones conservando el alto nivel de precisión obtenido por la triangulación a base de láser.

1.2 Antecedentes

El principio de la triangulación es el más usado para reconstrucción en tres dimensiones (3D) y se basa en la proyección de un patrón de luz, por lo general una línea generada por un láser, sobre el objeto que se quiere digitalizar y se captura la imagen mediante una cámara digital. Con esto se procesa la imagen y se calcula la distancia del objeto por trigonometría con la que se obtiene una nube de puntos, después el objeto es rotado para proyectar el patrón en otra zona del objeto, hasta escanearlo completamente (ver Figura 1.1).

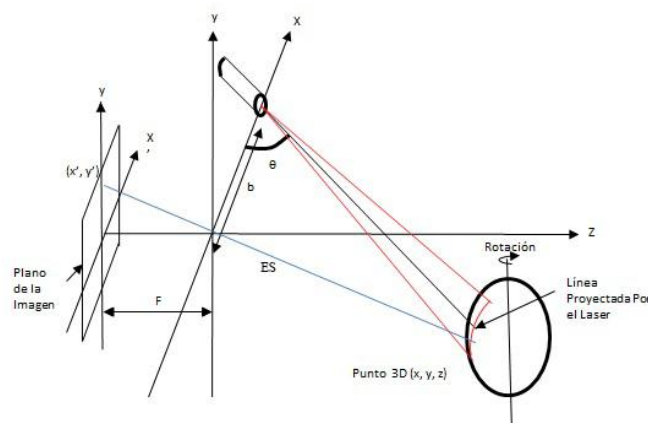


Figura 1.1: Geometría de la técnica de triangulación: El lente se sitúa en el origen con una distancia focal F al plano de la imagen, y una línea base b entre el proyector y la cámara.

Una nueva técnica de reconstrucción 3D basada en la triangulación es *Flying Triangulation*, la cual consta de un sensor guiado alrededor de un objeto o superficie mientras toma una serie de imágenes (ver Figura 1.2), después de aplicar la calibración, cada imagen entrega una vista parcial en tres dimensiones, la que es registrada de manera automática, junto a la adquirida previamente, y estos resultados son visualizados en tiempo real, lo que permite saber que parte del objeto aún no es medida, después se obtiene una nube de puntos densa de la topografía del objeto [6]. Una parte crítica de este proceso es la identificación de las líneas proyectadas [7]. Las principales características de *Flying Triangulation* son su robustez al movimiento, respuesta en tiempo real, escalabilidad y flexibilidad [8].

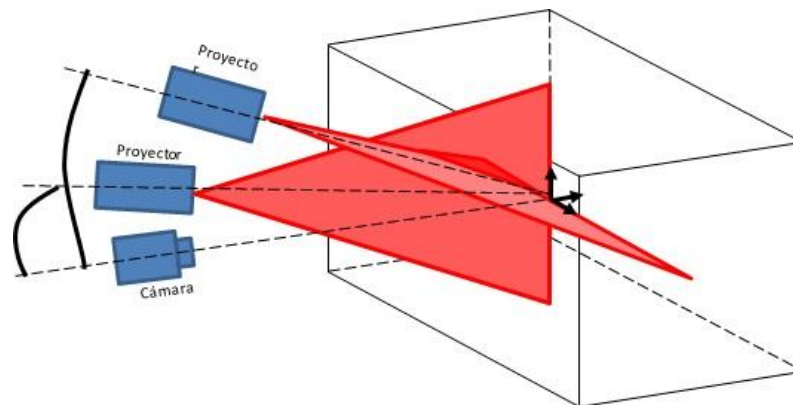


Figura 1.2: Configuración del sensor, consiste en dos proyectores ortogonales y una cámara.

1.3 Definición del problema

El proyecto intenta crear un sistema de visión computacional que sea capaz de reconstruir escenas u objetos en tres dimensiones casi en tiempo real. El desarrollo de este proyecto consta de la elaboración de un sensor guiado alrededor de un objeto o

superficie mientras toma una serie de imágenes, las cuales son procesadas por medio de la técnica de *Flying Triangulation*.

1.4 Objetivo

Diseñar un sistema de visión en tiempo real capaz de reconstruir una escena u objeto mediante la proyección de un patrón estructurado de luz sobre el entorno de tal modo que el sistema no requiera un punto de referencia fijo, mediante el método de *flying triangulation*.

1.5 Hipótesis

Mediante el uso de un patrón de luz estructurada sobre una superficie o un objeto es posible obtener la suficiente información para poder digitalizar la superficie u objeto en tiempo real con una secuencia de fotografías.

1.6 Metodología

- Analizar los diferentes métodos de proyección de luz estructurada y los diversos tipos de patrones para poder valorar las desventajas y ventajas de cada uno.
- Diseñar un patrón de luz estructurada que permita adquirir la mayor cantidad de puntos de interés y construir el sistema de proyección que será usado en el sistema de visión.
- Desarrollar un algoritmo capaz de filtrar la imagen para extraer o interpretar los puntos de interés contenidos en la imagen adquirida por el sistema de visión.

- Diseñar un visualizador que sea capaz de representar los datos adquiridos de manera gráfica y rápida de interpretar.
- Integración y optimización del sistema. Generación de propuestas para embeber el sistema de tal forma que se pueda implementar en sistemas de navegación.

1.7 Descripción de la Tesis

El resto de este documento está organizado de la siguiente manera: en el Capítulo 2 se describen los sistemas de visión 3D, así como las ventajas y desventajas que presentan. En el Capítulo 3 se explica el método de *Flying Triangulation*, el cual es el método de estudio en esta investigación. En el Capítulo 4 se describen las fases para la construcción del sistema de visión utilizado en esta Tesis. En el Capítulo 5 se muestran los resultados obtenidos por el sistema de visión construido y se discuten las conclusiones obtenidas de esta Tesis.

Capítulo 2

Sistemas de Visión 3D

El ser humano percibe el entorno que lo rodea en tres dimensiones con la capacidad de poder distinguir formas, luces, colores, entre otras características. Diversos investigadores en visión computacional han intentado emular esta capacidad mediante métodos matemáticos con el objetivo de restaurar la forma tridimensional del objeto en la imagen [9].

Los sistemas de visión en tres dimensiones se utilizan para adquirir, procesar y analizar imágenes, su objetivo principal es la extracción de información a partir de imágenes o fotografías mediante el uso de cámaras y sensores, utilizando algoritmos para realizar el análisis automático de las mismas, sin embargo, esto es un proceso complejo, ya que para reconstruir una escena es necesario el uso de múltiples de cámaras y sensores de alta velocidad que sean robustos y de alta precisión, así como también de equipos de cómputo de alta capacidad de procesamiento ya que a mayor velocidad, son más la cantidad de imágenes que el algoritmo requiere procesar.

Los sistemas para la digitalización en tres dimensiones pueden ser clasificados en 3 categorías, de contacto, de transmisión y de reflexión, sin embargo, las primeras dos

requieren de escáneres en un punto fijo donde el volumen del objeto que se puede digitalizar es proporcional al tamaño del escáner en si. Por otra parte, en cuanto a la reflexión, estos pueden ser a su vez subdivididos en sistemas ópticos y no ópticos, debido a las propiedades de precisión que poseen los sistemas ópticos este trabajo considera la construcción de un sistema óptico.

Los sistemas ópticos pueden ser a su vez divididos en dos categorías, activos y pasivos, esto dependiendo de si hay un emisor y un receptor o solo un receptor.

