

**INSTITUTO DE INVESTIGACIÓN EN CIENCIAS BÁSICAS Y APLICADAS  
(IICBA)**

**CENTRO DE INVESTIGACIÓN EN INGENIERÍA Y CIENCIAS APLICADAS  
(CIICAp)**

**TÍTULO DE LA TESIS**

**DISPOSITIVO DE RECONSTRUCCIÓN TRIDIMENSIONAL  
REMOTO MEDIANTE LUZ ESTRUCTURADA Y VISIÓN  
ESTEREOSCÓPICA PARA AMBIENTES NO  
CONTROLADOS.**

**TESIS PARA OBTENER EL GRADO DE:**

**DOCTORADO EN INGENIERIA Y CIENCIAS APLICADAS CON OPCION  
TERMINAL EN: TECNOLOGIA ELECTRICA**

**PRESENTA:**

**ROBERTO ALAN BELTRAN VARGAS**

**ASESOR:**

**DR J JESÚS ESCOBEDO ALATORRE**

**CUERNAVACA, MORELOS**

**MARZO, 2020**



## AGREDECIMIENTOS

Agradezco sinceramente a todas las personas que hicieron posible la realización de este trabajo de investigación.

Al Dr. J Jesús Escobedo Alatorre, por haberme aceptado como asesorado y por haberme brindado todos los recursos que en su momento se necesitaron en ésta investigación, gracias por su confianza y amistad, por compartirme sus conocimientos y apoyo en todo momento.

Al Dr. Álvaro Zamudio Lara, por compartir conmigo sus conocimientos, por todo su apoyo y brindarme la confianza de trabajar con un gran equipo multidisciplinario.

Al Dr. José Antonio Marbán, por mostrarme el sentido del trabajo en equipo, a compartir el conocimiento adquirido siempre de la mejor forma, gracias por su amistad y confianza.

Al Dr. Omar Palillero, por enseñarme que el trabajo de investigación no solo es en nuestro tema de dominio, por lo contrario, se requiere de leer y aprender de otros temas para generar soluciones desde otra perspectiva.

A la Dra. Margarita Tecpoyotl, por motivarnos a emprender, a soñar y a trabajar con el doble de esfuerzo para alcanzar nuestras metas.

También quiero agradecer a Jorge Arturo Sandoval quien, sin sus conocimientos y habilidades, no hubiese sido posible la implementación de la primera versión del prototipo del dispositivo de reconstrucción tridimensional.

Agradezco al Centro de Investigación en Ingeniería y Ciencias Aplicadas, así como al personal académico y administrativo por las facilidades otorgadas.

Agradecimientos al Consejo Nacional de Ciencia y Tecnología (CONACYT, México) el apoyo económico brindado durante el desarrollo de este trabajo de Doctorado.

A mi esposa Soraya y mis padres Roberto y Norma por su valioso apoyo.

Y de manera especial, agradezco y dedico la culminación de este trabajo a mis hijas Isabella y Daniela, quienes fueron siempre la motivación más grande para superar todas las adversidades.

# INDICE GENERAL

## Contenido

AGREDECIMIENTOS .....	i
INDICE GENERAL.....	ii
INDICE DE FIGURAS .....	iv
RESUMEN.....	vi
ABSTRACT .....	viii
MOTIVACIÓN .....	x
OBJETIVO.....	xi
Objetivos específicos:.....	xi
JUSTIFICACIÓN.....	xii
ALCANCE.....	xiii
LIMITACIONES .....	xiii
CAPÍTULO 1 INTRODUCCIÓN Y ANTECEDENTES .....	1
1. Reconstrucción tridimensional a partir de imágenes en ambientes no controlados.....	1
CAPÍTULO 2 MARCO TEÓRICO .....	6
2.1 Visión por computadora.....	6
2.2 Formación de la imagen.....	9
2.3 Modelo de cámara pinhole.....	10
2.4 Distorsiones asociadas más significativas.....	12
2.5 Calibración de cámaras estéreo.....	17
2.6 Reconstrucción pasiva y activa.....	22
2.6.1 Técnicas pasivas .....	23
2.6.1.1 Visión estereoscópica .....	23
2.6.1.2 Estructura desde movimiento (Structure from motion) .....	24
2.6.1.3 Forma desde sombreado (shape from shading) .....	24
2.6.1.4 Fotometría estéreo .....	25
2.6.1.5 Geometría epipolar.....	25
2.6.2 Técnicas activas.....	27
2.6.2.1 Luz estructurada .....	27
2.7 Técnicas de reconstrucción de luz estructurada .....	28
2.7.1.1 Cambio de fase (phase shifting).....	28
2.7.1.2 Código binario.....	30
2.8 Triangulación.....	33
2.9 Monitoreo y control en internet .....	36
2.10 Hardware en pequeña escala.....	40

2.11	Software utilizado .....	42
2.12	Aplicaciones de la reconstrucción tridimensional a partir de imágenes con luz estructurada.....	43
2.12.1	Medicina .....	44
2.12.2	Industria .....	45
2.12.3	Robótica .....	47
2.12.4	Reconstrucción de edificios y reliquias.....	49
2.12.5	Entretenimiento.....	51
CAPÍTULO 3 DESARROLLO TEÓRICO/EXPERIMENTAL .....		52
3.1	Componentes del sistema.....	52
3.1.1	Módulos del sistema .....	52
3.1.2	Comunicación escáner 3d - usuario .....	54
3.1.3	Configuración modo local.....	55
3.1.4	Configuración modo remoto .....	56
3.1.5	Arquitectura de software del sistema.....	57
3.1.6	Componentes de hardware.....	58
3.2	Geometría del arreglo óptico .....	60
3.3	Ajuste automático de parámetros de cámaras .....	63
3.4	Calibración de cámaras .....	65
3.5	Proyección y adquisición de patrones.....	66
3.6	Codificación de patrones.....	67
3.7	Segmentación.....	68
3.8	Decodificación de patrones.....	69
3.9	Método de triangulación .....	69
CAPÍTULO 4 RESULTADOS Y DISCUSIÓN .....		70
4.1	Pruebas de reconstrucción con visión estéreo y luz estructurada con el patrón de código binario. .	70
CAPÍTULO 5 CONCLUSIONES, RECOMENDACIONES Y TRABAJO FUTURO.....		75
6.1	Conclusiones.....	75
6.2	Recomendaciones .....	76
6.3	Trabajo futuro .....	78
BIBLIOGRAFÍA.....		79

## INDICE DE FIGURAS

FIGURA 1 EFECTO DE LA LUZ DEL SOL DIRECTA. ....	2
FIGURA 2 ARREGLO ÓPTICO PARA CONCENTRACIÓN DE LUZ. ....	3
FIGURA 3 TÉCNICA DE CONCENTRACIÓN DE LUZ. ....	4
FIGURA 4 PREPROCESAMIENTO DE CADA PROYECCIÓN DE PATRON. ....	5
FIGURA 5 TRANSFORMACIÓN DE LA CÁMARA (ROBERTS, 1965). ....	7
FIGURA 6 RECONOCIMIENTO DE PLACAS DE CARROS. ....	7
FIGURA 7 ESQUEMA DE LAS RELACIONES ENTRE LA VISIÓN POR ORDENADOR Y OTRAS DISCIPLINAS (IMAGEN DE DOMINIO PÚBLICO).8	
FIGURA 8. SENSOR DE UNA CÁMARA FOTOGRÁFICA. ....	10
FIGURA 9 MODELO DE CÁMARA PINHOLE. ....	12
FIGURA 10 DISTORSIÓN RADIAL DEBIDO ALLENTE DE LA CÁMARA. ....	13
FIGURA 11 DISTORSIÓN RADIAL (ZELINSKY, 2009). ....	14
FIGURA 12 DEFECTOS DE DISTORSIÓN TANGENCIAL EN CÁMARAS. ....	15
FIGURA 13 RESULTADOS DE DISTORSIÓN TANGENCIAL (ZELINSKY, 2009). ....	15
FIGURA 14 ESQUEMA DEL MODELO PINHOLE. ....	18
FIGURA 15 CALIBRACIÓN CON TABLERO DE AJEDREZ. ....	20
FIGURA 16 ROTACIÓN Y TRASLACIÓN ENTRE 2 CÁMARAS. ....	21
FIGURA 17 RECONSTRUCCIÓN ACTIVA. A) UN PROYECTOR Y UNA CÁMARA. B) UN PROYECTOR Y VISIÓN ESTEREOSCÓPICA .....22	
FIGURA 18. VISIÓN ESTEREOSCÓPICA. ....	23
FIGURA 19 GEOMETRÍA EPIPOLAR. ....	26
FIGURA 20 TÉCNICA DE LUZ ESTRUCTURADA. ....	27
FIGURA 21 PHASE SHIFTING CON 3 PATRONES. ....	28
FIGURA 22 IMAGEN CON INFORMACIÓN DE PATRONES Y TOPOGRAFÍA DEL PLANO (A) Y OBJETO (B). ....	29
FIGURA 23 IMAGEN CON FASE DESENVUELTA O CONTINUA EN PLANO DE REFERENCIA Y OBJETO. ....	30
FIGURA 24 PATRONES BINARIOS HORIZONTALES Y VERTICALES. ....	31
FIGURA 25 PÍXELES EN EL PLANO DE LA IMAGEN CON EL PATRON BINARIO PROYECTADO. ....	31
FIGURA 26 CODIFICACIÓN BINARIA DE UN PIXEL P EN EL PLANO DE LA IMAGEN. ....	32
FIGURA 27 UBICACION DE PIXEL P DE PROYECCIÓN. ....	32
FIGURA 28 CÁLCULO DEL PUNTO DE INTERSECCIÓN. ....	34
FIGURA 29 ESQUEMA DEL MÉTODO DE TRIANGULACIÓN POR LINEAS. ....	36
FIGURA 30 ARQUITECTURA IOT. ....	37
FIGURA 31 APLICACIÓN DE IOT EN LA INDUSTRIA. ....	39
FIGURA 32 APLICACIÓN DE IOT EN LA INDUSTRIA. ....	39
FIGURA 33 LEY DE MOORE, INCREMENTO DEL NÚMERO DE TRANSISTORES CADA 2 AÑOS. ....	40
FIGURA 34 COMPARACIÓN DE VERSIONES DE RASPBERRY EN DIFERENTES PRUEBAS DE BENCHMARKING. ....	42
FIGURA 35 MEDIT 1500 ESCANER 3D PARA ODONTOLOGÍA. ....	44
FIGURA 36 INGENIERÍA INVERSA PARA LA FABRICACIÓN DE SILLAS DE RUEDAS A MEDIDA. ....	45
FIGURA 37 INGENIERÍA INVERSA EN LA INDUSTRIA (STYL&TECH). ....	46
FIGURA 38 DETECCIÓN DE ERRORES EN LINEAS DE PRODUCCIÓN (KEYENCE, 2020). ....	46
FIGURA 39 MANIPULACIÓN DE MATERIALES AUTOMATIZADA MEDIANTE UN SISTEMA DE VISIÓN (FANUC). ....	47
FIGURA 40 GENERACION DE UN MAPA 3D DE UNA TUBERIA. ....	48
FIGURA 41 RECONSTRUCCIÓN 3D DE EDIFICIOS. ....	49
FIGURA 42 RECONSTRUCCIÓN 3D DE RELIQUIA (FARO). ....	49
FIGURA 43 MUSEO 3D EN LINEA DE RELIQUIAS DE VIETNAM. ....	50
FIGURA 44 DETALLE DE PIEZA HISTORICA DE MUSEO 3D. ....	50
FIGURA 45 RECONSTRUCCION 3D DE PERSONAS E IMPRESION 3D. ....	51
FIGURA 46 DIAGRAMA DE BLOQUES DEL FUNCIONAMIENTO GENERAL DEL ESCÁNER 3D. ....	52
FIGURA 47 PROCESO DE RECONSTRUNCCÓN 3D. ....	53
FIGURA 48 COMUNICACIÓN EN ENTORNO LOCAL Y REMOTO. ....	54

FIGURA 49 CONFIGURACIÓN LOCAL DEL ESCANER 3D. ....	55
FIGURA 50 CONFIGURACION DE ACCESO REMOTO. ....	56
FIGURA 51 ARQUITECTURA DEL SOFTWARE. ....	57
FIGURA 52 COMPONENTES DE HARDWARE DEL ESCÁNER 3D. ....	58
FIGURA 53 PRUEBAS PARA OBTENER GEOMETRÍA DE VISIÓN ESTÉREO. ....	61
FIGURA 54 GEOMETRÍA ESTÉREO EN ARREGLO DE CÁMARAS. ....	62
FIGURA 55 AJUSTE DE PARÁMETROS DE CÁMARAS. ....	63
FIGURA 56 ALGORITMO DE AUTO AJUSTE DE CÁMARAS. ....	64
FIGURA 57 FUNCIÓN DE CALIBRACIÓN ESTÉREO. ....	66
FIGURA 58 ALGORITMO DE ADQUISICIÓN DE PATRONES. ....	66
FIGURA 59 PROYECCIÓN Y CAPTURA DE PATRONES SOBRE OBJETO A RECONSTRUIR. ....	67
FIGURA 60 CODIFICACIÓN DE PATRONES CAPTURADOS. ....	68
FIGURA 61 SEGMENTACIÓN PARA SELECCIÓN DE REGIÓN DE INTERÉS. ....	68
FIGURA 62 MÉTODO DE TRINGULACIÓN PARA OBTENER Z. ....	69
FIGURA 63 NUBE DE PUNTOS DE UN PROTOTIPO ÁLABE HECHO CON FIBRA DE VIDRIO. ....	70
FIGURA 64 NUBE DE PUNTOS DE UN PROTOTIPO ÁLABE HECHO CON CASRTÓN. ....	71

## RESUMEN

La reconstrucción tridimensional a partir de imágenes tomadas desde cámaras fotográficas, comprende todas las técnicas ópticas para recuperar información de profundidad que se pierde en dichas imágenes. Su aplicación es muy extensa en un sinnúmero de áreas como la médica donde, por ejemplo, se pueden generar reconstrucciones tridimensionales a partir de imágenes TAC (Tomografía Axial Computarizada) y obtener un modelo 3d del órgano o parte del cuerpo del paciente que ayude a los médicos a comprender mejor una enfermedad y aumentar las posibilidades de curar al paciente. Otra área es la industrial, donde se aprovecha las ventajas de estas técnicas para exploración superficial de fallas y defectos en tuberías para el transporte de fluidos, defectos en superficies de objetos en línea de producción, con el objetivo de cuantificar daños o mejorar procesos. Otra de las áreas donde también se toma ventaja de dichas técnicas, es en la robótica, dotando a robots con una visión tridimensional con varios fines, un ejemplo, un robot que puede ir generando un mapa tridimensional de la ruta que va tomando, muy útil para reconocimiento de lugares donde es difícil el acceso o se expone la vida de una persona. También se puede encontrar la aplicación en video juegos, el ejemplo más conocido es el Kinect de Microsoft en sus versiones 1 y 2, donde se aprovecha la adquisición de la profundidad para detectar una o varias personas para jugar con gestos que son reconocibles sin un control físico. Cabe mencionar que el Kinect también ha sido aprovechado en temas de investigación científica, por ejemplo, el reconocimiento de gestos del lenguaje de señas. Así, se pueden citar muchas aplicaciones de tales técnicas.

Un problema en común con todas las técnicas de reconstrucción, es que necesitan de un ambiente controlado para aumentar el porcentaje de recuperación de profundidad de los píxeles de las imágenes. Para ambientes no controlados, donde se tiene la luz del sol iluminando al objeto a reconstruir en diferentes horarios, se han utilizado diversos dispositivos modificados que aumentan o concentran la luz



del proyector, donde se han conseguido resultados aceptables. Sin embargo, no es suficiente puesto que la energía emitida por el sol supera en 2:5 la energía de los proyectores. Otro problema que se aborda, es el de un monitoreo y control a distancia, ya que teniendo que reconstruir un objeto en horarios con temperaturas elevadas, es necesario controlar a distancia el dispositivo de reconstrucción.