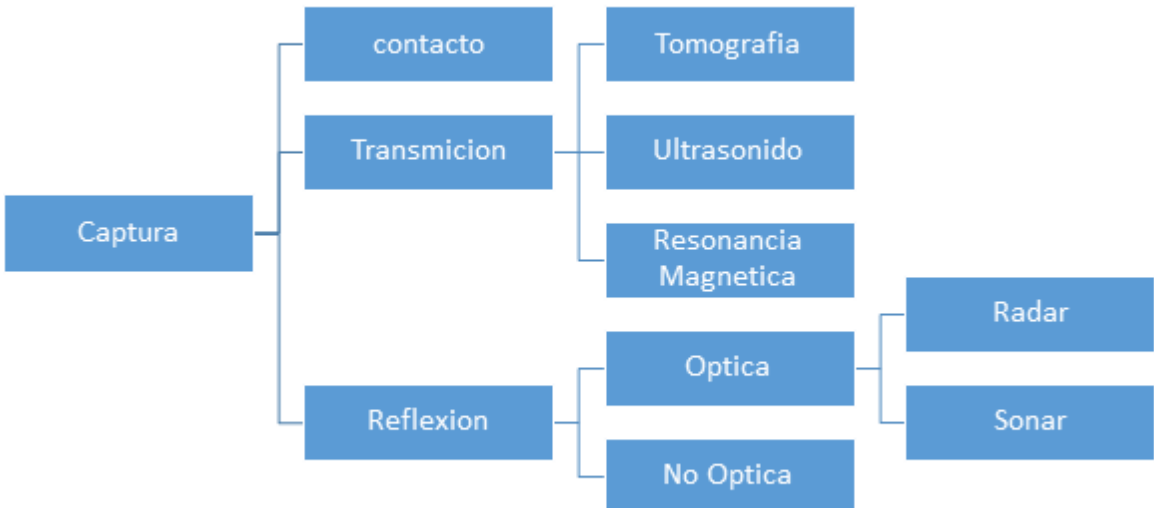


Figura 2.1: *Diagrama de los sistemas de digitalización*

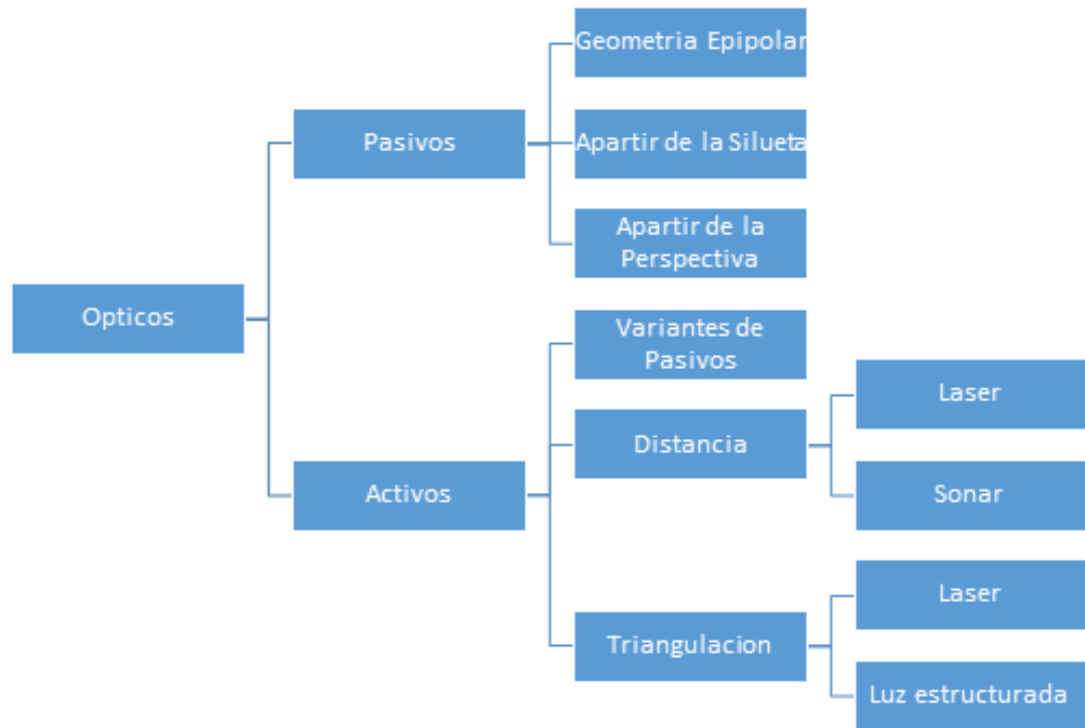


Figura 2.2: Diagrama de los sistemas ópticos.

2.1 Estereopsis

El principio de estereopsis fue descrito en 1838 por Charles Wheatstone [10], el cual se utilizó un siglo después para crear los primeros escáneres tridimensionales, es decir, es uno de los métodos más antiguos para la obtención de reconstrucción de superficies tridimensionales mediante la obtención de distancias tomadas mediante múltiples cámaras a los objetos u escena [11]. En la Figura 2.3 se muestra un diagrama de un sistema de visión basado en estereopsis, en donde se puede ver que se utilizan dos o múltiples cámaras separadas o en diferentes posiciones a una distancia relativa entre las mismas para con esto las cámaras captar múltiples imágenes con algo de desplazamiento entre las imágenes capturadas de una misma escena u objeto, con esto, la reconstrucción de una imagen en 3D se basa en la obtención de distancias de las cámaras con respecto al objeto, esto mediante un proceso de triangulación (ver Sección 2.3) [11].

Técnicamente consiste en el procesamiento de dos imágenes obtenidas mediante múltiples cámaras, enfocadas en la misma escena desde diferentes posiciones. Este procesamiento está orientado a reconstruir la escena en 3D a partir de las dos imágenes, izquierda y derecha. El trabajo del sistema consiste en identificar en ambas imágenes aquellos píxeles en las dos imágenes que se corresponden con la misma entidad física en la escena 3D, usando para ello algoritmos especializados. La distancia que separa estos píxeles se conoce como disparidad. La medida de la disparidad sirve para obtener la distancia a la que se sitúa físicamente ese objeto en la escena con respecto a las dos cámaras [5, 13].

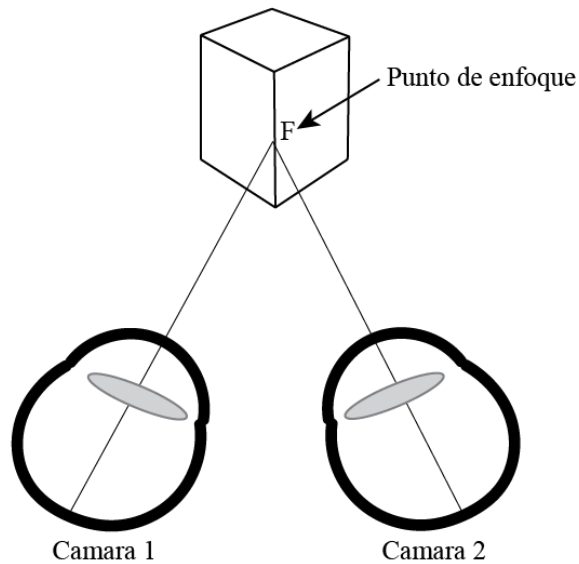


Figura 2.3: Diagrama de un sistema de visión basado en la estereopsis (editado a partir de Figura en [14]).

En resumen, un sistema de visión basado en estereopsis consta de las siguientes fases:

- Adquisición de múltiples imágenes tomadas por dos o más cámaras.
- Geometría del sistema: obtención de la distancia 3D.
- Extracción de características: extracción de los puntos de borde y regiones (superficie de interés de la escena en 3D delimitada por bordes).

- Correspondencia de características.
- Obtención de la distancia.

2.1.1 Ventajas

- Con el uso de múltiples cámaras es posible reconstruir un objeto completo en un instante.
- Es un sistema pasivo que no altera la escena.
- Puede capturar el color del objeto sin aumentar la carga computacional ni el tiempo de ejecución.

2.1.2 Desventajas

- Solo puede reconstruir escenas estáticas.
- Requiere de un sistema de sincronización preciso entre las cámaras.

2.2 Tiempo de vuelo

El escáner de tiempo de vuelo mide el tiempo que tarda un impulso del láser en salir del sensor llegar al objeto y regresar, con ello calcula la distancia de cada punto del panorama para obtener un relieve [5].

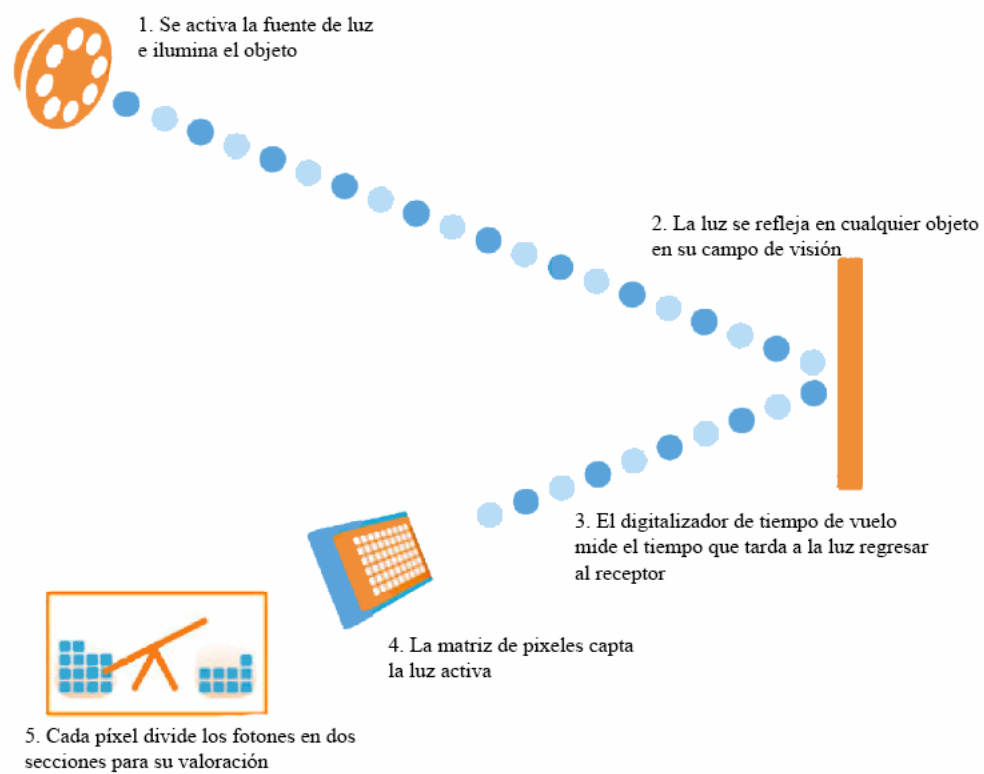


Figura 2.4: *Diagrama de un digitalizador de tiempo de vuelo (editado a partir de Figura en [15]).*

2.2.1 Ventajas

- Puede digitalizar objetos de gran volumen (dimensiones de 70m a 1,500m).
- Robusto al ruido de la iluminación externa.

2.2.2 Desventajas

- Solo sirve para objetos que reflejen en todas direcciones.
- Baja precisión (1mm a 10mm).

2.3 Triangulación

La triangulación en los sistemas de visión se utiliza para la reconstrucción de escenas, se basa principalmente en la técnica de fotogrametría [16], la cual consiste en la medición por medio de fotografías con el objetivo de crear un mapeo de entorno, sin embargo, utilizando una sola imagen solo es posible obtener dos dimensiones, por lo cual para crear coordenadas en tres dimensiones además de fotogrametría es necesario utilizar el principio de estereopsis [5], la cual se describe en la Sección 2.1. La triangulación consiste en determinar la profundidad a la que se encuentra la superficie de un objeto, mediante el análisis de una imagen de la superficie deseada sobre la cual se refleja un haz de luz láser, basándose en los planteamientos teóricos del comportamiento de la perspectiva. Esto se logra capturando en una cámara la posición del espectro del láser en la superficie del objeto deseado. Al acercarse la superficie al sensor, el espectro reflejado se desplaza en sentido horizontal si el observador se encuentra a un lado del láser. Esto se produce debido al ángulo de incidencia del láser con respecto al sensor. El desplazamiento registrado, se puede relacionar directamente con la distancia en la que se ubica la superficie respecto al sensor [17]. Al controlar la posición del objeto con respecto al haz de luz proyectado y el sensor es posible adquirir la profundidad del objeto en cualquier dirección. El esquema de triangulación es mostrado en la Figura 2.5.

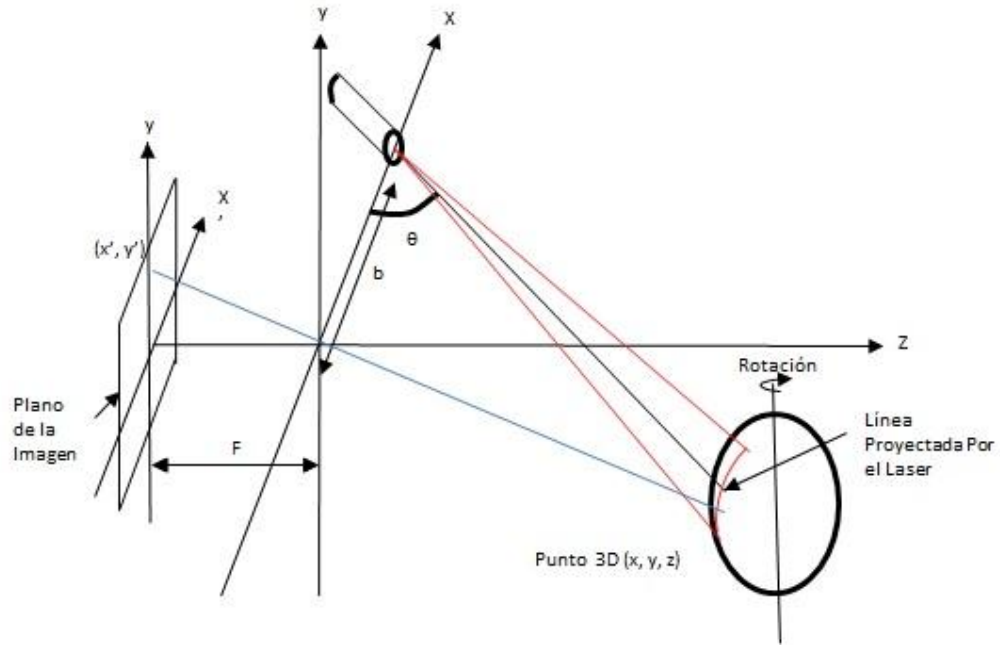


Figura 2.5: Diagrama de un digitalizador de triangulación.

2.3.1 Ventajas

- Alta precisión (0.05mm)
- La precisión no depende en su totalidad de la resolución de la cámara.

2.3.2 Desventajas

- El sistema necesita manipular el objeto sobre un punto de referencia conocido.
- La distancia máxima del sensor al objeto es baja (400cm)

2.4 Flying Triangulation

El sistema de *Flying Triangulation* es capaz de medir la superficie completa del objeto que se desee escanear moviéndose libremente a su alrededor tomando vistas parciales del

sensor rápidamente de tal manera que el sensor pueda evitar la pérdida de precisión causado por el movimiento. Esto es logrado mediante la combinación de un sensor simple con algoritmos sofisticados, cada vista parcial adquirida por el sensor genera una vista 3D parcial, lo que garantiza la robustez al movimiento, debido a que solo se genera una vista dispersa en lugar de una completa. La vista completa del objeto se genera tomando una serie de imágenes del objeto mientras el sensor se mueve libremente alrededor de él. Cada vista parcial generada es registrada y alineada con las demás para generar una vista 3D densa [3,4,5].

2.4.1 Ventajas

- Velocidad y precisión.
- Robustez al movimiento, respuesta en tiempo real, escalabilidad y flexibilidad

2.3.2 Desventajas

- Se requiere de cámaras de alta velocidad
- Equipos con una capacidad alta de procesamiento, debido que, a mayor velocidad, son más la cantidad de imágenes que se requiere procesar.

Capítulo 3

Flying Triangulation

Las cámaras y sensores utilizados en la actualidad son capaces de captar objetos en tres dimensiones, siendo utilizados en una gran variedad de aplicaciones, sin embargo, existe poco enfoque en el estudio de las cámaras y sensores con enfoque en mapeo en tres dimensiones en tiempo real [18,19]. Para poder obtener una superficie en tres dimensiones se requiere de la adquisición de múltiples vistas o imágenes, para esto es necesario tener un sensor que pueda estar moviéndose a través de la superficie, así como equipo de cómputo robusto y algoritmos sofisticados. En 2009 Ettl, S. et al [20] proponen un método conocido como *Flying Triangulation*.

El principio de *Flying Triangulation* es similar al proceso de triangulación normal, sin embargo, propone que si se capturan suficientes puntos en cada imagen adquirida no es necesario tener un punto de referencia fijo ni conocido.

El sistema de *Flying Triangulation* es capaz de medir la superficie completa del objeto que se desee escanear moviéndose alrededor del objeto libremente tomando vistas parciales del sensor rápidamente, de tal manera que el sensor pueda evitar la pérdida de precisión causada por el movimiento. Esto es logrado mediante la combinación de un

sensor simple con algoritmos de procesamiento de imágenes, tal como el algoritmo para detectar máximos locales y el algoritmo iterativo de vecinos cercanos (*Iterative Closest Point* por sus siglas en inglés ICP), los cuales se describen en las Secciones 4.2 y 4.3 respectivamente. Cada vista parcial adquirida por el sensor genera una vista 3D parcial, lo que garantiza la robustez al movimiento, debido a que solo se genera una vista dispersa en lugar de una completa. La vista completa del objeto se genera tomando una "película" del objeto mientras el sensor se mueve libremente alrededor de él, cada vista parcial generada es registrada y alineada con las demás para generar una vista 3D densa [3,4,20].

El proceso consiste en proyectar un patrón de luz estructurada a base de láser sobre un objeto o escena, capturar la imagen del área donde se proyecta el patrón de luz, filtrar la imagen para después extraer los puntos donde se encuentra proyectado nuestro patrón de luz para generar una nube de puntos la cual nos da una vista parcial en tres dimensiones del objeto o escena. Posteriormente se calculan los parámetros de rotación y traslación para que esta nube de puntos tenga el mismo origen que la nube de puntos anterior, a excepción de la primera, la cual se toma como referencia para las siguientes, y registrarla con la nube anterior de tal manera que la densidad de la nube de puntos aumente su densidad hasta ser lo suficientemente densa para la aplicación deseada. En la Figura 3.1 se muestra el diagrama de flujo del proceso descrito y en la Figura 3.2 se puede ver un ejemplo.

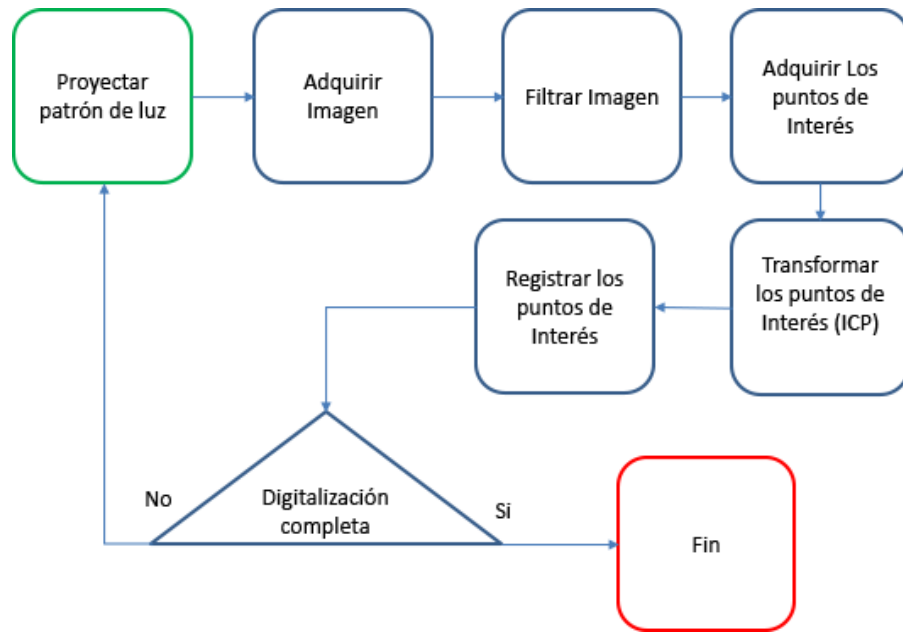


Figura 3.1: Diagrama de flujo del proceso del digitalizador del sistema Flying Triangulation.

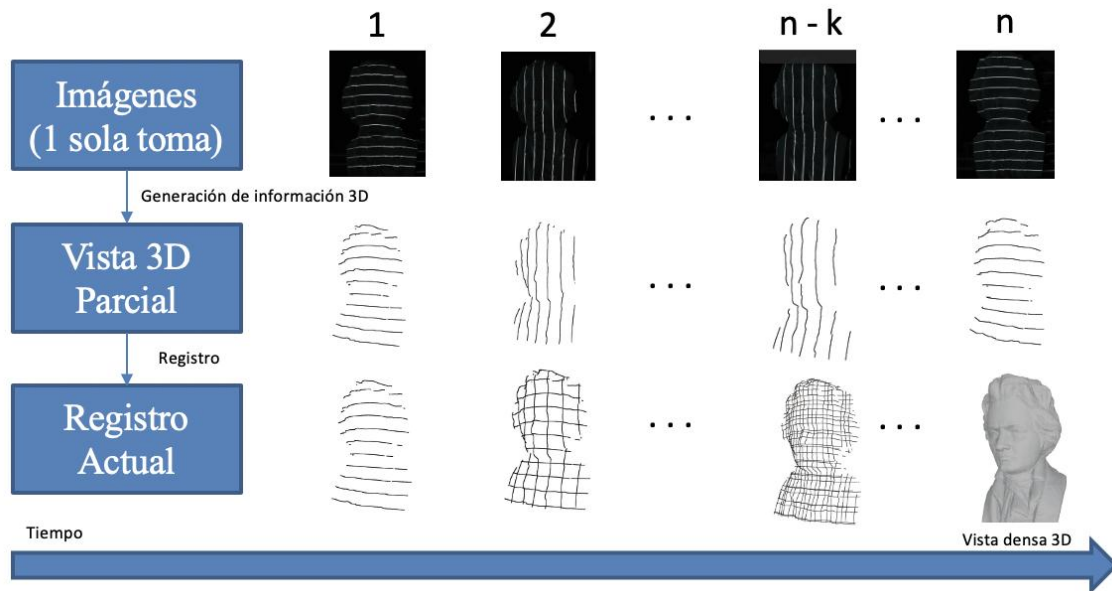


Figura 3.2: Ejemplo del proceso del digitalizador del sistema Flying Triangulation en varias iteraciones (editado a partir de Figura en [4]).