La presente tesis doctoral, se centra en el desarrollo de un prototipo de un dispositivo de reconstrucción tridimensional basado en técnicas de luz estructurada que consiste en un arreglo óptico de dos cámaras, un proyector de imágenes y el desarrollo de algoritmos lógicos y matemáticos que, en conjunto, obtienen datos de profundidad de imágenes en dos dimensiones de objetos en ambientes hostiles y no controlados que pondrían en riesgo al usuario. Se desarrolla una interfaz de programación de aplicaciones abierta para que cualquier software se conecte y obtenga las reconstrucciones. El problema de reconstrucciones a la luz del sol, se aborda con la creación de un algoritmo inteligente que ajusta automáticamente los parámetros de las cámaras y ajusta las imágenes tomadas mediante un procesamiento de digital, lo que genera información que aumenta el desempeño de las reconstrucciones en ambientes no controlados. La arquitectura del sistema permite un acceso remoto, inalámbrico y en tiempo real desde cualquier dispositivo con comunicación WIFI a través del estándar IEEE 802.11x, donde el resultado final es un archivo con un objeto de nube de puntos que representa una reconstrucción tridimensional del objeto y datos que permiten realizar mediciones en las coordenadas x, y, z.

## ABSTRACT

Three-dimensional reconstruction from images taken from cameras, includes all optical techniques to recover depth information that is lost in these images. Its application is very extensive in countless areas such as medical where, for example, three-dimensional reconstructions can be generated from images CAT (Computed axial tomography) and obtain a 3d model of the organ or part of the patient's body that helps doctors better understand a disease and increase the chances of curing the patient. Another area is the industrial one, where the advantages of these techniques are explored for superficial exploration of faults and defects in pipes for the transport of fluids, defects in surfaces of objects in the production line, in order to quantify damages or improve processes. Another area where these techniques are also taken advantage of, is in robotics, providing robots with a three-dimensional vision for several purposes, an example, a robot that can generate a three-dimensional map of the route it is taking, very useful for recognition of places where access is difficult or a person's life is exposed. You can also find the application in video games, the best known example is the Microsoft Kinect in versions 1 and 2, where the acquisition of depth is used to detect one or more people to play with gestures that are recognizable without a control physical. It should be mentioned that the Kinect has also been used in scientific research topics, for example, the recognition of sign language gestures. Thus, many applications of such techniques can be cited.

A problem in common with all reconstruction techniques is that they need a controlled environment to increase the percentage of depth recovery of the image pixels. For uncontrolled environments, where there is sunlight illuminating the object to be rebuilt at different times, various modified devices have been used that increase or concentrate the projector's light, where acceptable results have been achieved. However, it is not enough since the energy emitted by the sun exceeds

the power of the projectors by 2: 5. Another problem that is addressed is that of remote monitoring and control, since having to reconstruct an object at times with high temperatures, it is necessary to remotely control the reconstruction device.

The present doctoral thesis focuses on the development of a prototype of a three-dimensional reconstruction device based on structured light techniques consisting of an optical arrangement of two cameras, an image projector and the development of logical and mathematical algorithms that, in together, they obtain depth data of images in two dimensions of objects in hostile and uncontrolled environments that would put the user at risk. An open application programming interface is developed for any software to connect and get reconstructions. The problem of reconstructions in sunlight, is addressed with the creation of an intelligent algorithm that automatically adjusts the parameters of the cameras and adjusts the images taken by digital processing, which generates information that increases the performance of the reconstructions in uncontrolled environments. The system architecture allows remote, wireless and real-time access from any device with WIFI communication through the IEEE 802.11x standard, where the end result is a file with a point cloud object that represents a three-dimensional reconstruction of the object and data that allow measurements at x, y, z coordinates.

## MOTIVACIÓN

La reconstrucción tridimensional se encuentra en un punto en donde las técnicas ofrecen resultados con gran calidad. Se aplica cada técnica dependiendo el tipo de objeto y el ambiente en el cual se va a reconstruir. El campo de aplicaciones es extenso ya que la recuperación de la profundidad nos permite comprender como son o cómo funcionan los objetos que son reconstruidos de una forma diferente. El desarrollo de equipos de cómputo cada vez más pequeños y con poder de procesamiento hace que incluso se ejecuten algoritmos de reconstrucción en tiempo real, permitiendo hacer reconstrucciones de hasta 30 fpm los protocolos y estándares de comunicación, nos permiten ver y controlar a distancia estas reconstrucciones y nos hace pensar en la infinita variedad de aplicaciones que podemos aplicar estas técnicas.

Con lo anterior, resulta de gran interés mostrar el potencial que tienen las diferentes técnicas de reconstrucción y en este caso la técnica de luz estructurada con visión estereoscópica, desarrollando un dispositivo que es capaz de reconstruir en ambientes no controlados u hostiles, con hardware de poder de procesamiento moderadamente bajo y con componentes que son fácilmente asequibles para propósitos de investigación.

## OBJETIVO

Se propone el desarrollo e implementación de un dispositivo con algoritmos para la reconstrucción tridimensional de objetos en ambientes no controlados, remoto y en tiempo real que permite el control a distancia y el acceso libre a las funciones que obtienen imágenes y generan los objetos reconstruidos. En una arquitectura IoT (Internet of things, por sus siglas en inglés), que permite el acceso a sensores a través de internet sin la intervención de personas.

Objetivos específicos:

- Codificar un algoritmo de calibración para obtener los parámetros intrínsecos y extrínsecos del arreglo de cámaras.
- Desarrollar e implementar los algoritmos de ajuste automático de parámetros de cámaras, procesamiento de imágenes para aumentar el porcentaje de puntos reconstruidos, codificación y decodificación de patrones, proceso de triangulación para obtener información de profundidad de cada pixel y proceso de generación de nube de puntos.
- Desarrollar una interfaz de programación de aplicaciones abierta para convertir al dispositivo en un sensor que pueda ser aprovechado por otras aplicaciones.

## JUSTIFICACIÓN

El número de aplicaciones en diversas áreas de los dispositivos de reconstrucción tridimensional va en aumento, por ejemplo: En ingeniería, en la modificación de productos, ingeniería inversa, medición sin contacto de prácticamente cualquier objeto; en la medicina, tomografías axiales computarizadas en 3d, para la creación de prótesis; en la arqueología, para el estudio de piezas muy frágiles; aplicaciones para mejora de ergonomía, diseño de calzado a medida, creación de modelos en 3d para tiendas en línea o entrenamiento. Cual sea la aplicación, es importante para los diseñadores de estos dispositivos entender la teoría que hay detrás, y también tener las herramientas para crear nuevas aplicaciones que ayuden a resolver problemas o entenderlos mejor.

Con la importancia que representa esta herramienta, es necesario la creación de dispositivos que puedan ser asequibles a costos menores y puedan servir de ayuda a poblaciones con menos recursos.

México es uno de los países con mayor número de pobres a nivel mundial, solo una mínima parte de la población concentra gran parte de ésta riqueza. Sólo las grandes empresas pueden pagar por estos dispositivos. Del otro lado, donde está la mayoría, por ejemplo, los hospitales públicos no cuentan con éstas tecnologías, y que, de poder contar con ellas, representaría una gran herramienta para el tratamiento de enfermedades. En las universidades públicas no se cuenta con el recurso suficiente para proveer a laboratorios de ingeniería con estos dispositivos, y que, de poder tenerlos, se podría empezar a conocer el funcionamiento y poco a poco ir mejorándolos. Faltarían muchas áreas que mencionar, pero éstas dos últimas son una cadena que depende una de otra, el tener ingenieros, maestros o doctores con la capacidad de inventar dispositivos para cualquier tipo de aplicación, detonaría el crecimiento tecnológico de México en el área de reconstrucción tridimensional.

Es por esto que, que parte de la presente investigación, se trata de explicar de la manera más detallada el proceso de reconstrucción tridimensional de la técnica de luz estructurada con visión estereoscópica, errores comunes se describen, otro tipo de técnicas se mencionan y de manera general se describe el contexto para el que fue creado un prototipo.

## **ALCANCE**

La presente investigación tiene como alcance comprender teórica, matemática y computacionalmente la reconstrucción tridimensional de objetos mediante la técnica de luz estructurada con visión estereoscópica, conocer todas las etapas del proceso para obtener los mejores resultados que nos permita el hardware utilizado y así poder realizar mejoras en trabajos futuros.

## **LIMITACIONES**

Durante el desarrollo experimental, se trabajó con hardware de bajo costo y por consecuencia de bajo rendimiento, no obstante, permitió comprobar que no es necesario hardware de última generación para obtener resultados con precisión aceptable. También, no fue posible adaptar al dispositivo un proyector que permitiera la concentración de luz para obtener mejores resultados en presencia de luz solar, lo que limitó el desarrollo a procesamiento digital de las imágenes adquiridas.

## CAPÍTULO 1 INTRODUCCIÓN Y ANTECEDENTES

### 1. Reconstrucción tridimensional a partir de imágenes en ambientes no controlados

Las técnicas de reconstrucción tridimensional a partir de imágenes es un campo de la visión por computadora y tienen ya un largo recorrido. De hecho, la primera publicación de reconstrucción de objetos 3d a partir de imágenes 2D fue con el trabajo de Lawrence G. Roberts, quien en su tesis de investigación (Roberts, 1965), plasmó las primeras reglas y supuestos de la percepción de profundidad que fueron cuidadosamente analizadas y mecanizadas. Con este trabajo para obtener su título de doctorado, puede decirse que nace el concepto de visión por computadora.

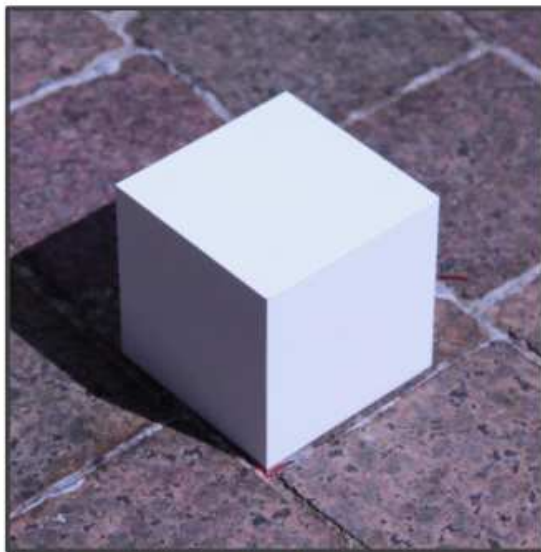
Existen diversas técnicas para la adquisición de los datos de profundidad, sin embargo, la técnica de reconstrucción con luz estructurada, es la más usada, pues un 90% de escáneres 3d comerciales se basan en esta técnica, debido su rapidez y precisión en el proceso de reconstrucción. La línea que siguen las investigaciones recientes, apuntan al desarrollo de técnicas y algoritmos para reducir el tiempo de adquisición y aumentar la precisión de las reconstrucciones, así como mejorar las técnicas de reconstrucción en ambientes no controlados.

La técnica de reconstrucción tridimensional a partir de imágenes con proyección de luz estructurada, tiene varias ventajas respecto a otras técnicas: 1) no requiere contacto directo con el objeto; 2) realiza una rápida adquisición de captura con gran precisión; 3) tiene un escaneo de campo completo; 4) es fácil de automatizar; 5) puede ser portable y 5) hay seguridad para el usuario al hacer uso de luz no dañina.

Sin embargo, uno de los problemas que se sigue tratando de mejorar, es la reconstrucción en ambientes no controlados, donde el objeto puede estar expuesto a iluminación indirecta como luz solar, que puede alcanzar hasta 90000 lux, o efectos de reflexión y dispersión subsuperficial que dominan la iluminación directa



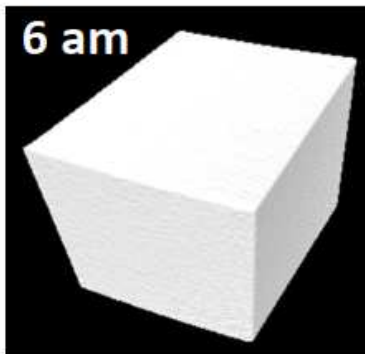
y dependen en gran medida de la forma y las propiedades del material del objeto o escena, resultando en una saturación del sensor disminuyendo la relación señal ruido o SNR (SNR Signal to Noise Relation, por sus siglas en ingles) que opacan la luz estructurada emitida por los proyectores, dando como resultado reconstrucciones de una menor cantidad de puntos 3d (Figura 1).



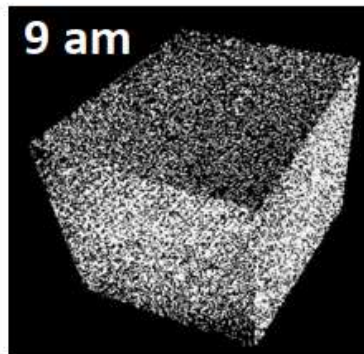
(a) An object placed outdoors



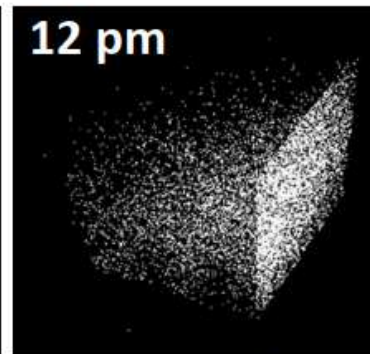
(b) Image of the sky



6 am



9 am



12 pm

(c-e) 3D reconstructions at different times of the day

Figura 1 Efecto de la luz del sol directa.

En (Gupta et al., 2013), hacen referencia a los efectos que tiene la luz solar directa en objetos que se reconstruyen a partir de la técnica de luz estructurada, ellos proponen el uso de un proyector personalizado con luz láser (Figura 2), donde

enfocan la luz láser en franjas que van cambiando de tamaño sobre el objeto a través de un arreglo óptico que incrementa la relación señal ruido, ellos lo llaman concentración de luz (Figura 3) para para lidiar con una fuerte iluminación ambiental, logrando resultados sorprendentes a cualquier hora del día.

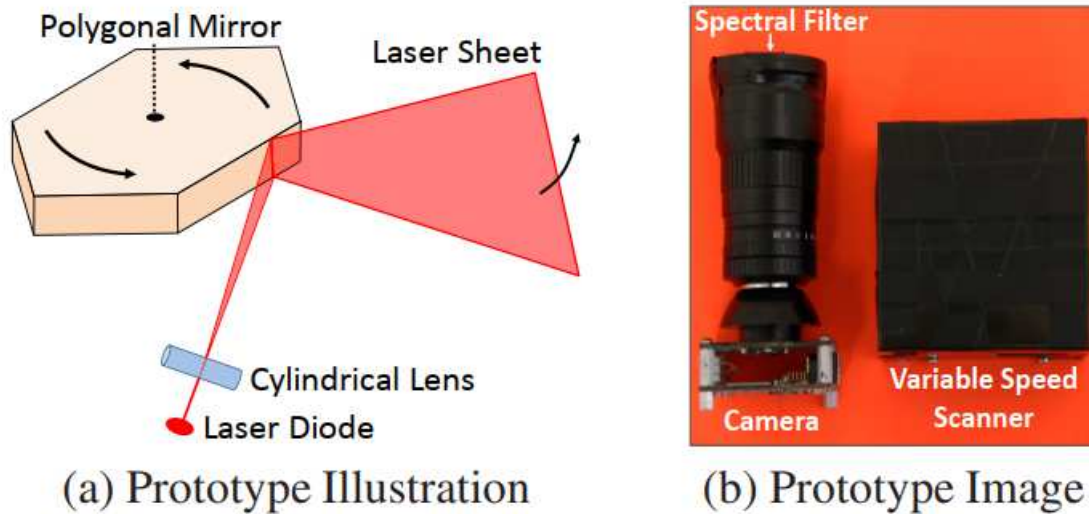


Figura 2 Arreglo óptico para concentración de luz.

En sus pruebas, ellos aseguran que los resultados se pueden alcanzar hasta con proyecciones de baja luminosidad. El problema es, que se necesitan de muchas más imágenes para completar los patrones de luz estructurada que es proporcional al tiempo que necesita la técnica para la adquisición de las imágenes.

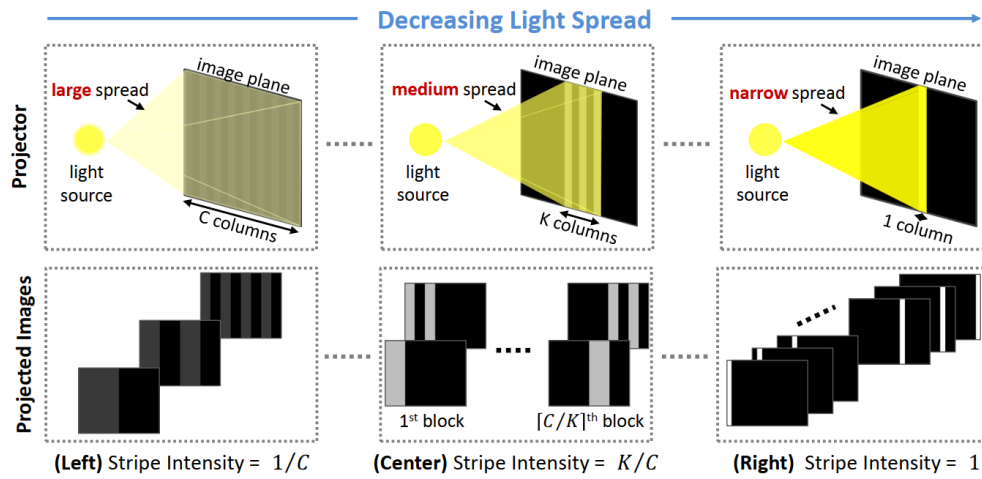


Figura 3 Técnica de concentración de luz.