3.1 Principio

El proceso de *Flying Triangulation* consta de un sensor guiado alrededor de un objeto o superficie mientras toma una serie de imágenes. Después de aplicar la calibración, cada imagen entrega una vista parcial en tres dimensiones, la que es registrada de manera automática, junto a la adquirida previamente, hasta que se obtiene una nube de puntos densa de la topografía del objeto [6].

El sensor utilizado en el proceso de *Flying Triangulation* consta de un proyector de líneas láser sobre la escena y una cámara de alta definición para capturar la escena. El proyector deberá estar colocado paralelo a la cámara y rotado en el eje paralelo a las líneas proyectadas para que sea posible visualizar la deformación en el patrón de luz, como se muestra en la Figura 3.3.

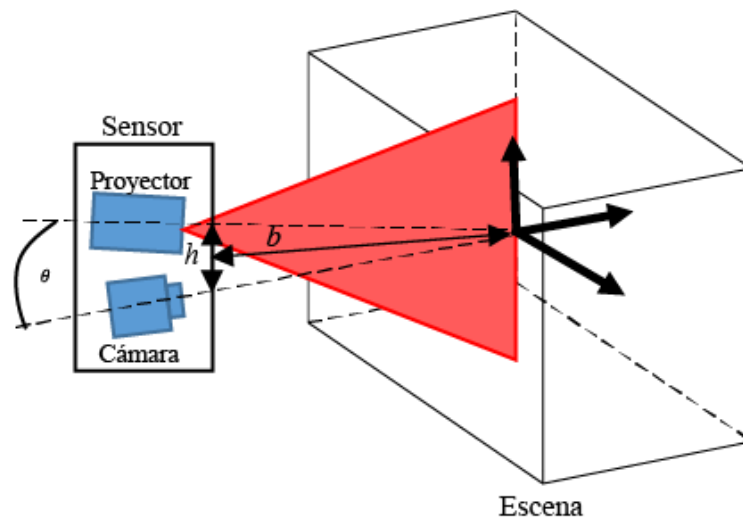


Figura 3.3: Configuración del sensor.

Como se puede observar en la Figura 3.3 para poder realizar la triangulación y medir la deformación sobre el patrón de luz es necesario conocer la distancia h definida como la distancia entre el proyector de nuestro sensor y la cámara, junto con el ángulo de rotación del eje proyector con respecto al eje de sensor de la cámara, y la distancia de nuestro sensor a la escena b , con estos tres parámetros es posible calcular la distancia de cada punto con respecto al sensor.

3.2 Registro entre imágenes

El registro entre imágenes permite reconstruir una escena, objeto o superficie y es una técnica utilizada para la fusión o unión entre imágenes para con esto realizar una reconstrucción de la imagen tridimensional [12], así como en la detección de cambios al momento de relacionar las posiciones espaciales entre imágenes. El principio del registro se puede observar en la Figura 3.4, en donde en la Imagen A aparecen diversos objetos en una determinada posición, y en la Imagen B se observa que la posición de los objetos cambia en rotación, translación y en el tamaño de los objetos debido a un posible alejamiento de la cámara, entonces, el registro entre imágenes consiste en transformar los datos que se obtienen de una imagen a un conjunto de coordenadas, para con esto estimar las coordenadas de la Imagen B a partir de la posición o coordenadas de la Imagen A [12].

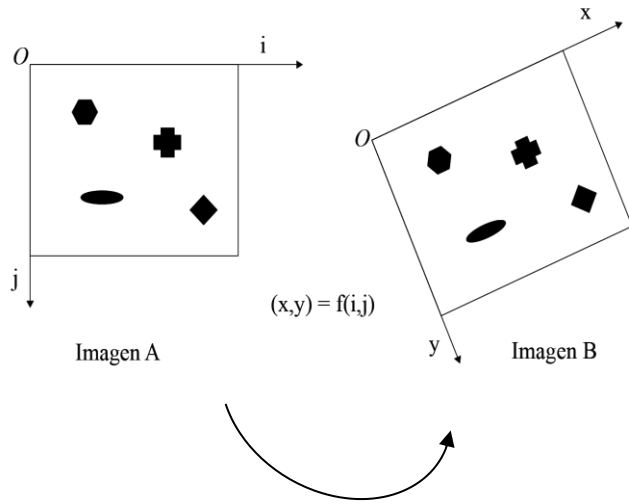


Figura 3.4 Registro entre imágenes (editado a partir de Figura 3.4 en [12]).

Capítulo 4

Sistema de visión construido

El proceso para el desarrollo del sistema de visión se dividió en 3 fases, seleccionar el tipo de proyector que se debería de utilizar, construcción del proyector, desarrollo del algoritmo de escaneo. Este último se divide en captura y extracción de información, transformación de la información obtenida al punto de referencia establecido por el algoritmo y registro en nuestro objeto de salida.

4.1 Patrón de luz estructurada

El patrón de luz estructurada está formado por un grupo de líneas paralelas de forma horizontal y otro vertical que son alternados en dos fotos consecutivas donde la relación de posición entre la cámara y la escena es la misma. Para generar estas líneas hay dos métodos muy comunes, LED y láser, cuyas características se muestran en la Tabla 4.1.

LED	LÁSER
Fácil de construir.	Construcción requiere de espejos en movimiento.
La luz ambiental puede ocasionar que los valores de intensidad en cualquier punto de la imagen sean muy cercanos a los que contiene la información del patrón de luz.	Baja sensibilidad a la luz ambiental.
La divergencia de la luz en el patrón inyecta ruido a las mediciones.	Poca divergencia.
	El patrón puede ser adquirido a mayor distancia.

Tabla 4.1: *Comparación entre un proyector LED y uno láser.*

Después de comparar ambas técnicas se optó por utilizar el láser para generar nuestro patrón. Para generar el patrón de luz estructurada se construyó un arreglo de láseres los cuales apuntan a un espejo que gira a alta velocidad reflejando el haz del láser hacia nuestra superficie creando una línea por cada láser.

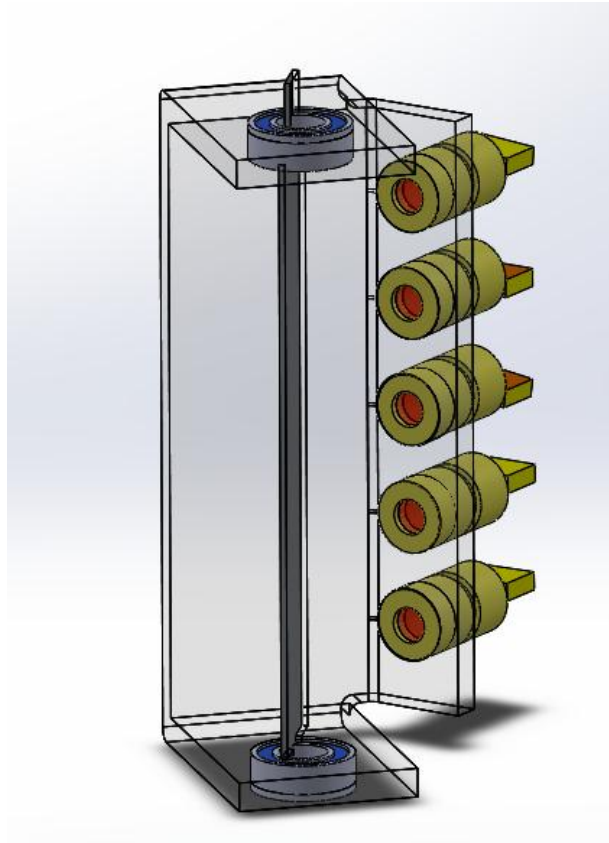


Figura 4.1: CAD del proyecto construido.

4.2 Extracción de los puntos de interés

Para la detección del patrón de luz se aprovecha que la distribución de intensidad del haz de luz generado por el láser sobre la escena es una gaussiana, como se muestra en la Figura 4.2, lo cual nos permite utilizar la correlación cruzada entre una función gaussiana de dos dimensiones y la información de la imagen [22]. La correlación cruzada resalta los puntos de la imagen en los cuales la variación de intensidad se asemeje a la función gaussiana, la cual está dada por la Ecuación 4.0.1.

$$g(r) = \exp\left(\frac{-x^2}{w_x} - \frac{y^2}{w_y}\right) \quad (4.0.1)$$

Donde x y y representan las coordenadas de un punto arbitrario r en el plano, w_x y w_y son el ancho de la función gaussiana en cada dimensión [22]. Sin embargo, dado que la distribución de intensidad del haz de luz generado por el láser tiene una forma circular, ambos parámetros w_x y w_y son iguales.