En (Gupta et al., 2011), se analizan los efectos de la iluminación indirecta que inciden sobre el objeto y que degradan los resultados en las reconstrucciones con las técnicas de luz estructurada. Ellos proponen hacer un pre procesamiento de cada imagen del patrón proyectado (Figura 4) y capturado por las cámaras y realizar un análisis y en función de los efectos detectados, lanzar un patrón idóneo para obtener las mejores tomas de los patrones.

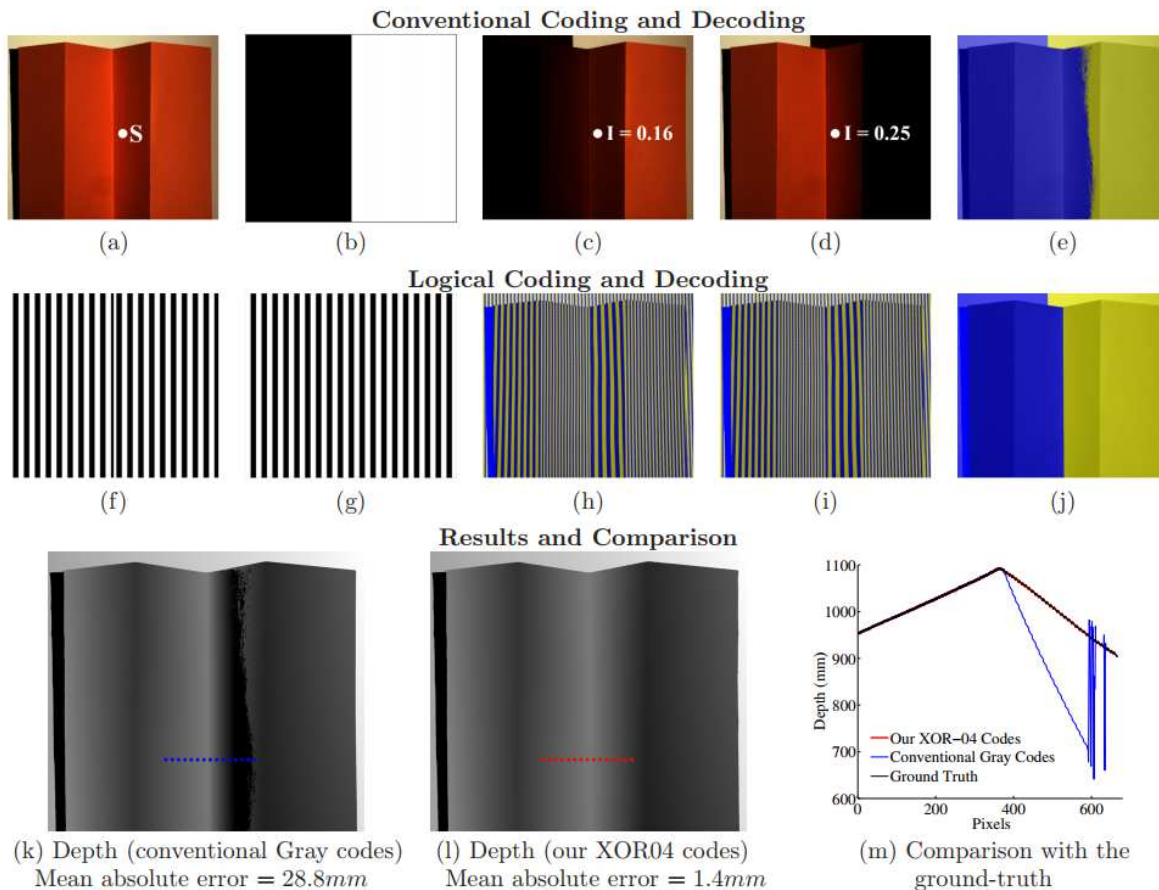


Figura 4 Preprocesamiento de cada proyección de patrón.

El resultado es una decodificación correcta del patrón lanzado, eliminando de esta forma los efectos que disminuyen el éxito de reconstrucción.

En (Wang, 2010), se toma como ejemplo la perfilometría de medición de fase (PMP) y se propone la estrategia de patrón de borde que logra la relación señal / ruido máxima (SNR) para los patrones proyectados con el fin de eliminar el mayor ruido posible en las imágenes capturadas.

Como hemos visto en estos trabajos, la mayoría realiza un pre procesamiento de las imágenes que capturan las cámaras para obtener la máxima relación señal / ruido y obtener buenos resultados en ambientes no controlados.

En esta investigación, se propone un algoritmo ligero de pre ajuste automático de las imágenes captadas por las cámaras para obtener resultados con número similar de puntos 3d en ambientes no controlados.

## CAPÍTULO 2 MARCO TEÓRICO

### 2.1 Visión por computadora

La visión es uno de los sentidos de la mayoría de los organismos, a través de la visión, se conoce el mundo de elementos que nos rodean. Una persona, es capaz de percibir emociones a través de sus ojos, ha permitido entender y plasmar en libros de cualquier tipo el conocimiento que interpretan su visión.

A partir de (Roberts, 1965), donde se discuten por primer vez la extracción de información 3d sobre objetos sólidos a partir de fotografías 2D de dibujos lineales y se mencionan palabras como transformaciones de cámara, efectos de perspectiva y reglas y supuestos de la percepción de profundidad, se puede decir que nace el concepto de visión por computadora o visión artificial.

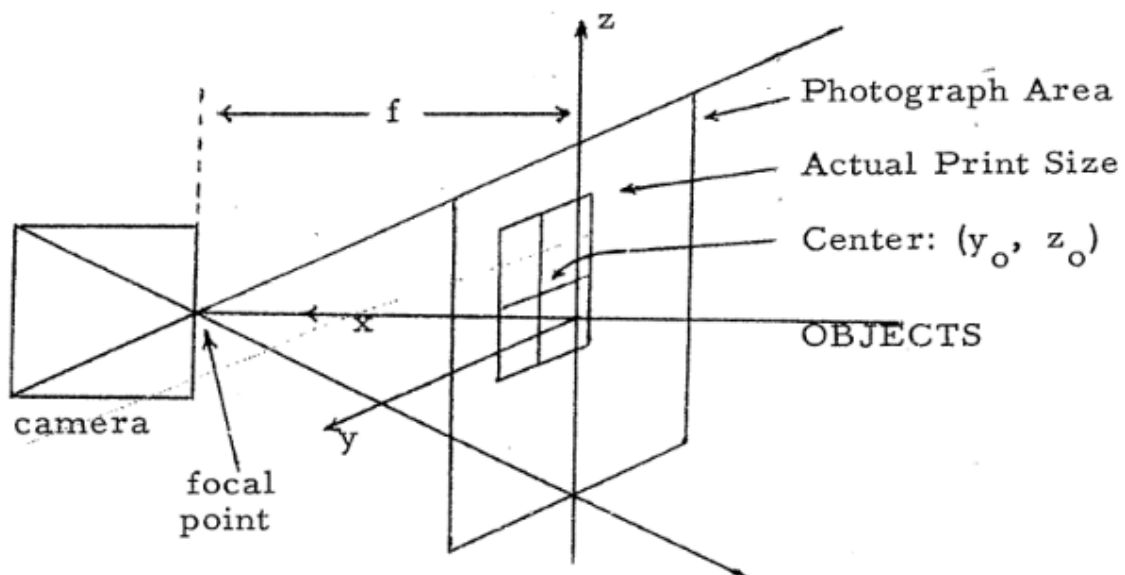


Figura 5 Transformación de la cámara (Roberts, 1965).

En la **¡Error! No se encuentra el origen de la referencia.**, se muestra una de las figuras donde se explica la transformación de una escena del mundo real a 2D, y es muy parecida a los diagramas que vemos hasta hoy en varios libros y artículos de procesamiento de imágenes y visión por computadora.

A diferencia del procesamiento digital de imágenes donde el objetivo es hacer mejoras a imágenes a través de procesos de funciones matemáticas para obtener una mejor imagen para su posterior interpretación, la visión por computadora tiene como objetivo extraer características de una imagen para su descripción e interpretación.



Figura 6 Reconocimiento de placas de carros.

En la Figura 6, se puede observar el resultado de un algoritmo de visión por computadora (Masood et al., 2017), utilizando imágenes con automóviles donde se detecta la posición de la placa y los caracteres que tienen impresos, esto resulta muy útil para automatizar, por dar un ejemplo, la búsqueda de algún automóvil en carretera, el algoritmo se encuentra procesando los frames de un video en tiempo real, sin la intervención humana.

Entonces, la visión por computadora es una disciplina científica que incluye métodos para adquirir, procesar, analizar y comprender las imágenes del mundo real con el fin de producir información numérica o simbólica para que puedan ser tratados por una computadora, tal como lo hace el cerebro a través de los ojos, trata de emular el funcionamiento del cerebro para que las computadoras puedan percibir y comprender una imagen o video y actuar ante una determinada situación.

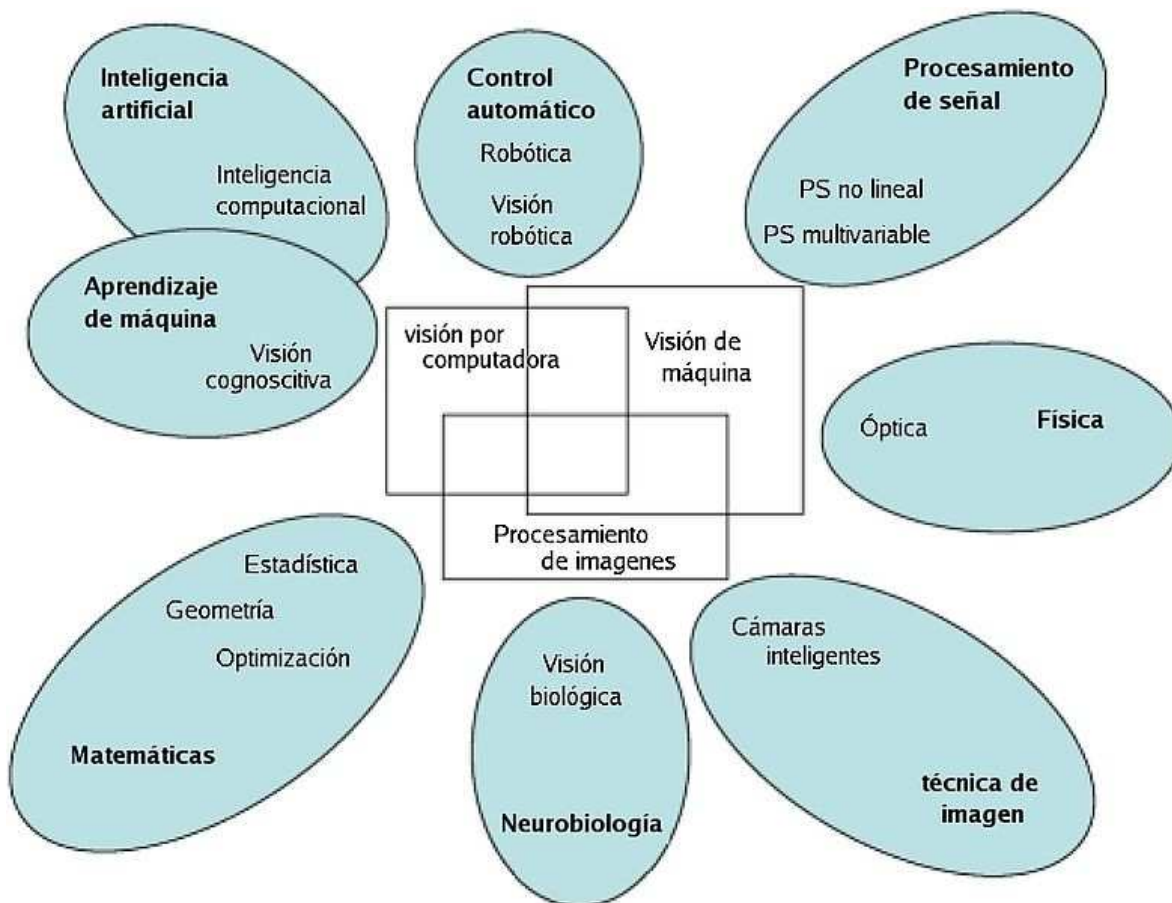


Figura 7 Esquema de las relaciones entre la visión por ordenador y otras disciplinas (imagen de dominio público).

En la Figura 7, se muestra la relación entre la visión por computadora y otras disciplinas, se puede observar que aunque tiene campos exclusivos, también necesita de otros campos de otras áreas para la solución de problemas.

Se podrían citar muchos más ejemplos, pero ahora, nos enfocaremos en un área de la visión por computadora que como hemos visto, fue el inicio de todo. La reconstrucción tridimensional o 3d a partir de imágenes bidimensionales o 2D.

Se puede decir que, a partir de la salida al mercado de cámaras digitales con un precio más accesible, por ejemplo, el caso de las cámaras web, aumentó el número de investigaciones en el tema de reconstrucción 3d.

## 2.2 Formación de la imagen

Una imagen es una representación bidimensional de una escena del mundo tridimensional. La imagen es el resultado de la adquisición de una señal proporcionada por un sensor, que convierte la información del espectro electromagnético en codificaciones numéricas. Y su transformación en el formato elegido de representación de imágenes, que constituye información discreta tanto en los valores que puede tomar, como en los parámetros que la definen. De forma genérica, una imagen digital se define como una matriz (o vector) de dimensiones  $N \times M$ , conteniendo en cada elemento de la matriz un valor discreto que cuantifica el nivel de información del correspondiente elemento, representado con un número finito de bits (Alegre et al., 2016).

Así, una imagen es una función bidimensional que proporciona cierta información electromagnética para cada uno de sus valores. A cada uno de estos elementos discretos se le denomina punto o píxel y generalmente contiene el nivel de iluminación o el color de un punto en la escena. El conjunto de puntos o píxeles forman la imagen o fotografía de la escena. Se dice que una imagen es fruto de la representación espectral recibida por el sensor. En este sentido por ejemplo el canal rojo en una imagen RGB correspondería a la representación espectral de las



correspondientes longitudes de onda. El color se genera por la superposición de tres componentes espectrales.

El proceso de formación y captura de una imagen depende de un conjunto de elementos cada uno de los cuales cumple una misión determinada. Básicamente, el proceso consiste en generar una señal electromagnética en forma de luz que se transmite, refracta y refleja en la escena hasta que alcanza el área del sensor de la cámara. En el sensor se produce la transducción a una señal eléctrica y a partir de

ese momento se dispone de un conjunto de elementos electrónicos que digitalizan, formatean, transmiten y almacenan las imágenes digitales.

## 2.3 Modelo de cámara pinhole

Las cámaras fotográficas están compuestas por un arreglo óptico que permite el paso de ondas de luz reflejadas de un objeto o escena y las proyecta sobre un sensor compuesto por una matriz bidimensional de elementos fotosensibles que las convierte en señales eléctricas (Figura 8. Sensor de una cámara fotográfica.Figura 8) y que posteriormente pueden ser analizadas, almacenadas y representadas como un patrón analógico o digital.



Figura 8. Sensor de una cámara fotográfica.

Éste patrón es la representación en dos dimensiones de la escena capturada del mundo real. En el mundo real, tenemos un plano  $x,y,z$ , donde  $z$  es la profundidad de los objetos en la escena, cuando esta escena es captada por el sensor bidimensional, se pierde la información de  $z$ , ya que una línea en 3d se convierte en un punto o pixel de la imagen, de esta forma, aunque se puede estimar el valor de  $z$  mediante las técnicas de forma de sombreado o a partir de la textura, en un sentido matemático no es posible la recuperación de dicha información con exactitud.