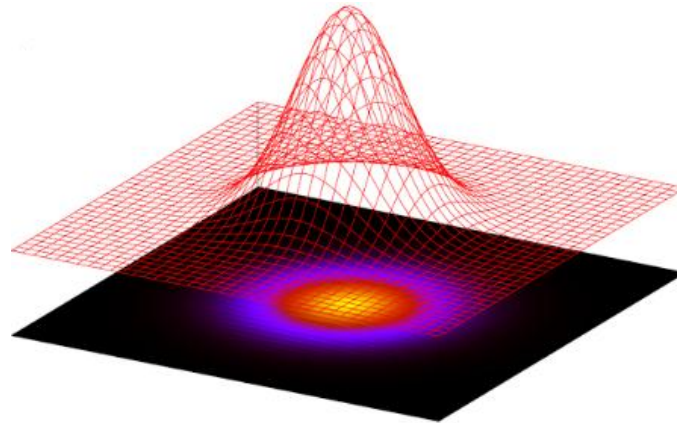


Figura 4.2: Gráfica de la intensidad del haz de luz generado por el láser.

La correlación cruzada entre dos funciones $f(r)$ y $g(r)$ esta dada por

$$f(r) \otimes g(r) = J\{F\{f(r)\} * F\{g(r)\}\}. \quad (4.0.2)$$

Donde F representa la transformada de Fourier y $f(r)$ es la componente del color rojo de la imagen adquirida.

Al calcular la correlación se resaltan los píxeles sobre los cuales se encuentra las líneas láser proyectadas, y se disminuyen los píxeles que no se asemejan a la función gaussiana lo que ayuda a eliminar el ruido introducido a la escena por la iluminación ajena al sensor, como iluminación solar, reflejo de la escena sobre otro objeto. Posteriormente es posible

alcanzar una resolución subpíxel al hacer una interpolación entre los píxeles resaltados por la correlación para encontrar el máximo de la campana de Gauss [7].

Para identificar que píxeles pertenecen a cada línea láser del patrón y obtener solo el píxel donde se encuentra el centro del haz de luz de cada línea proyectada en nuestro patrón es necesario implementar un detector de máximos locales con una distancia mínima ∂ entre los píxeles sobre los cuales pasa el haz de luz. Este ∂ se define dependiendo del tamaño y niveles de deformación de nuestra escena. En el detector de máximos locales se ordenan los valores de cada renglón de píxeles en la imagen de manera descendente y se toma el primer píxel como un máximo, después se toma el siguiente píxel en la lista ordenada y se calcula la distancia con respecto a los máximos anteriores, si la distancia es mayor al ∂ establecido este píxel se considera como un máximo, después se pasa al siguiente píxel en la lista, y se repite el proceso hasta que el número de máximos locales sea igual al número de líneas proyectadas en nuestro patrón.

El algoritmo para detectar los máximos locales identifica los puntos por los cuales pasa el láser sobre nuestra imagen toma como entrada el renglón de la imagen en el que se va a buscar, la distancia mínima que tienen que tener los láseres entre sí, el mínimo de intensidad para que sea considerado como un máximo y como resultado entrega la posición de los máximos donde se encuentran los láseres, la cual es obtenida ordenando los valores de intensidad en el rango de mayor a menor y midiendo la distancia entre la posición de los máximos que sobrepasen el nivel de intensidad mínimo establecido y la distancia mínima establecida entre los láseres, seleccionando los valores más altos que cumplan con estas condiciones, tomando en cuenta que el máximo absoluto siempre se

considera como un punto donde pasa un láser. En la Figura 4.3 se muestra el pseudocódigo del algoritmo para detectar los máximos locales y en el Anexo A.2 se puede observar el diagrama de flujo del código desarrollado.

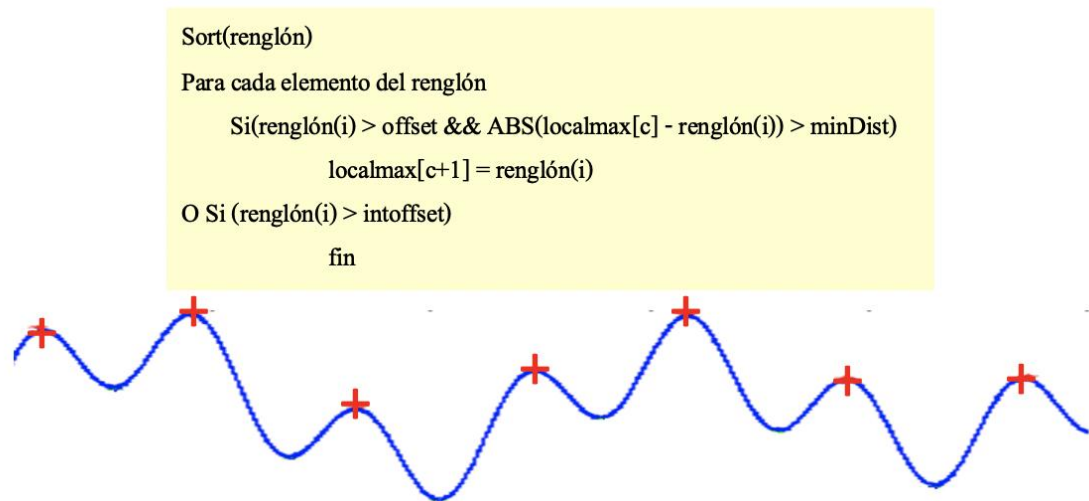


Figura 4.3: Pseudocódigo del algoritmo para detectar máximos locales.

Una vez obtenidos los pixeles donde se encuentra proyectado el patrón de luz es necesario calcular la deformación para poder obtener la profundidad o distancia a la cual se encuentra el objeto. Para calcular la deformación se realiza una comparación entre el patrón de luz como se detectaría en una superficie plana y el que se obtuvo en esta medición, lo que nos daría el cambio en el relieve.

En la parte superior de la figura 4.4 se muestra un ejemplo del proceso de escaneo de una superficie, en la parte inferior se muestra la captura de la cámara (a), el patrón original (b), y el resultado de la comparación de la muestra obtenida y la original, donde la diferencia es la deformación en la superficie obtenida.

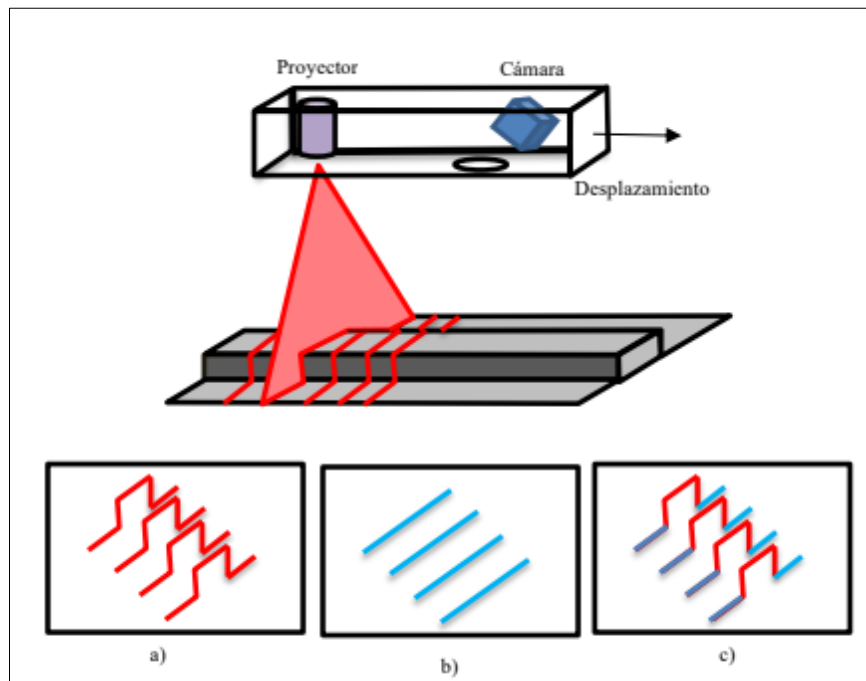


Figura 4.4: *Ejemplo del uso del escáner, en donde a) muestra los puntos que obtiene el sistema proyectados en la superficie, b) es lo que se obtendría de una superficie plana y c) compara a) con b) y la diferencia representa el relieve de la superficie escaneada .*

4.3 Transformación y registro

Para transformar los puntos adquiridos en cada imagen de tal manera que todas las nubes de puntos m y s tengan el mismo origen que la primera nube adquirida se utiliza un algoritmo iterativo de vecinos más cercanos (ICP) [4]. En el algoritmo ICP se calcula la distancia entre los puntos de la nube m con los puntos más cercanos de s después se eliminan las relaciones sobre el mismo punto seleccionando la que tenga la menor

distancia, de tal manera que sea una relación uno a uno, una vez teniendo estas relaciones se calculan los parámetros de transformación que harían coincidir los puntos en cada relación y se calcula el error de mínimos cuadrados para cada posible transformación, y se aplica la transformación del que tenga el menor error. El proceso anterior se repite hasta que el error de la transformación sea tan pequeño como el deseado o se llegue a un número máximo de iteraciones.

Como se puede ver en detalle en el Anexo A.3 nuestro algoritmo toma dos parámetros de entrada l_i y m , donde l_i es el grupo de puntos extraídos de la última toma, y m es el conjunto de puntos obtenidos en todas las tomas anteriores, y un parámetro de salida s , el cual es el resultado de la transformación de los puntos de entrada l_i de tal manera que su punto de referencia coincida con el de los puntos que se encuentran en m , los parámetros de rotación y traslación son calculados mediante el error de mínimos cuadrados.

Para el registro de los nuevos puntos simplemente se anexa la nube transformada s a la nueva nube m [21], de tal manera que:

$$m = m + s. \quad (4.0.3)$$

Una vez que m sea lo suficientemente robusta como nuestra aplicación lo requiera el algoritmo se da por terminado, de lo contrario se sigue repitiendo el ciclo y aumentando el tamaño de m .

De tal manera que el ciclo principal del sistema, así como su desarrollo completo se muestra en los diagramas de flujo del Anexo A.

En la Figura 4.5 y 4.6 se puede ver ejemplificado en un diagrama el proceso de registro construido para esta investigación. En donde la Figura 4.5 representa dos barridos o capturas de una superficie por el sistema de visión que se construyó utilizando una matriz de láser en dos posiciones diferentes de la misma escena, las cuales se pueden volver a ver en la Figura 4.6 en donde la imagen en a) se presentan las 2 mediciones tomadas de la misma superficie, pero de diferentes puntos, y en la imagen b) se presenta el resultado de la implementación del algoritmo ICP.

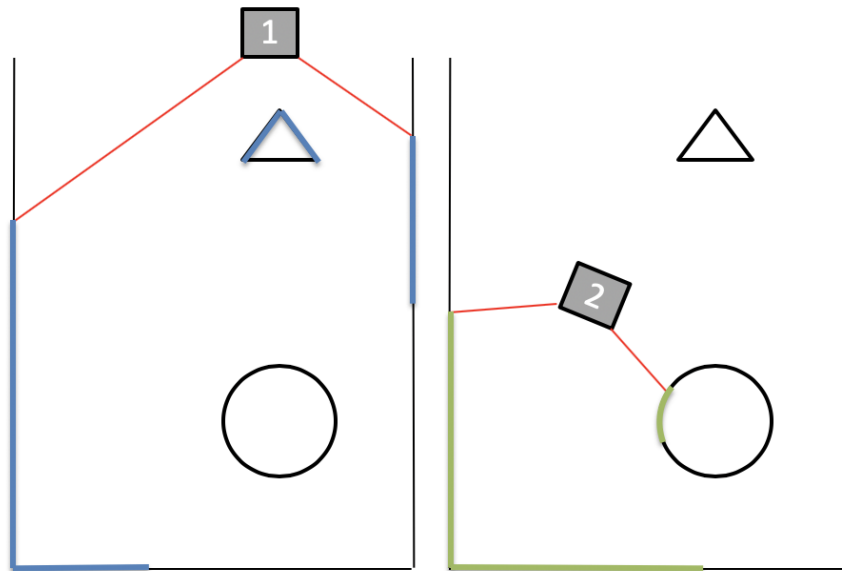


Figura 4.5: *Diagrama de representación para mostrar dos barridos del sistema de visión construido de la misma superficie en dos diferentes posiciones.*

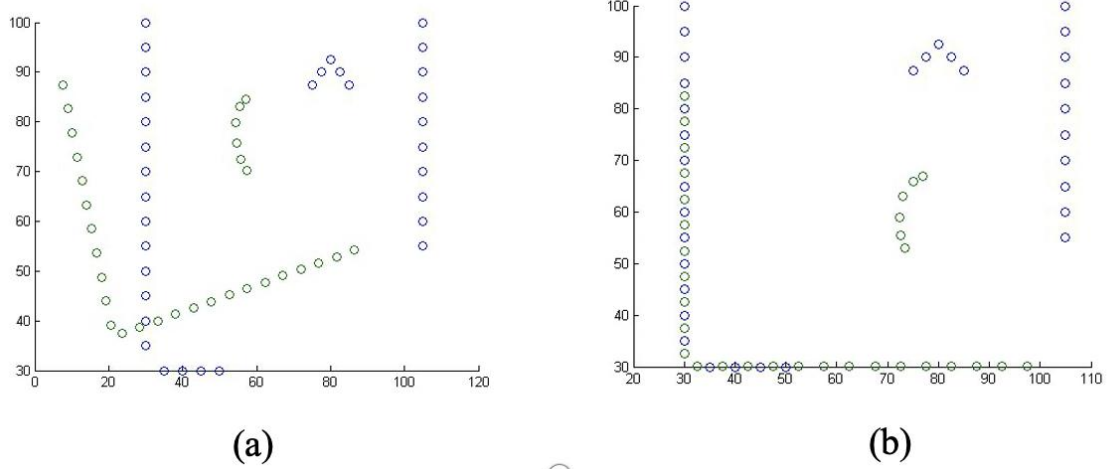


Figura 4.6: Diagrama para ejemplificar el uso del escáner construido y el resultado obtenido del algoritmo ICP. En a) se muestran los puntos obtenidos en dos tomas distintas, en donde el conjunto de puntos azules es la primera toma, el conjunto de puntos verdes es la segunda toma vista desde otro punto o ángulo; en b) es el resultado esperado, el cual es alinear los puntos obtenidos en la primera toma (azul) con los puntos de la segunda toma (verde) a manera de ampliar la información que se tiene del área u objeto.

Capítulo 5

Resultados

En la Figura 5.1 se puede observar el resultado de la proyección luz estructurada del sensor con láseres y en la Figura 5.2 se puede ver el sensor y cámara, es decir, el sistema de visión que se construyó para el desarrollo de esta investigación, así como la descripción de cada uno de sus componentes.



Figura 5.1: *Ejemplo de láseres proyectados en el objeto a escanear.*

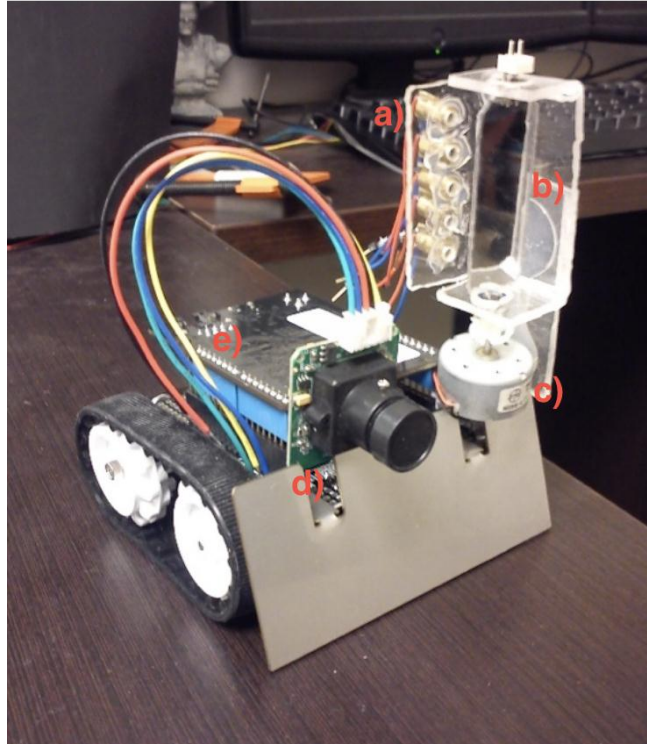


Figura 5.2: Sistema de visión construido. a) arreglo de diodos láser de 5mW con disipador; b) arreglo triangular de tres espejos; c) motor dc de 5V; d) cámara RGB STD; e) tablat de desarrollo netduino plus.

5.1 Casos de estudio.

A continuación, se presentan dos casos de estudio para mostrar el desempeño del sistema propuesto.

En la Figura 5.3 y 5.5 se muestran los objetos con el cual se realizaron pruebas y en la Figura 5.4 y 5.6 se muestra el resultado de la digitalización con el método implementado en este trabajo para ambos objetos. Como se puede apreciar en las Figuras 5.4 y 5.6 se obtuvo una nube de puntos lo suficientemente densa como para interpretar y reconstruir el objeto, en este resultado las unidades de profundidad están en pixeles, sin embargo,

estableciendo la distancia del sensor al objeto se puede hacer una triangulación y convertir las unidades de pixeles a la unidad de distancia deseada.



Figura 5.3: Objeto de prueba A para ser escaneado.

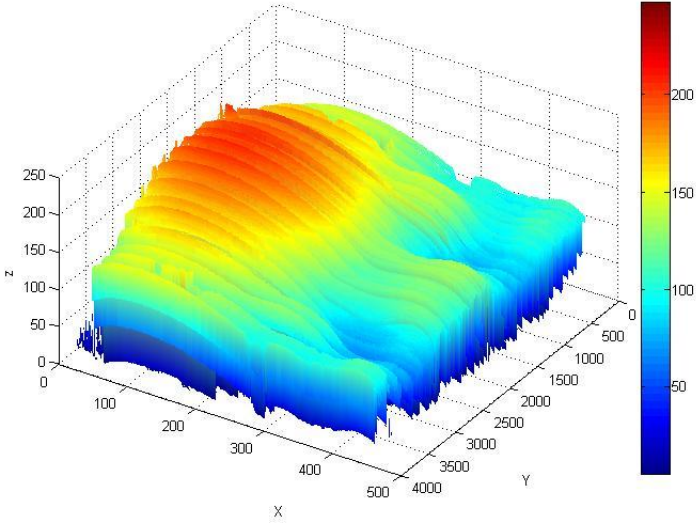


Figura 5.4: Resultados del escaneo de objeto de prueba A.



Figura 5.5: Objeto de prueba B para ser escaneado.

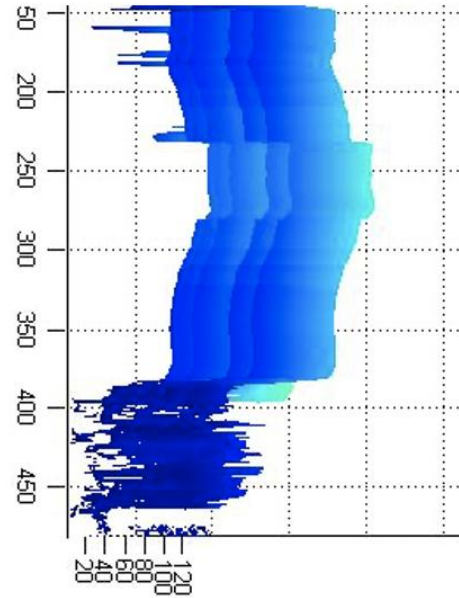
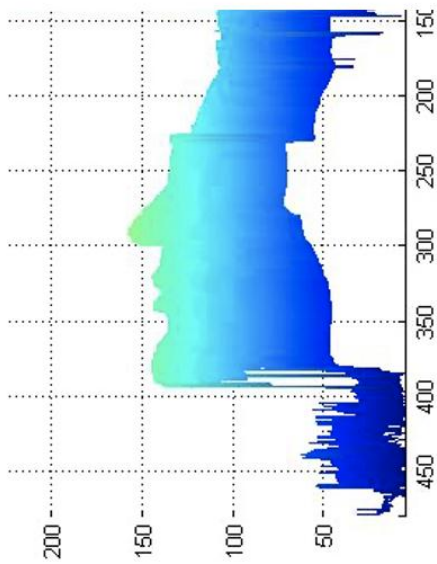


Figura 5.6: Resultados obtenidos del escaneo de objeto B, vista de 90° y 45°.