Para éste trabajo, ya que se va a utilizar un arreglo óptico de visión estéreo, primero se debe entender algunos aspectos que se utilizarán más adelante. El modelo de cámara pinhole es el que nos muestra los aspectos básicos para comprender el modelo estéreo.

En la Figura 9, se muestra el modelo de cámara pinhole, modelado por un centro óptico  $O$  y el plano de la imagen  $\Pi$ . Un punto 3d del mundo real  $W$  se proyecta en un punto de imagen  $p$ , que es una intersección de  $\Pi$  con la línea que contiene  $O$  y  $W$ . Sea  $W=[x \ y \ z]^T$  las coordenadas en el sistema de coordenadas del mundo real,  $q=[x,y]^T$  un punto en el plano de la imagen. El mapeo de coordenadas 3d a coordenadas 2D es la proyección en perspectiva, que se representa mediante una transformación lineal en coordenadas homogéneas (S. Zhang, 2013).

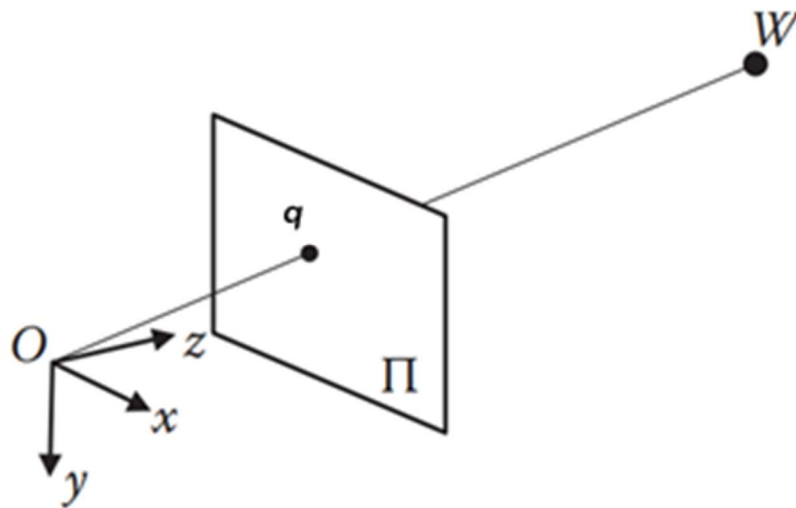


Figura 9 Modelo de cámara Pinhole.

## 2.4 Distorsiones asociadas más significativas

Existen cámaras con un sofisticado proceso de fabricación, con componentes de muy alta calidad que permiten hacer tomas con mínimas deformaciones, sin embargo, la calidad de las cámaras es proporcional al costo de estos dispositivos. Actualmente, para cuestiones de investigación académica, se pueden conseguir cámaras con costos accesibles, pero con un incremento de distorsiones.

Los principales tipos de distorsiones son: la distorsión radial y la distorsión tangencial y tienen que ver con el objetivo (conjunto de lentes) de la cámara.

La distorsión radial se incrementa debido a la forma de las lentes, en la práctica ninguna lente es perfecta, las líneas rectas aparecerán curvas. Su efecto es mayor a medida que nos alejamos del centro de la imagen.

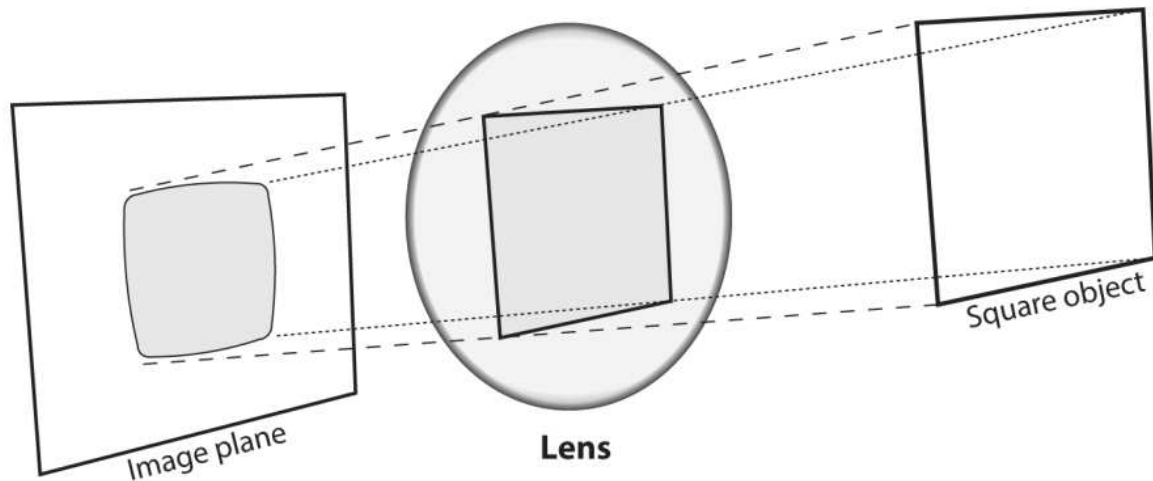


Figura 10 Distorsión radial debido al lente de la cámara.

En la Figura 10 los rayos más alejados del centro de una lente simple se doblan demasiado en comparación con los rayos que pasan más cerca del centro; así, los lados de un cuadrado parecen arquearse en el plano de la imagen.

Para modelar la distorsión radial, y por tanto poder corregir este tipo de problemas, existen muchos modelos de distorsión (Ricolfe-Viala & Sánchez-Salmerón, 2010), aunque se suele emplear los tres primeros términos de la serie de Taylor calculada alrededor del punto principal 0.

La serie matemática de Taylor se emplea para modelar de forma aproximada funciones complejas alrededor de un punto de forma que con cada nuevo término polinomial, la aproximación es mejor. En este caso, se expande la función de Taylor polinomial de la siguiente forma:  $f(r) = a_0 + a_1r_1 + a_2r_2^2 + \dots + a_6r_6^6$ .

Debido a que se calcula en el punto principal 0,  $f(r) = 0; r = 0; a_0 = 0$  y al ser una función simétrica, sólo los coeficientes pares 2, 4 y 6 son necesarios ( $a_2, a_4$  y  $a_6$ ) a los cuales llamaremos respectivamente  $k_1, k_2, k_3$ . Una vez calculados estos coeficientes, es posible corregir la distorsión radial siguiendo la ecuación:

$$\begin{bmatrix} u' \\ v' \end{bmatrix} = (1 + k_1 r^2 + k_2 r^4 + k_3 r^6) \begin{bmatrix} u \\ v \end{bmatrix} \quad (1)$$

Donde  $(u', v')^T$  son las coordenadas de los píxeles corregidos y  $(u, v)^T$  las de los píxeles de la imagen distorsionada.

Por ejemplo, en la Figura 11, se observan dos bordes de un tablero de ajedrez que están marcados con líneas rojas. Se puede observar que el borde no es una línea recta y no coincide con la línea roja. Todas las líneas rectas esperadas están abultadas.

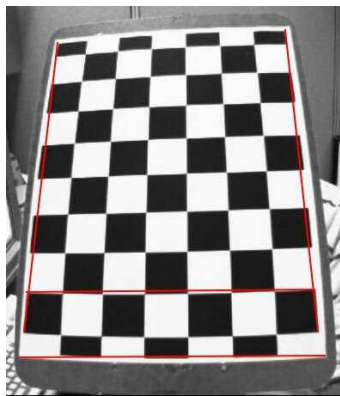


Figura 11 Distorsión radial (Zelinsky, 2009).

La distorsión tangencial ocurre cuando la lente y el plano de la imagen no son paralelos. Los coeficientes de distorsión tangencial modelan este tipo de distorsión. Se caracteriza mínimamente por dos parámetros adicionales  $p_1$  y  $p_2$ .

Para corregir la distorsión tangencial, se tienen otras dos ecuaciones:

$$x_{\text{corregido}} = x + [2p_1xy + p_2(r^2 + 2x^2)] \quad (2)$$

$$y_{\text{corregido}} = y + [p_1(r^2 + 2y^2) + 2p_2xy] \quad (3)$$

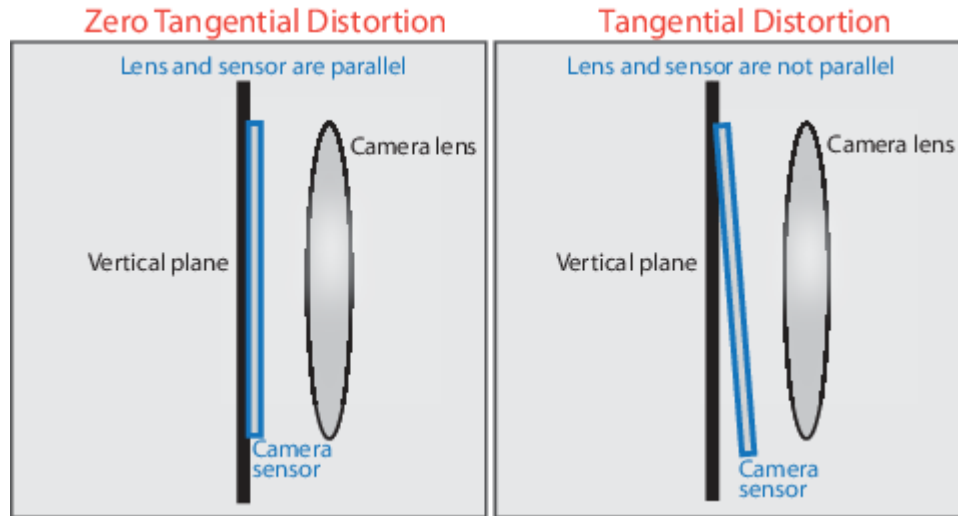


Figura 12 Defectos de distorsión tangencial en cámaras.

En la Figura 12, se observa del lado izquierdo la correcta posición de sensor y lente óptico, se encuentran paralelos con un índice cero de distorsión. Del lado derecho, el sensor y el lente no se encuentran paralelos, y genera la distorsión tangencial.

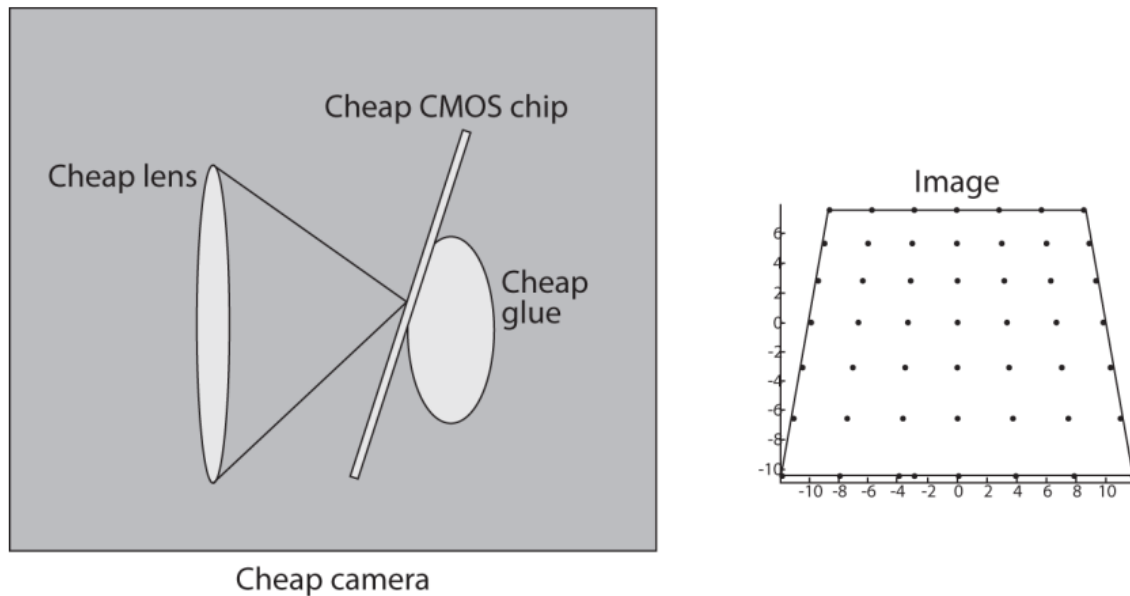


Figura 13 Resultados de distorsión tangencial (Zelinsky, 2009).

En la Figura 13, se puede observar el resultado de la distorsión radial. En la imagen capturada, algunas áreas pueden verse más cerca de lo esperado.

En general, se necesitan 5 coeficientes de distorsión y la mayoría de las librerías de visión por computadora, tienen las funciones para obtenerlas. La derivación de estos coeficientes va más allá del objetivo de esta investigación y pueden verse en (Brown, 1966).

$$\text{coeficientes de distorsión} = (k_1 k_2 p_1 p_2 k_3) \quad (4)$$

Teniendo estos coeficientes de distorsión necesitamos ahora conocer las características internas de la cámara. Incluyen información como distancia focal ( $f_x$ ,  $f_y$ ) y centros ópticos ( $c_x$ ,  $c_y$ ). La distancia focal y los centros ópticos se pueden usar para crear una matriz de la cámara, que se puede usar para eliminar la distorsión debida a las lentes de una cámara específica. La matriz de la cámara es exclusiva de una cámara específica, por lo que una vez calculada, puede reutilizarse en otras imágenes tomadas por la misma cámara. Se expresa como una matriz 3x3:

$$\text{matriz de parámetros intrínsecos} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} = C \quad (5)$$

Donde:

$f_x, f_y$  representan la longitud focal en dirección horizontal y vertical respectivamente en píxeles.

$c_x, c_y$  representa el centro óptico (el punto principal), en píxeles.

Las imágenes con distorsión radial y tangencial pueden ser corregidas mediante procesos y técnicas matemáticas que en conjunto se le conocen como calibración de cámaras.

## 2.5 Calibración de cámaras estéreo

El objetivo de la calibración de un arreglo de cámaras con visión estéreo, es para conocer los parámetros intrínsecos y extrínsecos de las cámaras.

La geometría para la calibración de cámaras, también conocida como rectificación de cámara, estima los parámetros de una lente y un sensor de imagen de una cámara. Se Pueden usar estos parámetros para corregir la distorsión de la lente, medir el tamaño de un objeto en unidades reales o determinar la ubicación de la cámara en la escena. Estas tareas se utilizan en aplicaciones como la visión por computadora para detectar y medir objetos. También se utilizan en robótica, para sistemas de navegación y reconstrucción de escenas en 3d (Matlab, 2019).

Los parámetros intrínsecos tienen que ver con la estructura interna de la propia cámara. A través de procesos matemáticos, se pueden obtener la distancia focal, el centro óptico, las distorsiones anteriormente detalladas, con el fin de corregir las imágenes para su posterior reconstrucción.