Capítulo 6

Conclusiones y trabajo futuro.

6.1 Conclusiones

Mediante el uso de un patrón de luz estructurada sobre una superficie o un objeto es posible obtener la suficiente información para poder digitalizar la superficie u objeto en tiempo real con una secuencia de fotos.

Con este trabajo observamos que es posible crear un sistema digitalizador de entornos el cual no dependa de un punto de referencia fijo en el espacio, ni sensores que lo calculen, acelerómetros, giroscopios, entre otros, sino que sea capaz de establecerlo a través de software, lo que nos da como principal ventaja la eliminación de un volumen máximo para ser digitalizado puesto que ya no es necesario manipular el entorno para controlar el punto de referencia, sin embargo para que este método sea viable y posible en tiempo real es necesario utilizar una cámara de alta velocidad en nuestro sensor, puesto que esto ayuda a que el sistema requiera de menos ciclos para poder ajustar los datos obtenidos.

6.2 Trabajo futuro

Como trabajo futuro queda el desarrollo de un sistema de navegación que incluya la toma de decisiones en base al modelo tridimensional generado por la técnica de *Flying Triangulation*.

Bibliografía

- [1] E.R. Davies, *Computer and Machine Vision. Principle Algorithms applications*, 4th ed. Academic Press, 2012. ISBN: 9780123869913.
- [2] S. Prince, *Computer Vision: Models, Learning and Inference*, Cambridge University Press, 2012. DOI: 10.1017/CBO9780511996504.
- [3] S. Ettl, O. Arold, P. Vogt, O. Hybl, Z. Yang, W. Xie and G. Häusler, "Flying Triangulation" – a new optical 3D sensor enabling the acquisition of surfaces by freehand motion”, *APPLIED OPTICS*, Vol. 51, No. 2, 2012.
- [4] F. Willomitzer, S. Ettl, O. Arold and G. Häusler, *Flying Triangulation - A Motion-Robust Optical 3D Sensor For The Real-Time Shape acquisition Of Complex Objects*, *AIP Conf. Proc.* 1537, pp. 19-26, 2013. DOI: 10.1063/1.4809687.
- [5] D. Forsyth and J. Ponce, *Computer Vision. A modern Approach*, 2nd ed. Prentice Hall, 2003. ISBN: 0131911937, 9780131911932.
- [6] S. Ettl, O. Arold, Z. Yang, and G. Häusler, *Flying triangulation—an optical 3D sensor for the motion-robust acquisition of complex objects*, *Applied optics*, Vol. 51, No. 2, 2012. DOI: 10.1364/AO.51.000281.
- [7] M. Izquierdo, M. Sanchez, A. Ibañez and L. Ullate, *Sub-pixel measurement of 3D surfaces by laser scanning*, *Sensors and Actuators A: Physical*, Elsevier, Vol. 76, issues 1-3, pp. 1-8, 1999. DOI: 10.1016/S0924-4247(98)00283-0.
- [8] S. Ettl, F. Willomitzer, O. Arold, G. Häusler, *Robust pattern indexing methods for “Flying Triangulation”*, *Conference: Annual Meeting of DGaO*, pp. 24-26, 2012.
- [9] R. Szeliski, *Computer Vision. Algorithms and Application*, Springer-Verlag London, 2011. DOI: 10.1007/978-1-84882-935-0.
- [10] B. Rogers, and P. Howard, *Binocular Vision and Stereopsis*, Oxford University Press, 1995.
- [11] G. Pajares, J. de la Cruz, *Ejercicios resueltos de Visión por Computador*, Alfaomega Grupo Editor, pp. 239-250, 2008 .ISBN: 978-970-15-1372-9.

- [12] G. Pajares, J. de la Cruz, *Ejercicios resueltos de Visión por Computador*, Alfaomega Grupo Editor, pp. 47-48, 2008 .ISBN: 978-970-15-1372-9.
- [13] E. Trucco and A. Verri, *Introductory Techniques for 3-D Computer Vision*, Prentice Hall, 1998. ISBN: 978-0132611084.
- [14] BC Ministry of Healthy Living and Sport, *Provincial Vision Screening Training Manual*, de British Columbia, 2010. <https://www2.gov.bc.ca/gov/content/health/conducting-health-research-evaluation/health-and-human-services-library>.
- [15] Design World Staff, *IFM Efector Introduces New 3D Image Sensor*, Design World, 2009. <https://www.designworldonline.com/ifm-efector-introduces-new-3d-image-sensor/>.
- [16] C. Wöhler, *3D computer vision Efficient Methods and Applications*, Springer-Verlag London, 2013. DOI: 10.1007/978-1-4471-4150-1.
- [17] W. Jüpner and W. Osten, *Automatic Processing of Fringe Patterns*, 3rd ed. Bremen, Akademie Verlag Serie in Optical Metrology, 1997.
- [18] F. Willomitzer, S. Ettl, C. Faber, and G. Häusler, *Flying Triangulation – Towards the 3D Movie Camera*, Springer-Verlag Berlin, 2014. DOI: 10.1007/978-3-642-36359-7_167.
- [19] S. Ettl, F. Schiffers, F. Willomitzer, S. Rampp, C. Strauss and G. Häusler, *Flying Triangulation: a tool for automated medical diagnosis*, Annual Meeting of DgaO, 2014.
- [20] S. Ettl, O. Arold, P. Vogt, O. Hybl, Z. Yang, W. Xie, G. Häusler , “Flying Triangulation”: A motion-robust optical 3D sensor principle, Springer Berlin, pp. 1-4, 2009. DOI: 978-3-642-03051-2.
- [21] P. J. Besl and N. D. McKay, *A method for registration of 3-D shapes*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 14, No. 2, pp. 239-256, 1992. DOI: 10.1109/34.121791.
- [22] G. Rodríguez-Morales, V. Guzmán-Ramos, H. Martínez, P. López and A. Garza, *Integración de Sistema de Digitalización 3D Usando Matlab*, 10º Congreso Nacional de Mecatrónica, Jalisco, noviembre 2011.

Anexo A

Diagramas de flujo del proceso del algoritmo *Flying Triangulation*

desarrollado para el sistema de visión construido.

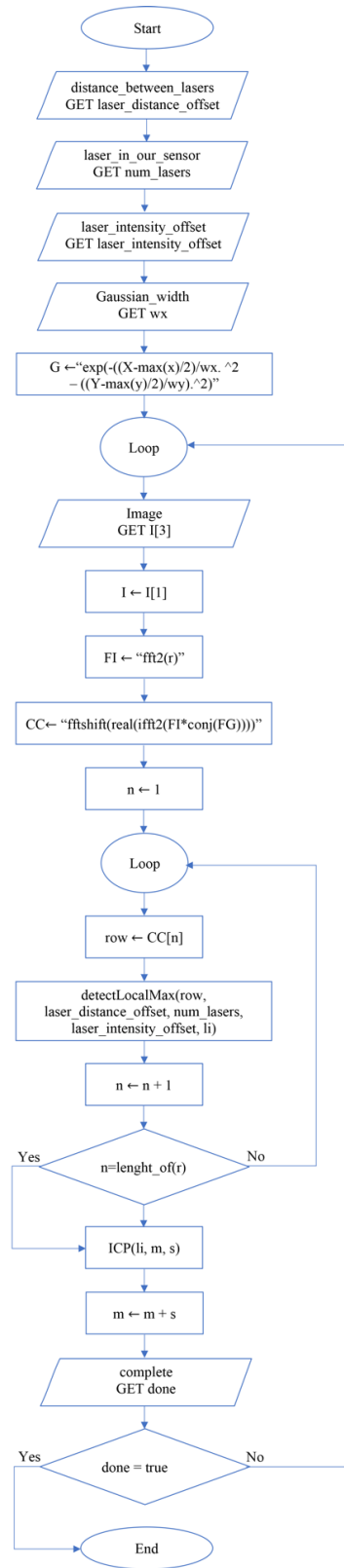


Figura A.1: Diagrama de flujo del algoritmo para el ciclo principal del sistema para el procesamiento de las imágenes del sensor construido

detectLocalMax

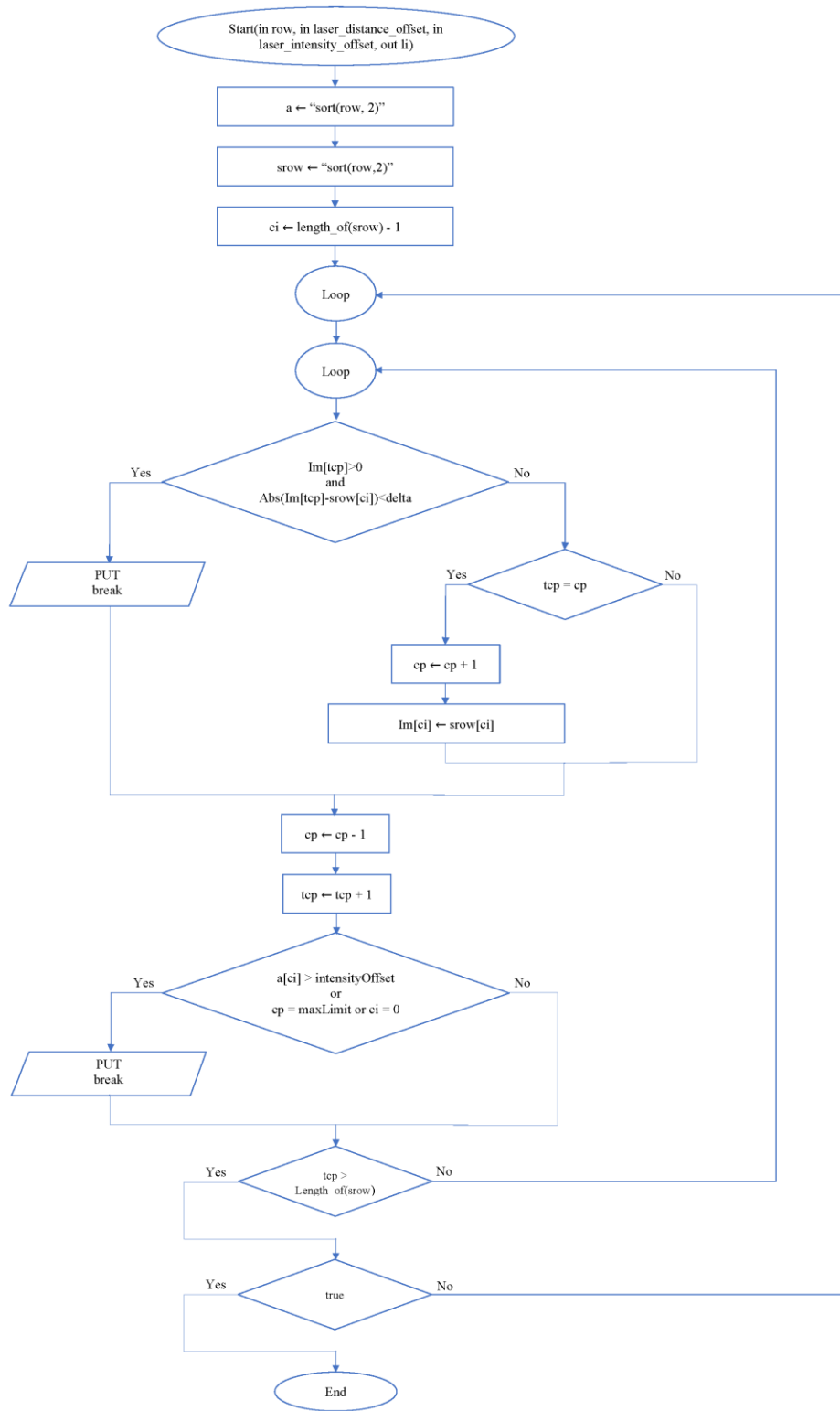


Figura A.2: Diagrama de flujo del algoritmo para detectar máximos locales.

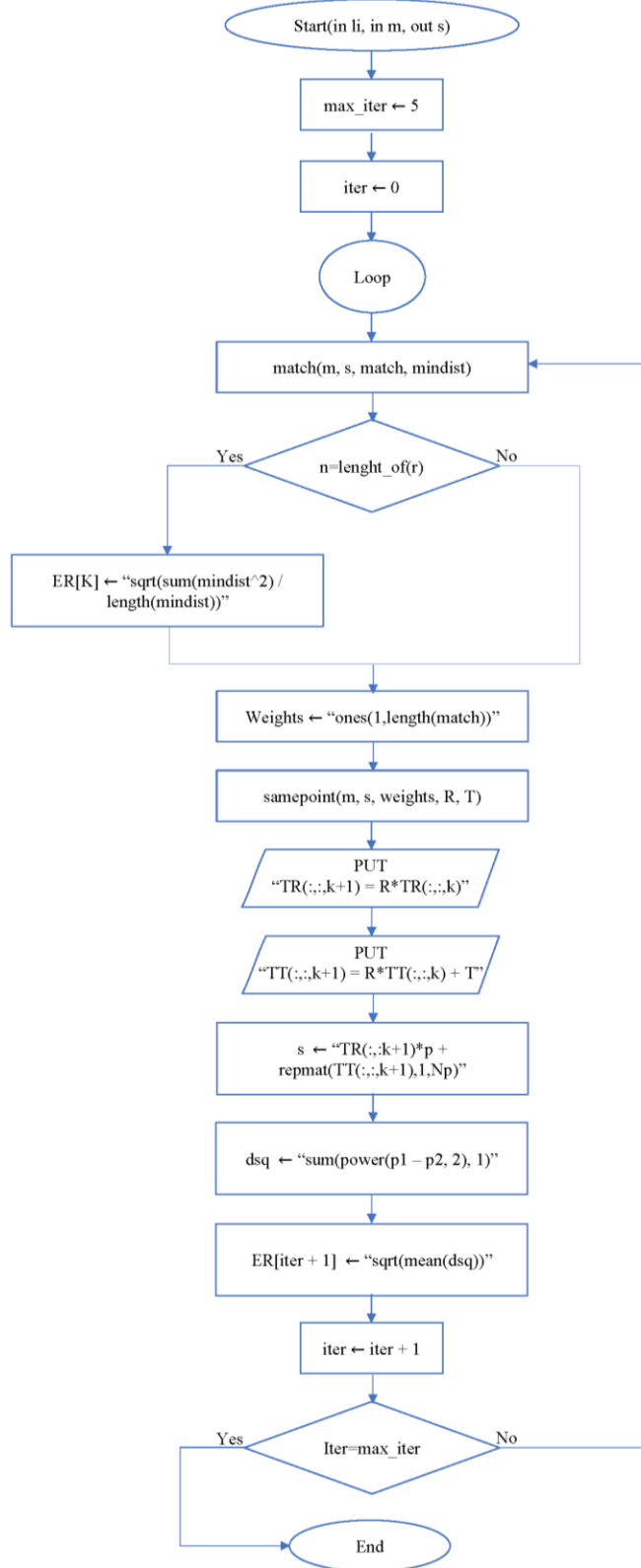


Figura A.3: Diagrama de flujo del algoritmo ICP desarrollado.

Match

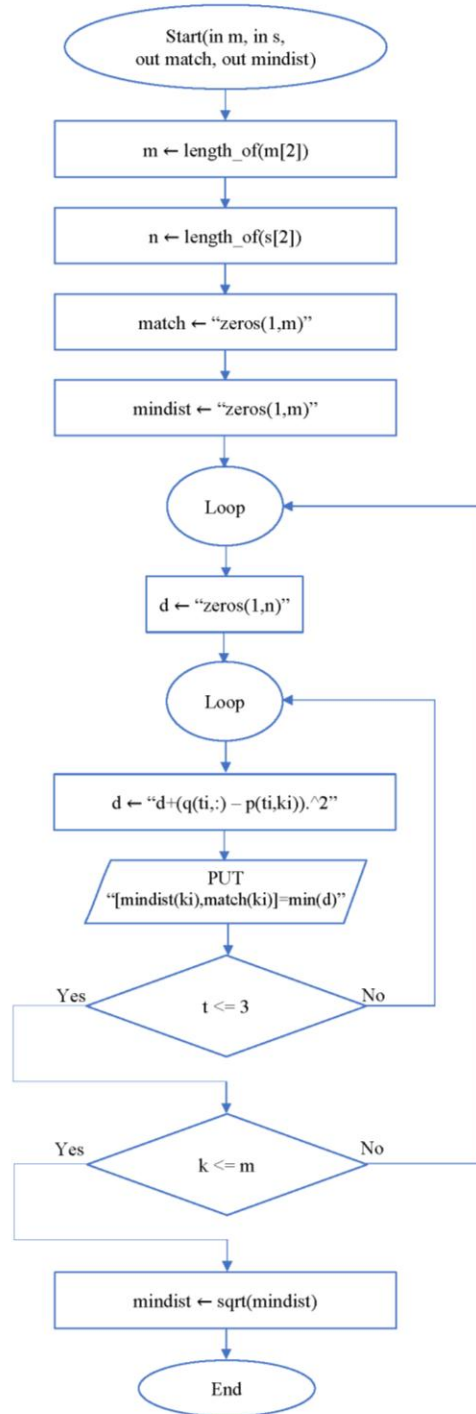


Figura A.4: Diagrama de flujo de la función match utilizada en el algoritmo ICP desarrollado.

Samepoint

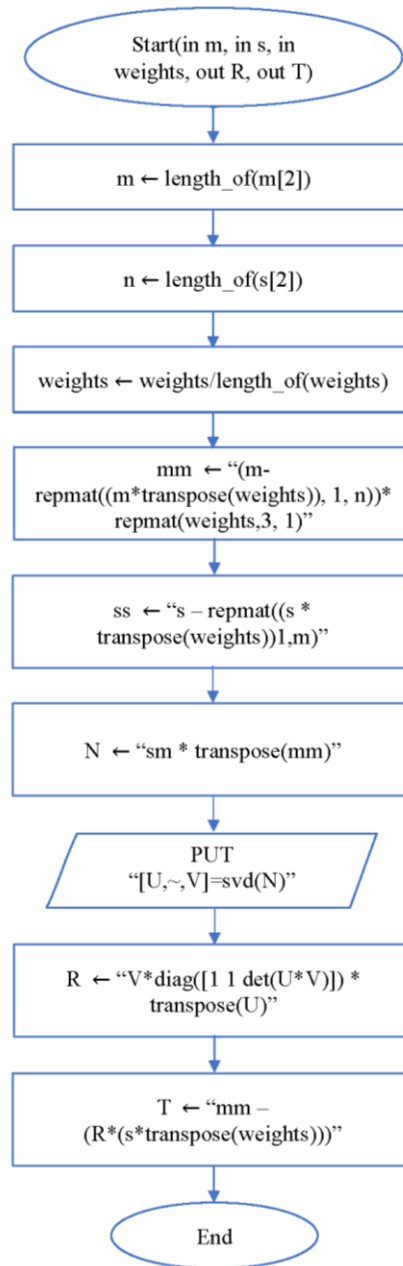


Figura A.5: Diagrama de flujo de la función `samepoint` utilizada en el algoritmo ICP desarrollado.