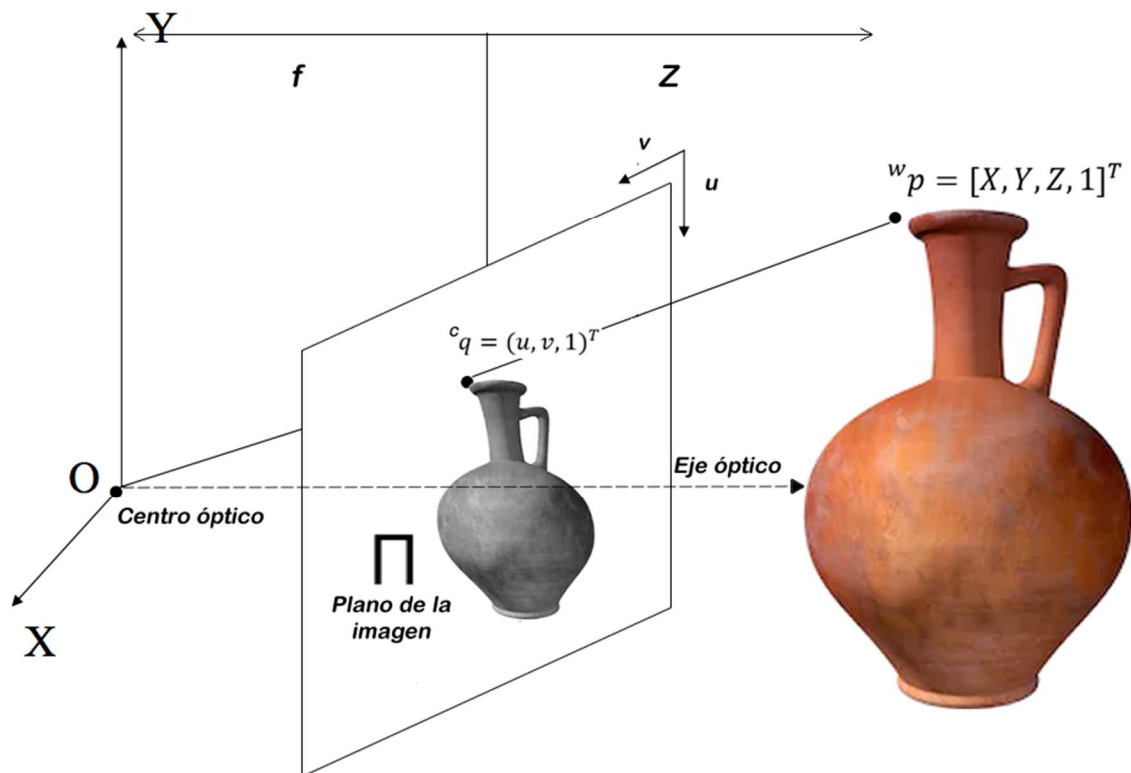


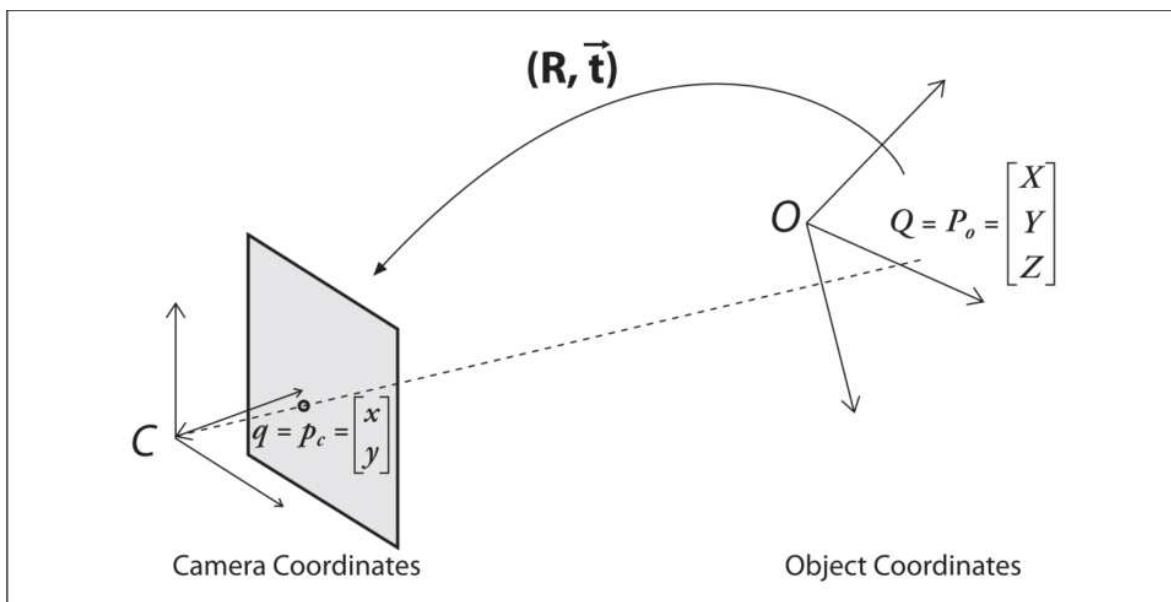
Figura 14 Esquema del modelo pinhole.

Los parámetros extrínsecos, tienen que ver con la disposición de la cámara o arreglo de cámaras respecto al objeto. Se obtienen datos como la rotación y traslación.

Dado el sistema de referencia  $W$  de los puntos 3d denotado por  ${}^w p = [X, Y, Z, 1]^T$  y el sistema de referencia  $C$  asociados a la cámara (Figura 14), la matriz de transformación homogénea  $T$ , es una matriz de  $4 \times 4$  que define la localización del sistema de coordenadas del mundo  $W$  con respecto al sistema de coordenadas de la cámara  $C$ , de tal manera que,  ${}^c q = {}^c T_w \cdot {}^w p$ , de este modo, obtenemos el punto  ${}^c q$  del plano de la cámara a partir del punto  ${}^w p$  (Figura 14).

$$T = {}^cT_w = \begin{bmatrix} R_{3 \times 3} & t_{3 \times 1} \\ 0_{1 \times 3} & 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (6)$$

Donde  $R$  es la matriz de rotación de  $3 \times 3$  que define la orientación de  $W$  con respecto a  $C$ ,  $t$  es el vector columna de traslación  $3 \times 1$  que define donde está el origen de  $W$  con respecto a  $C$ ; y  $0$  es el vector fila nulo  $1 \times 3$ .



Con la información anterior, ahora tenemos que calibrar el arreglo de cámaras con visión estéreo, que no es más que el proceso de encontrar la matriz de rotación  $R$  y el vector de traducción  $T$  entre las dos cámaras.

Matlab, OpenCV (utilizado en éste trabajo de investigación), contienen herramientas para obtener las matrices antes utilizadas. Comúnmente se utiliza el método del tablero de ajedrez (Z. Zhang, 2000), éste método hace uso de un tablero de ajedrez en donde se tiene un patrón de cuadros blancos y negros con dimensiones conocidas. La función, ejecuta un conjunto de algoritmos de procesamiento de

imágenes como: detección de bordes, detección de esquinas, detección de manchas, detección de crestas y otras funciones que permiten obtener todos los parámetros necesarios para la correcta reconstrucción. Se necesita la adquisición de al menos 10 imágenes del tablero en distinta posición (Figura 15).

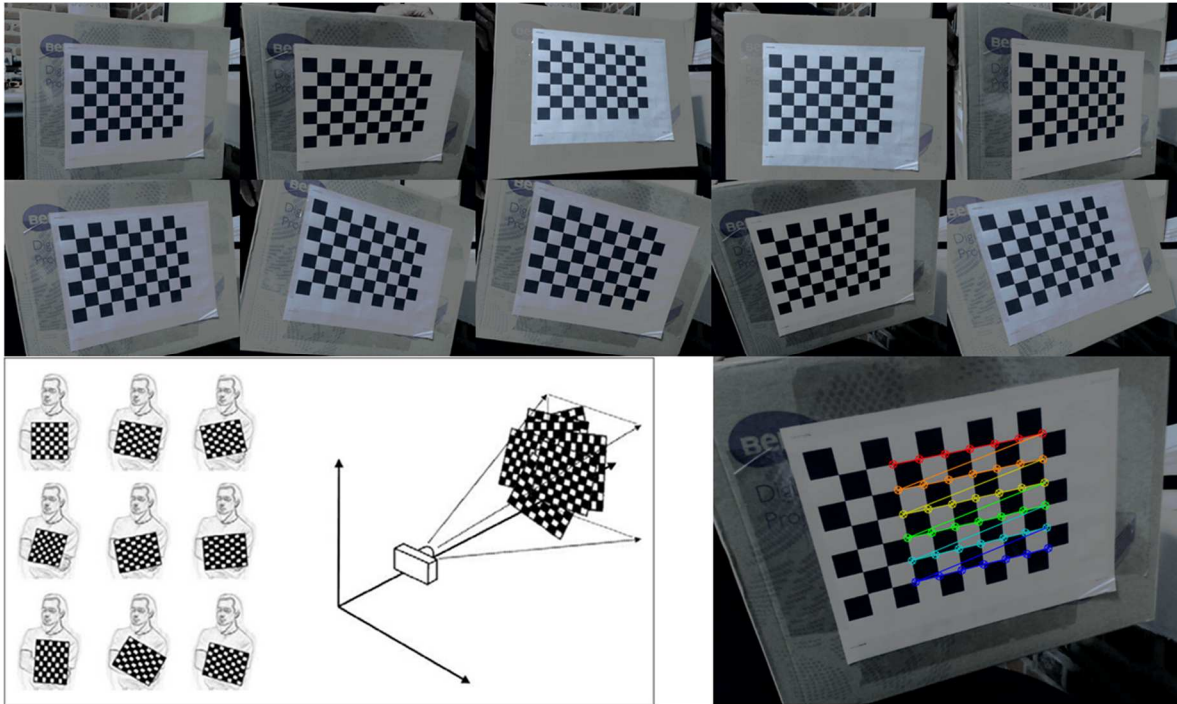


Figura 15 Calibración con tablero de ajedrez.

Una vez finalizado el proceso de captura y procesamiento del patrón en diferentes posiciones, obtenemos la matriz de la cámara, los coeficientes de distorsión, y los vectores de rotación y traslación de la cámara.

Se puede automatizar el proceso obteniendo las capturas a un mismo tiempo en ambas cámaras, logrando obtener los parámetros de una sola vez. En la Figura 16, se muestra el esquema de cómo podemos posicionar ambas cámaras.

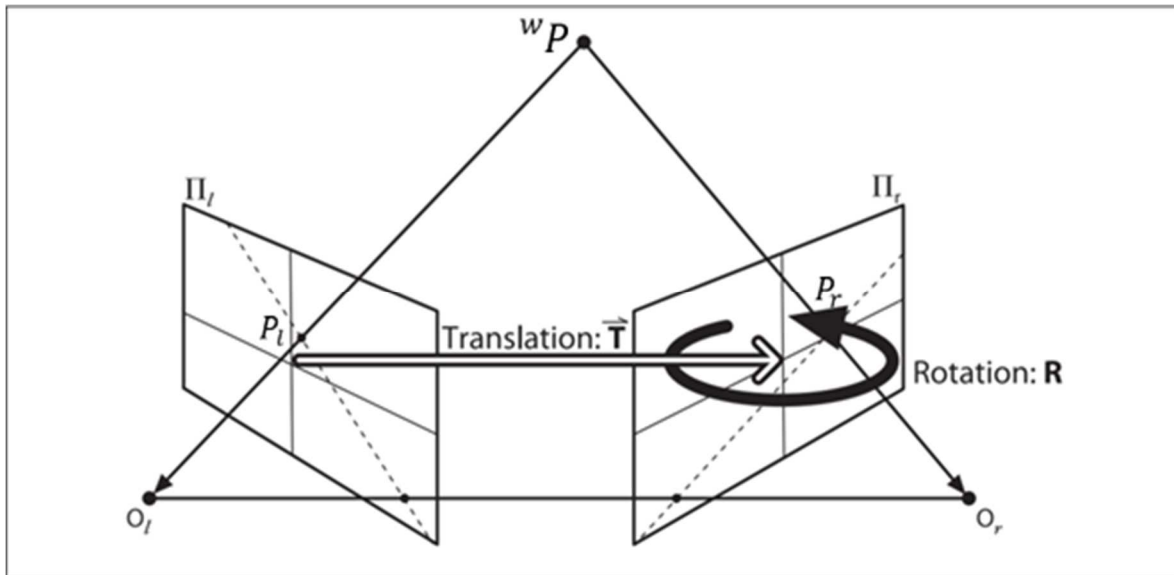


Figura 16 Rotación y traslación entre 2 cámaras.

Sea  $P_l = R_l {}^w P + T_l$  el punto en el plano de la cámara izquierda, donde  $R_l$  es la matriz de rotación de la cámara izquierda,  ${}^w P$  algún punto de la escena del mundo real y  $T_l$  la matriz de traslación de la cámara izquierda, tenemos también  $P_r = R_r {}^w P + T_r$  para la segunda cámara o la cámara derecha.

De acuerdo a la Figura 16, se puede deducir  $P_l = R^T (P_r - T)$ , donde  $R$  y  $T$  son los vectores de rotación y traslación entre ambas cámaras. Tomando las tres anteriores ecuaciones, obtenemos:

$$R = R_r (R_l)^T \quad (7)$$

$$T = T_r - R T_l \quad (8)$$

Con estos datos obtenidos, podemos obtener dos nuevas matrices: la matriz esencial  $E$ , que contiene información de la rotación y traslación entre las dos cámaras y la matriz fundamental, que contiene a  $E$  y adicionalmente la información de los parámetros intrínsecos de ambas cámaras.

## 2.6 Reconstrucción pasiva y activa.

Existen principalmente dos tipos de reconstrucciones a partir de imágenes: pasivas y activas.

La diferencia entre ambas, es que en la reconstrucción activa se utilizan proyecciones de luz (Figura 17), con el objetivo de aumentar la probabilidad de encontrar puntos coincidentes de  ${}^wP$ .

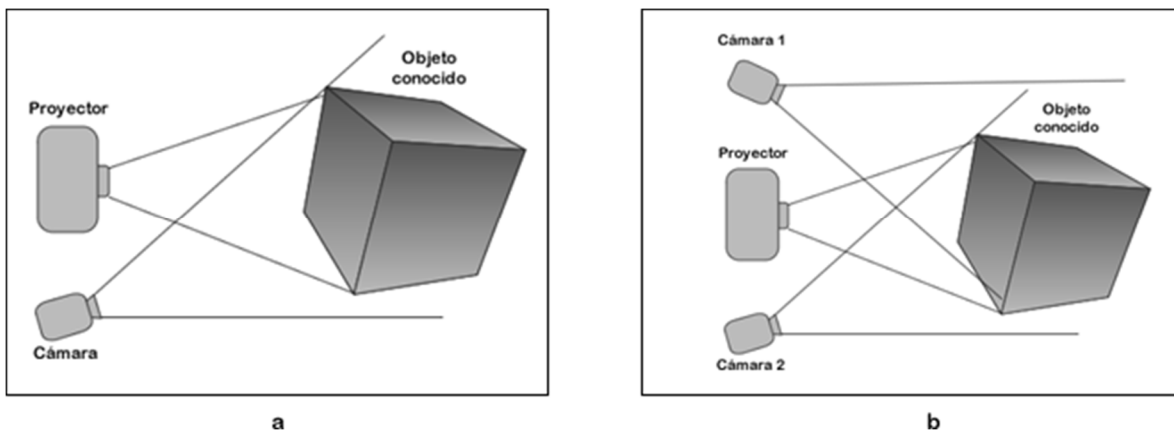


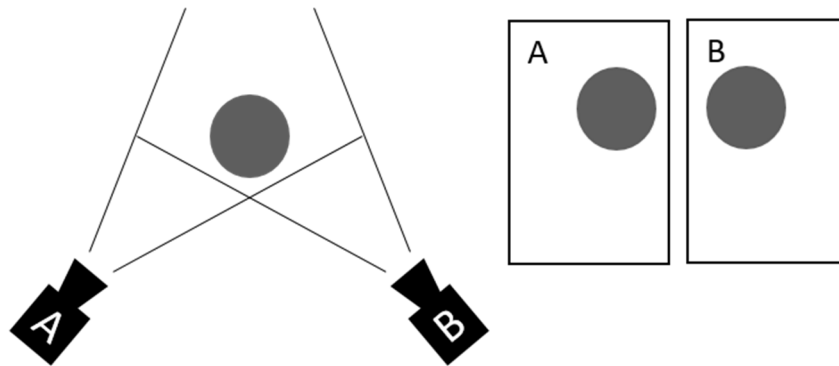
Figura 17 Reconstrucción activa. a) Un proyector y una cámara. b) Un proyector y visión estereoscópica

La reconstrucción pasiva en contraste de la visión activa, se trabaja con la radiación natural presente en la escena. Aunque la visión activa es mucho más efectiva, no siempre es posible hacer uso de ella y por esto, es importante mencionar algunas técnicas pasivas.

## 2.6.1 Técnicas pasivas

### 2.6.1.1 Visión estereoscópica

Consiste en colocar dos cámaras en distintas posiciones a una distancia conocida, en donde se obtienen dos imágenes del mismo punto con las cuales se puede conocer con exactitud el valor de Z. Es una técnica que emula el sistema de visión biológico, la calibración previa, genera reconstrucciones sin distorsiones y muy aproximadas a los objetos reales.



*Figura 18. Visión estereoscópica.*

En la Figura 18, del lado izquierdo se muestra la disposición de las cámaras respecto al objeto, del lado derecho es la imagen del objeto en cada cámara.

Cuando se visualiza un punto desde dos puntos de vista diferentes, su proyección de imagen sufre un desplazamiento desde su posición en la primera imagen a la de la segunda imagen. La cantidad de desplazamiento, alternativamente llamada paralaje o flujo óptico, es inversamente proporcional a la distancia y, por lo tanto, puede usarse para calcular la geometría 3d. Dada una correspondencia entre los puntos de imagen desde dos puntos de vista conocidos, es posible calcular la profundidad por triangulación: intersecar los rayos de cada centro óptico a través de la proyección del punto en el plano de la imagen. El problema de establecer correspondencia es una dificultad fundamental y es el tema de una gran cantidad

de literatura sobre visión estereoscópica. Un enfoque destacado es correlacionar píxeles de intensidades similares en dos imágenes, suponiendo que cada punto de escena refleje la misma intensidad de luz en las dos vistas. El documento incluido de (Okutomi & Kanade, 1993) extiende este enfoque de correlación a dos tres o más imágenes y demuestra que el uso de varias cámaras en diferentes separaciones de cámaras o líneas de base produce una mejora significativa en la precisión de la reconstrucción.

### **2.6.1.2 Estructura desde movimiento (Structure from motion)**

Es una alternativa que, en lugar de capturar una imagen de una escena con dos o más cámaras, se adquieren imágenes, que generan un video, de una sola cámara en movimiento para reconstruir un modelo 3d. El mayor desafío en este enfoque es que la ruta de la cámara debe estimarse ya que la calibración precisa de la cámara que se mueve libremente, es extremadamente difícil (Pollefeys et al., 1999).

El problema de recuperar la geometría de la escena a partir del movimiento de puntos en el plano de imagen de una cámara en movimiento, se denomina estructura a partir del movimiento o structure from motion. Como uno de los problemas clásicos en la visión por computadora, existen numerosas estructuras de algoritmos de movimiento. Existe una técnica (Pollefeys et al., 1999), que mediante un elegante método de factorización, ensambla mediciones puntuales en una matriz de medición que se factoriza en el producto de una matriz de movimiento y forma, usando una descomposición de valores singulares. Este enfoque se basa en un modelo de proyección ortográfica.

### **2.6.1.3 Forma desde sombreado (shape from shading)**

La forma del sombreado o “shape from shading”, es uno de los problemas más simples de expresar y uno de los más complicados de resolver. Plantea una solución para recuperar la forma 3d mediante una imagen de intensidad única de un objeto

curvo liso. Este problema y su solución fue investigada por Berthold Horn en su tesis doctoral (Horn, 1970) y mejorada después en (Brooks & Horn, 1989). Dada la intensidad de un punto en la imagen y una fuente de luz direccional conocida, la ley de Lambert ( $I = kNL$ , donde  $I$  es la intensidad,  $k$  la reflectancia,  $N$  la unidad normal y  $L$  la dirección de la fuente de luz) produce un conjunto de un parámetro de soluciones para la superficie normal. Se necesitan restricciones adicionales para hacer que el problema esté bien planteado: generalmente se resuelve asumiendo una similitud de reflectancia de superficie y orientación en puntos cercanos.

#### **2.6.1.4 Fotometría estéreo**

Las dificultades con la forma del sombreado pueden superarse adquiriendo dos o más imágenes del objeto bajo diferentes ambientes de iluminación. Este es el enfoque de la fotometría estéreo. Cada imagen proporciona una restricción sobre la normal y, por lo tanto, dos imágenes son suficientes para recuperar la normalidad hasta un pequeño número de posibles soluciones, y tres imágenes producen una solución única para cada píxel de imagen. La fotometría estéreo, permite relajar las fuertes condiciones de suavidad impuestas por la forma clásica de los enfoques de sombreado y, por lo tanto, proporciona estimaciones de forma más confiables.

#### **2.6.1.5 Geometría epipolar.**

Ya se ha mencionado que, en el proceso de adquisición de imágenes con una cámara, se pierde información de profundidad, y que podemos hacer uso de 2 o más cámaras para recuperar dicha información.

La geometría epipolar (Zelinsky, 2009), representa la geometría de la visión estereoscópica y plantea la forma de encontrar un punto  ${}^w p$  en ambas cámaras sin recorrer toda la imagen como en otros métodos o técnicas.



En la Figura 19, se representa dicha geometría. Si observamos en el plano de la imagen izquierda el punto  $x$ , podemos observar que existen otros puntos  $X$ , que se pierden por la perspectiva de la cámara hacia el objeto de la línea  $OX$ . Sin embargo, con la ayuda de otra cámara con rotación conocida y traslación hacia la derecha, podemos recuperar todos los puntos  $X$ , que se pierden en la cámara izquierda.

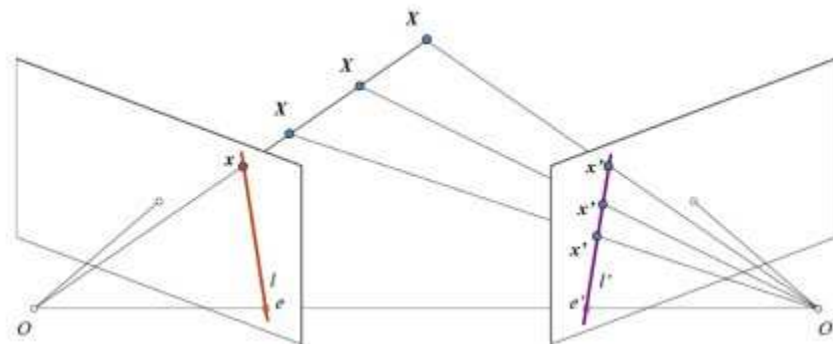


Figura 19 Geometría epipolar.

Ahora diferentes puntos de la línea  $OX$  se proyectan a diferentes puntos  $x'$  en la imagen del plano derecho. Así que con estas dos imágenes podemos triangular el punto 3d correcto.

La proyección de los diferentes puntos en  $OX$  forman una línea  $l'$  en el plano de la imagen derecha denominada epilínea correspondiente al punto  $x$ . Esto quiere decir, que para encontrar el punto  $x$  en la imagen derecha, se busca a lo largo de esta epilínea, lo que se traduce en menor tiempo y costo computacional. A este proceso, se le denomina Restricción Epipolar. Del mismo modo, todos los puntos  $x$  tendrán sus correspondientes epilíneas en la otra imagen. El plano  $XOO'$  se denomina Plano epipolar.

$O$  y  $O'$  son los centros de la cámara, respectivamente. A partir de la configuración detallada anteriormente, se puede ver que la proyección de la cámara derecha  $O'$  se ve en la imagen izquierda en el punto,  $e$  llamado epipolo. El epipolo es el punto de intersección de la línea a través de los centros de la cámara y los planos de la

imagen. Similarmente  $e'$  es el epipolo de la cámara izquierda. En algunos casos, no podrá localizar el epipolo en la imagen, puede que esté fuera de la imagen (lo que significa que una cámara no ve la otra).

Todas las epilíneas pasan por su epipolo. Así que para encontrar la ubicación del epipolo, podemos encontrar muchas epilíneas y encontrar su punto de intersección.

Para encontrar la información de la geometría epipolar, se requiere de las matrices esencial y fundamental, que fueron presentadas con anterioridad.

## 2.6.2 Técnicas activas.

### 2.6.2.1 Luz estructurada

La técnica de luz estructurada en un sistema de reconstrucción tridimensional, tiene como ventajas la elevación del porcentaje de detección de puntos y la rapidez en el proceso de captura de imágenes por las cámaras fotográficas. En el arreglo óptico, se agrega un proyector de luz, regularmente se utilizan proyectores de video que proyectan la imagen correspondiente en una pantalla de proyección usando un arreglo de una o más cámaras, permitiendo así mostrar imágenes fijas o en movimiento (Figura 20).

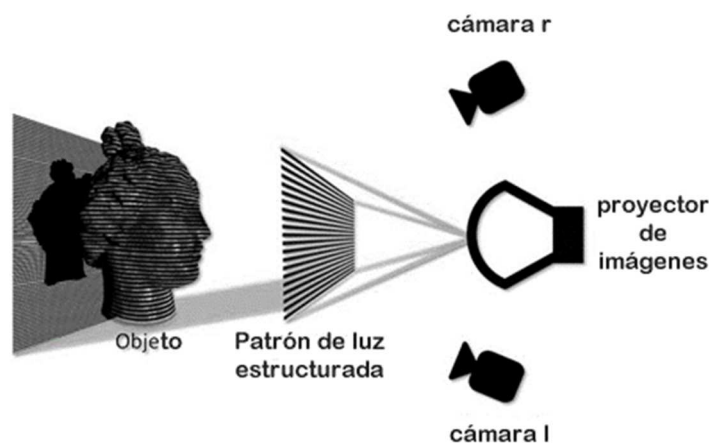


Figura 20 Técnica de luz estructurada.

Dependiendo la técnica utilizada, se pueden realizar reconstrucciones con una o más cámaras.

## 2.7 Técnicas de reconstrucción de luz estructurada

### 2.7.1.1 Cambio de fase (phase shifting)

Con ésta técnica, pueden realizarse reconstrucciones de alta calidad, prácticamente en tiempo real (AU - Ekstrand et al., 2013). El proceso consiste en proyectar un patrón de luz estructurada con señales cosenoidales, agregando un desplazamiento de fase en cada imagen sobre el objeto que se va a reconstruir. El arreglo óptico, consiste en una cámara y un proyector de imágenes con una rotación y traslación que deben ser calculadas con la calibración de la cámara para obtener los parámetros intrínseco y extrínsecos, como ya se explicó anteriormente. Al proyectarse sobre el objeto, el patrón se deforma por la topografía superficial del objeto.

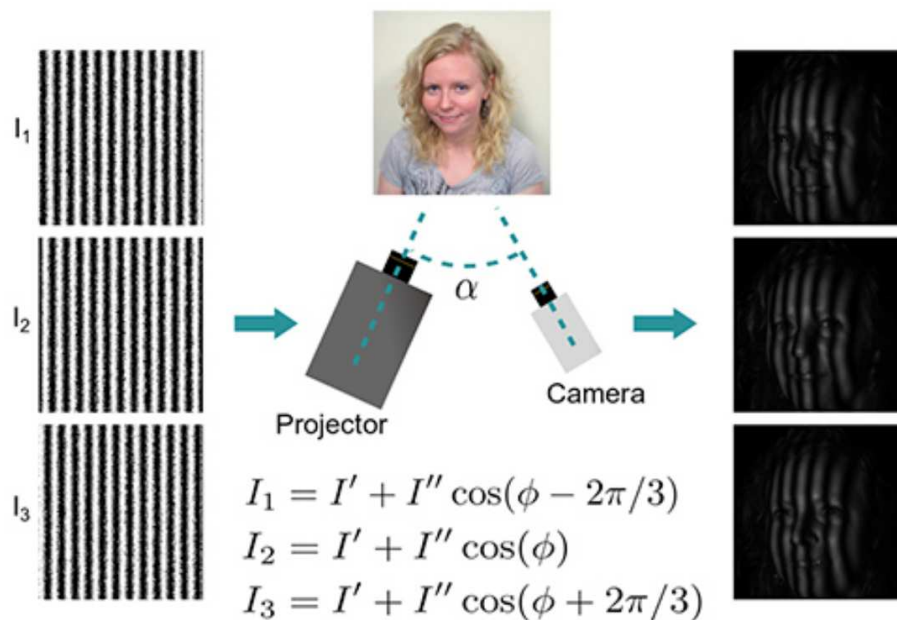


Figura 21 Phase Shifting con 3 patrones.

Deben proyectarse los patrones sobre un plano de referencia y sobre el objeto a reconstruir.

Las imágenes adquiridas son codificadas para obtener una sola imagen (Figura 22) con la información de las señales cosenoidales y la topografía del objeto en forma discontinua (Gonzalez et al., 2012). Ésta nueva imagen, debe ser decodificada para poder realizar la correspondencia entre el patrón proyectado con el patrón capturado por la cámara.

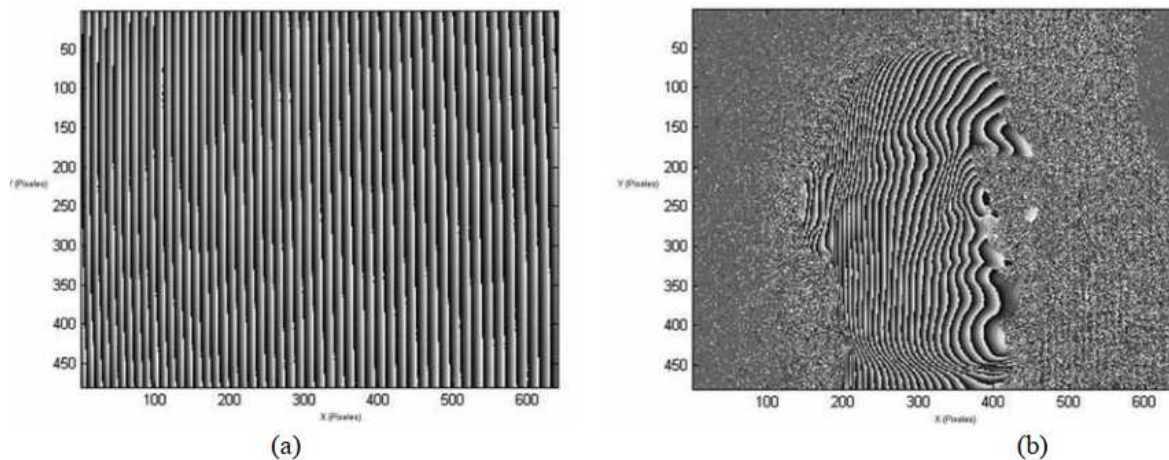


Figura 22 Imagen con información de patrones y topografía del plano (a) y objeto (b).

Posteriormente, el proceso de decodificación, consiste en aplicar un algoritmo de desenvolvimiento de fase que detecte los saltos de  $2\pi$  en la fase y los elimine sumando o restando múltiplos de  $2\pi$ , obteniendo una nueva imagen con la fase de forma continua (Figura 23).

Para finalizar, la profundidad se puede obtener calculando la diferencia entre la fase desenvuelta del objeto y la fase desenvuelta del plano de referencia.

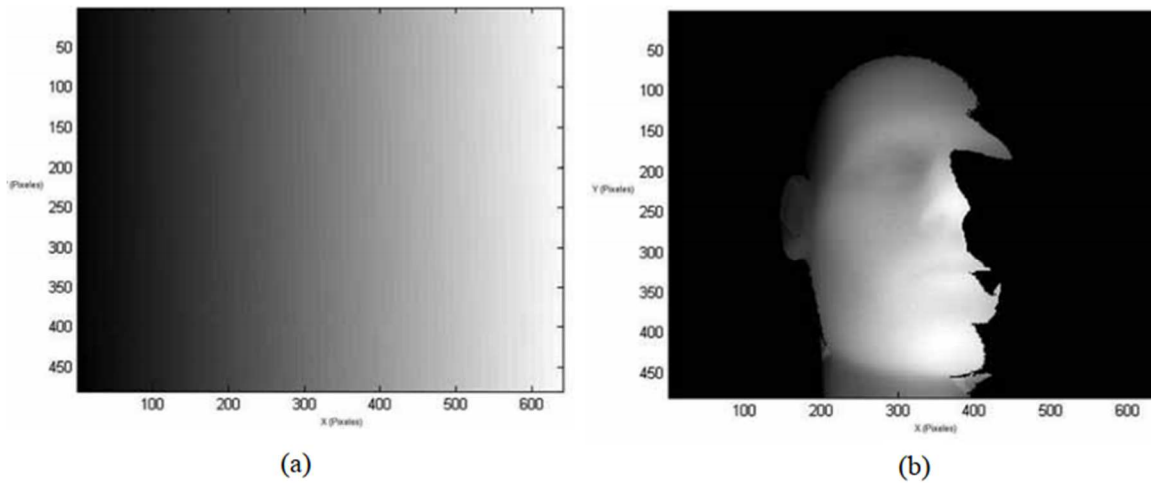


Figura 23 Imagen con fase desenvuelta o continua en plano de referencia y objeto.

Ésta técnica permite realizar reconstrucciones muy rápidas y con alta calidad, pero debemos siempre realizar una toma del plano de referencia para obtener la profundidad a través de una resta entre las fases desenvueltas.

### 2.7.1.2 Código binario

La técnica de proyección de luz estructurada con patrones binarios consiste, básicamente, en codificar una secuencia de líneas horizontales y verticales que permite identificar cada pixel correspondiente a las dimensiones de la resolución del proyector en un arreglo de visión estéreo. Por ejemplo, con un proyector con una resolución nativa de 1280x720, tenemos una proyección de 921,600 pixeles, los cuales pueden ser identificados de manera única en ambas imágenes de las cámaras.

Las proyecciones verticales identifican a la coordenada  $x$  del pixel y las proyecciones horizontales identifica a la coordenada  $y$  del mismo pixel en el plano de la imagen en ambas cámaras.

Podemos obtener el número de patrones horizontales mediante  $n = \log(1280) / \log(2)$ , y para el número de patrones verticales mediante  $m = \log(720) / \log(2)$ . Quedando  $n=10$  y  $m=10$  que es el número de bits necesarios para obtener 1024 y 720.

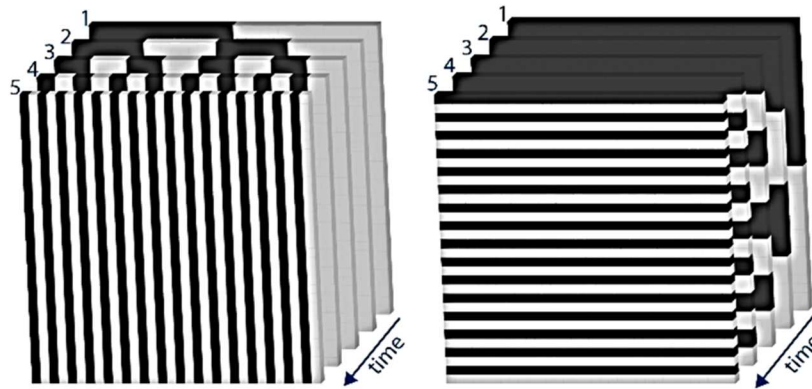


Figura 24 Patrones binarios horizontales y verticales.

En la Figura 24, se muestran los patrones que se van a proyectar, el color en negro es un bit en 0 y el color blanco es un bit en 1. La imagen captada identifica a cada pixel con un 1 o un 0 (Figura 25).

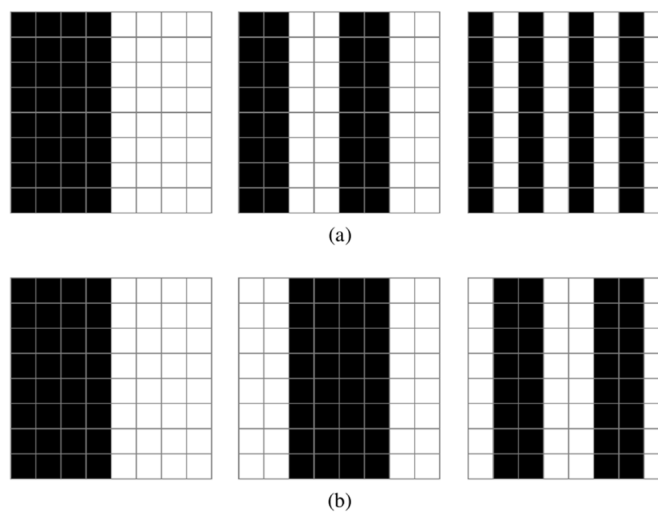


Figura 25 Píxeles en el plano de la imagen con el patrón binario proyectado.

El objetivo es que, para cualquier pixel  $p$  del plano de la imagen se genere un número binario con la concatenación de los patrones captados.

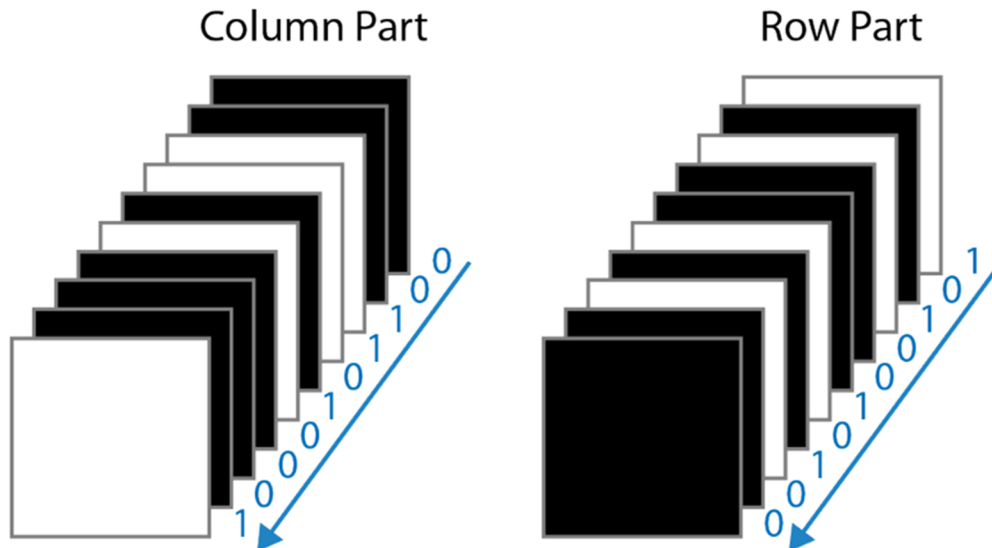


Figura 26 Codificación binaria de un pixel  $p$  en el plano de la imagen.

Se determina entonces  $B_{c_p} = 1000101100$  un número binario para el valor de columna del pixel  $p$ , y del mismo modo,  $B_{r_p} = 0010100101$  un número binario para el valor del renglón del mismo pixel, dando como resultado la identificación del pixel proyectado en coordenadas  $(x,y) = (209,660)$ .

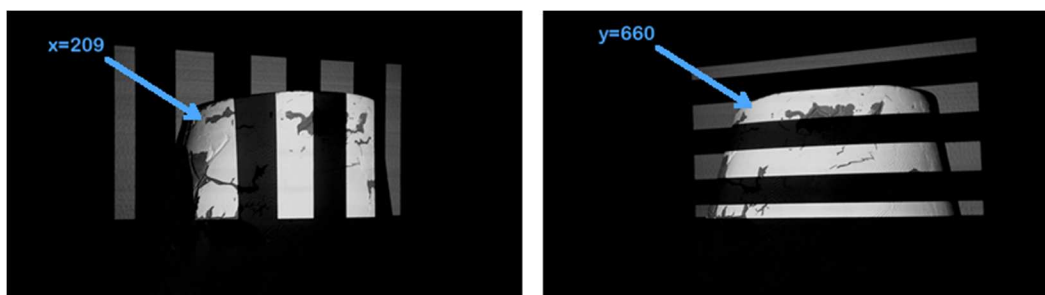


Figura 27 Ubicación de pixel  $p$  de proyección.

Habiendo identificado las coordenadas del pixel la posición (209,660) corresponde a un número decimal correspondiente al número de pixel ordenado de izquierda a derecha y de arriba hacia abajo, que sería en este caso 843,729.

## 2.8 Triangulación

La triangulación es el último paso en el proceso de reconstrucción, y consiste en recuperar el valor de X en  ${}^w p = [X, Y, Z, 1]$ .

Dado un pixel en  $I_1$  (Imagen 1) y su correspondiente pixel en  $I_2$  (Imagen 2), se forman dos líneas que intersectan en un punto que puede ser calculado. En la práctica, la triangulación de dos líneas en el espacio 3d se puede considerar como un problema mal planteado (ill-posed), ya que con frecuencia las dos líneas no se cruzan exactamente en el punto a reconstruir, sino que se pasan uno al lado del otro en estrecha proximidad.

Para superar esta limitación, se calcula el segmento perpendicular a las dos líneas con la distancia más corta, y el punto medio del segmento se considera el punto de intersección. La Figura 28, muestra un ejemplo donde el segmento ab es la línea más corta que conecta las dos líneas A y B. El punto medio p se considera el punto de intersección.



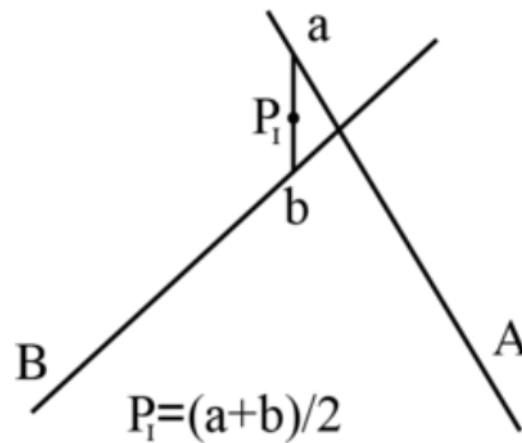


Figura 28 Cálculo del punto de intersección.

Considerar ahora las dos líneas A,B que intersectan en un espacio 3d y que pasan por los puntos  ${}^c p$  y  ${}^c q$ , con sus respectivos vectores  $\vec{u}$  y  $\vec{v}$ . Se definen a y b como:

$$a = p + s * \vec{u} \quad (9)$$

$$b = q + t * \vec{v} \quad (10)$$

El segmento que conecta a y b es perpendicular a las líneas, por lo tanto, el producto escalar de sus vectores es igual a 0 de la siguiente manera:

$$(a - b) \cdot \vec{u} = 0 \quad (11)$$

$$(a - b) \cdot \vec{v} = 0 \quad (12)$$

Entonces, de las ecuaciones (9) y (10), las ecuaciones (11) y (12) se convierten en:

$$\vec{w} \cdot \vec{u} + s * \vec{u} \cdot \vec{u} - t * \vec{v} \cdot \vec{u} = 0 \quad (13)$$

$$\vec{w} \cdot \vec{v} + s * \vec{v} \cdot \vec{u} - t * \vec{v} \cdot \vec{v} = 0 \quad (14)$$

Donde

$$\vec{w} = p - q \quad (15)$$

De las ecuaciones (13) y (14) se deduce que:

$$s = \frac{\vec{w} \cdot \vec{u} * \vec{v} \cdot \vec{v} - \vec{u} \cdot \vec{v} * \vec{w} \cdot \vec{v}}{\vec{v} \cdot \vec{u} * \vec{v} \cdot \vec{u} - \vec{v} \cdot \vec{v} * \vec{u} \cdot \vec{u}} \quad (16)$$

$$t = \frac{\vec{v} \cdot \vec{u} * \vec{w} \cdot \vec{u} - \vec{u} \cdot \vec{u} * \vec{w} \cdot \vec{v}}{\vec{v} \cdot \vec{u} * \vec{v} \cdot \vec{u} - \vec{v} \cdot \vec{v} * \vec{u} \cdot \vec{u}} \quad (17)$$

El punto de intersección es el promedio de los 2 puntos finales del segmento de la siguiente manera:

$$P_i = \frac{(p + s * \vec{u}) + (q + t * \vec{v})}{2} \quad (18)$$

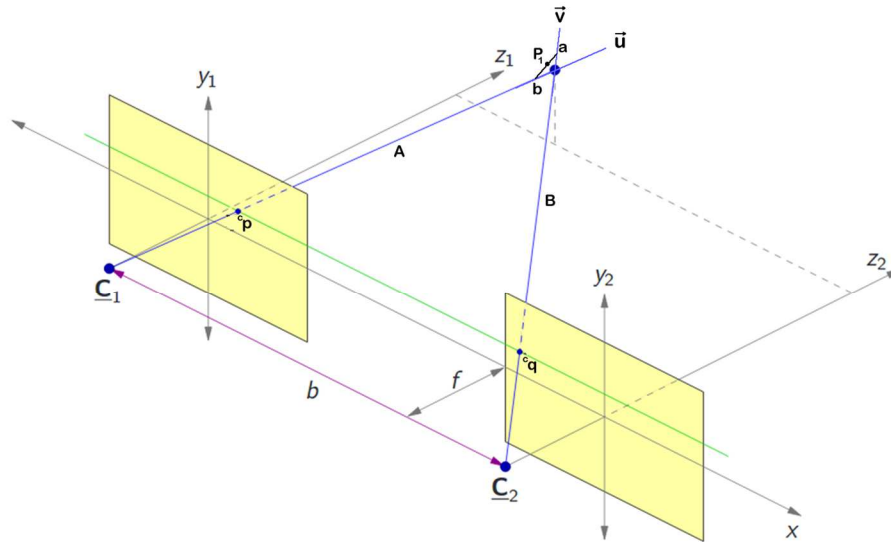


Figura 29 Esquema del método de triangulación por líneas.

$P_i$ , se calcula para todos los puntos correspondientes en las dos imágenes, es decir, se recupera el valor de  $Z$  en  ${}^w p = [X, Y, Z]$ .

## 2.9 Monitoreo y control en internet

La presente investigación, tiene como objetivo mostrar la integración de un trabajo de investigación con las tecnologías que dominan actualmente el modo en que se comparte la información.

Para tal efecto, se ha diseñado una arquitectura de hardware y software, que permite al usuario realizar reconstrucciones desde cualquier dispositivo conectado a internet.

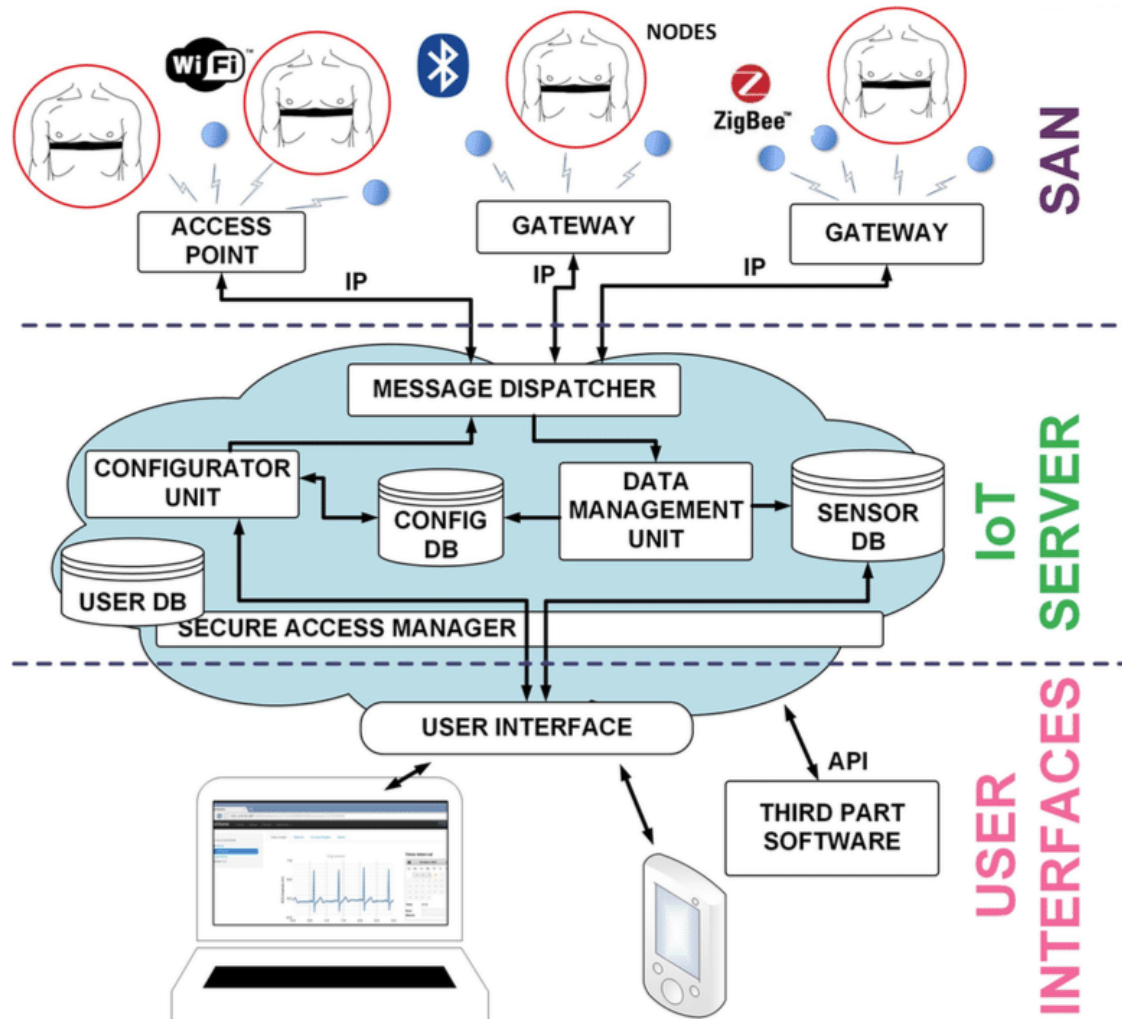


Figura 30 Arquitectura IoT.

En la Figura 30, se puede apreciar los componentes del sistema que se pretende desarrollar. Tal arquitectura se conoce como el internet de las cosas (IoT, por sus siglas en inglés).

El Internet of Things (IoT) describe la red de objetos físicos (cosas) que llevan sensores integrados, software y otras tecnologías con el fin de conectar e intercambiar datos con otros dispositivos y sistemas a través de Internet. Estos dispositivos abarcan desde objetos domésticos cotidianos hasta sofisticadas herramientas industriales. Con más de 7.000 millones de dispositivos de IoT

conectados en la actualidad, los expertos prevén que este número aumentará hasta llegar a 10.000 millones en 2020 y a 22.000 millones en 2025 (Oracle, 2020).

En la industria, la implementación de IoT va en aumento, pues la posibilidad de automatizar procesos, monitorearlos y tener un control de ellos a distancia. Genera grandes ahorros, además que puede implementarse en procesos de alto riesgo para una persona.

IoT industrial (IIoT) hace referencia a la aplicación de la tecnología de IoT en ámbitos industriales, en especial con respecto a la instrumentación y el control de los sensores y dispositivos que participan en las tecnologías en la nube. Recientemente, las industrias han utilizado la comunicación entre máquinas (M2M) para lograr la automatización y el control inalámbricos. Pero con el surgimiento de la nube y tecnologías relacionadas (como la analítica y el machine learning), las industrias pueden lograr una nueva capa de automatización y crear con ella nuevos ingresos y modelos de negocio. A veces, IIoT se denomina la cuarta ola de la revolución industrial o Industria 4.0. A continuación se indican algunos usos habituales de IIoT:

- Fabricación inteligente
- Mantenimiento preventivo y predictivo
- Redes eléctricas inteligentes
- Ciudades inteligentes
- Logística conectada e inteligente
- Cadenas de suministro digitales inteligentes



Figura 31 Aplicación de lot en la industria.

En (Kanagachidambaresan et al., 2020) una computadora de placa única (SBC) se utiliza para resolver el proceso obligatorio. El SBC se usa para controlar si las opciones de conmutación deben habilitarse o deshabilitarse en la fuente de alimentación de una máquina grande después de verificar que el trabajador esté usando todos los tipos de equipos de seguridad necesarios. Utilizando Deep Learning, se realiza un procesamiento de imágenes para detectar los objetos de seguridad, como casco, guantes, chaleco, etc. El lenguaje de programación python se utiliza para acceder al proceso anterior con la cámara para el procesamiento de imágenes. En cada nivel de transformación, los datos de entrada serán un poco más abstractos y un modo compuesto de representación.

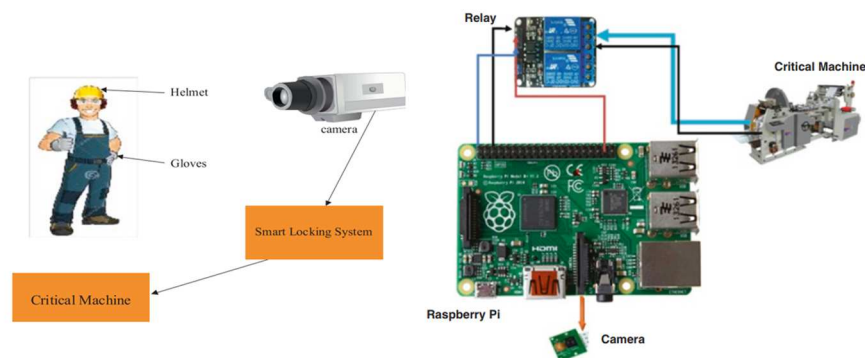


Figura 32 Aplicación de IoT en la industria.

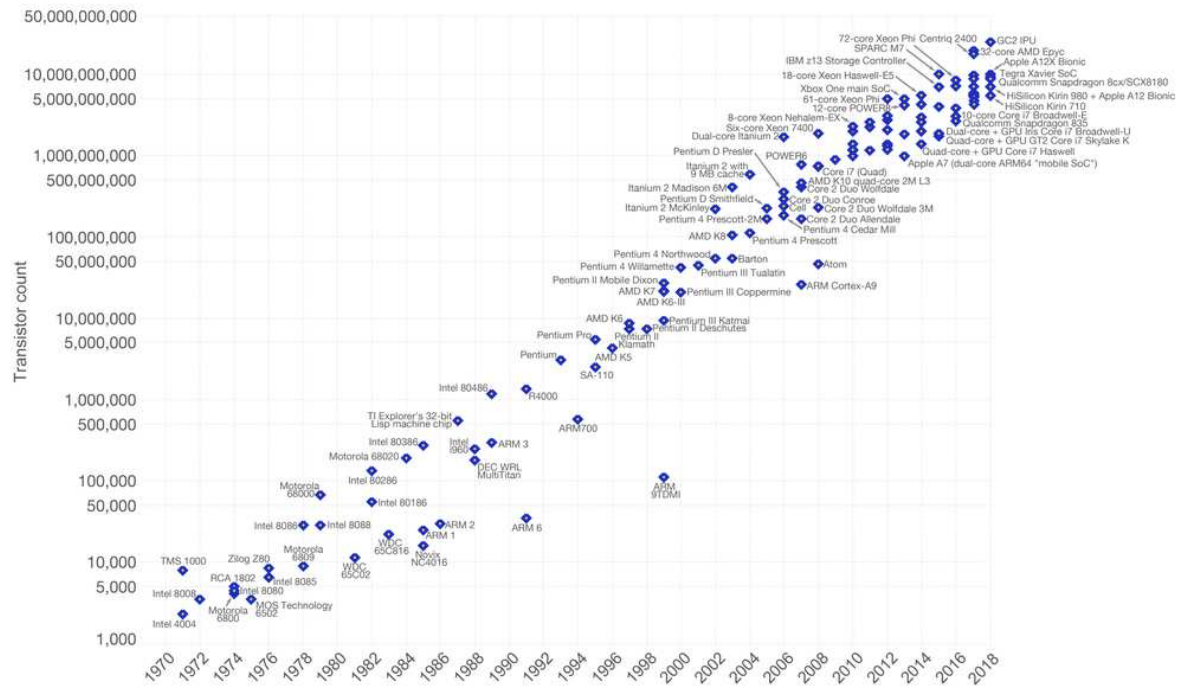
De este modo, se establece un entorno de trabajo seguro para los trabajadores utilizando el método de Deep Learning y el procesamiento de imágenes, que ayudan a garantizar la seguridad de los trabajadores en el entorno industrial. Además, ayuda a evitar accidentes graves en la industria y a reducir la mano de obra para el monitoreo, y también se logra este método a un costo mínimo.

## 2.10 Hardware en pequeña escala

Se sabe que es una tendencia el incremento de poder de procesamiento en procesadores. La ley de Moore es la observación de que el número de transistores en un circuito integrado denso se duplica aproximadamente cada dos años. Con lo cual, aumenta el poder de procesamiento.

### Moore's Law – The number of transistors on integrated circuit chips (1971-2018)

Moore's law describes the empirical regularity that the number of transistors on integrated circuits doubles approximately every two years. This advancement is important as other aspects of technological progress – such as processing speed or the price of electronic products – are linked to Moore's law.



Data source: Wikipedia ([https://en.wikipedia.org/wiki/Transistor\\_count](https://en.wikipedia.org/wiki/Transistor_count))

The data visualization is available at OurWorldinData.org. There you find more visualizations and research on this topic.

Licensed under CC-BY-SA by the author Max Roser.

Figura 33 Ley de Moore, incremento del número de transistores cada 2 años.

Aunque en una reciente entrevista (Courtland, 2015) al creador de esta ley, Gordon Moore, comenta que esta ley podría llegar a su fin entre los años 2020 o 2022, y explica que existen 3 principales obstáculos técnicos: la velocidad de la luz, naturaleza atómica de los materiales y costos crecientes.

Sin embargo, estos tres obstáculos, hacen que se busquen nuevas alternativas que no consideren electrones y transistores hechos de silicio.

Una de las alternativas, es el uso de las GPU y FPGA. Las GPU ya se han utilizado para el entrenamiento de IA durante más de una década. En los últimos años, Google ha introducido el término TPU (Google, 2019) con el impulso de la inteligencia artificial en mente, y en este momento hay más de 50 empresas que fabrican chips de inteligencia artificial. Los FPGA están siendo cada vez más usados, especialmente con el uso en centros de datos como lo hicieron Intel y Microsoft (Intel & Microsoft, 2020). Microsoft también usó FPGA para acelerar la búsqueda de Bing.

Otra alternativa, es el desarrollo de aplicaciones con algoritmos “ligeros” en computadoras de placas reducidas como las raspberry pi, también llamadas computadoras de placa simple (SBC, por sus siglas en inglés), desarrollada en el Reino Unido por la Raspberry Pi Foundation, que tienen la potencia suficiente para ejecutar en poco tiempo los algoritmos desarrollados en la presente investigación.

Estas placas, han venido evolucionando, incorporando mayor capacidad de procesamiento en un tamaño relativamente pequeño. En la Figura 34, se muestran diferentes pruebas realizadas a la placa raspberry en diferentes versiones (Halfacree, 2019).



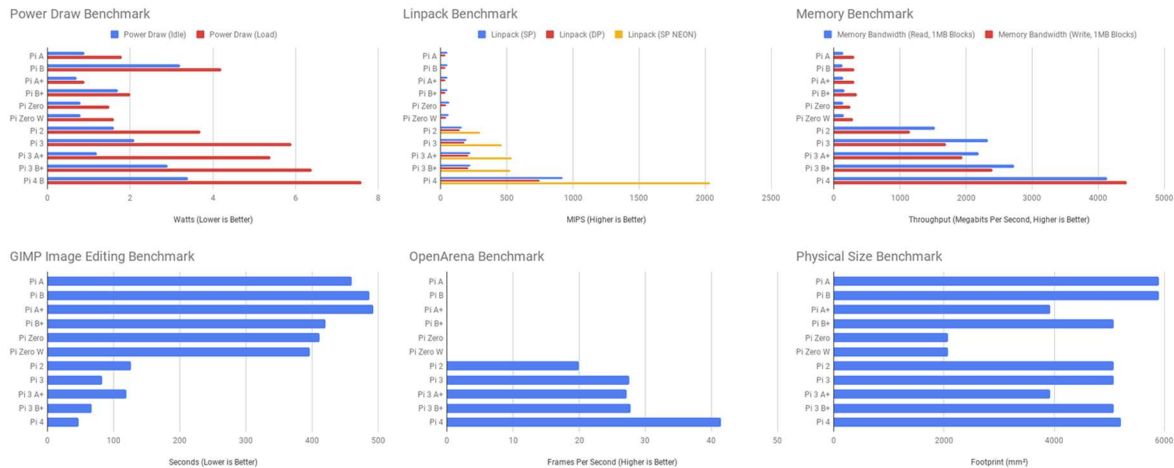


Figura 34 Comparación de versiones de raspberry en diferentes pruebas de benchmarking.

De este modo, se puede decir que, la selección de una determinada tecnología de hardware depende del proyecto a desarrollar. Para este caso, el desarrollo de un dispositivo compacto con la implementación de algoritmos de procesamiento ligero, las placas tipo raspberry, cumplen con los recursos de hardware necesarios.

## 2.11 Software utilizado

La elección del software a utilizar, para este caso, va de acuerdo a las necesidades del proyecto y las limitaciones del hardware. Se ha diseñado una arquitectura de software, que utiliza librerías ligeras implementadas en un sistema operativo sin interfaz gráfica. De tal modo, que queden disponibles la mayoría de los recursos de hardware (cpu, memoria ram, principalmente), para la ejecución de los procesos de reconstrucción 3d.

- Angular Framework, es conjunto de herramientas con los lenguajes de programación HTML, CSS y Javascript para crear aplicaciones eficientes y sofisticadas de una sola página. Se utiliza para la creación de la interfaz entre usuario y algoritmos de reconstrucción 3d.

- Django Framework, es un conjunto de herramientas basado en el lenguaje de programación Python. Se utiliza para codificar las funciones de la interfaz de programación abierta (API), que utiliza la interfaz de usuario.
- Python, lenguaje de programación interpretado, dinámico y multiplataforma de alto nivel con extensas librerías para investigaciones de carácter científico.
- OpenCV, es una de las librerías más utilizadas para el desarrollo de proyectos de visión por computadora. Ofrece una interfaz entre el lenguaje de programación Python y las funciones propias de OpenCV desarrolladas en C y C++, lo que permite generar código más entendible y ejecuciones más rápidas.
- SQLite, es una base de datos ligera, que se utiliza, en este caso, para guardar información de reconstrucciones y hacer diferentes análisis con el fin de mejorar los algoritmos de reconstrucción.
- Ubuntu 18.04.4 LTS Server de 64 bits, Sistema operativo que soporta la arquitectura del software.

## **2.12 Aplicaciones de la reconstrucción tridimensional a partir de imágenes con luz estructurada.**

Hoy en día, las aplicaciones con la técnica de luz estructurada son muy amplias, el 90% de las aplicaciones comerciales usan ésta técnica. Debido a sus ventajas a otro tipo de técnicas de reconstrucción, como: no requerir de contacto directo con el objeto, rápida adquisición de captura con gran precisión, escaneo de campo completo, fácil de automatizar, portabilidad y seguridad para el usuario al emitir luz no dañina, hacen de ésta técnica la más utilizada para la solución de problemas de

medición de superficies y análisis de la propia topografía de prácticamente cualquier objeto.

### 2.12.1 Medicina

El campo médico tiene un amplio uso de diferentes escaneos como rayos X, tomografía computarizada, resonancia magnética y ultrasonido. Estas técnicas son bastante útiles para proporcionar información sobre los órganos internos. Sin embargo, hay una brecha para obtener información de las partes externas del cuerpo que ahora pueden ser atendidas por las tecnologías de escaneo 3d. Se ha estudiado una gran cantidad de trabajos de investigación sobre diferentes técnicas de escaneo y escaneo 3d para identificar el impacto de ellos y la investigación posterior (Haleem & Javaid, 2019).

Aplicaciones como (Medit, 2020), ofrecen diferentes tipos de escáneres en el área de odontología. Es una herramienta de muy útil para revisar el estado de toda la dentadura del paciente.

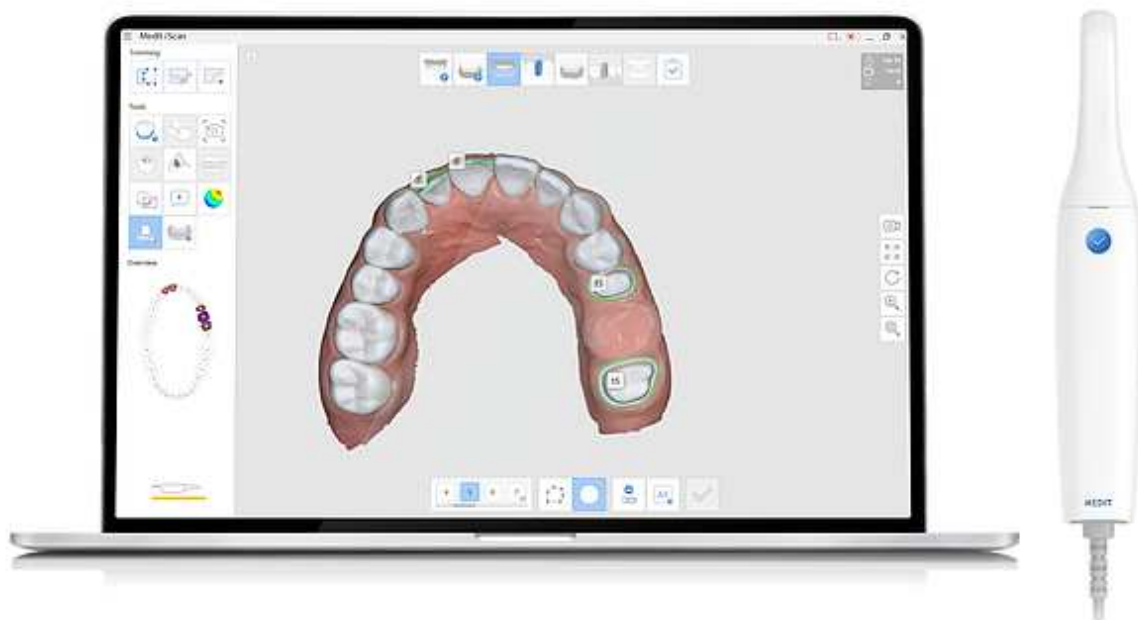


Figura 35 Medit i500 escaner 3d para odontología.

Entre sus principales características, se encuentran: comunicación abierta para exportación de archivos STL, alta velocidad de escaneo, alta precisión, ligero y con software de pos procesamiento.

Artec (Artec, 2020), es una empresa que ofrece aplicaciones de escaneo 3d para el cuidado de la salud, entre sus principales aplicaciones en el área médica están: creación de prótesis ortopédicas, análisis para cirugías plásticas, e ingeniería inversa, como la creación de sillas de ruedas a medida.



Figura 36 Ingeniería inversa para la fabricación de sillas de ruedas a medida.

La ingeniería inversa en la reconstrucción 3d, consiste en construir un modelo 3d a partir de una pieza u objeto, en este caso, el objeto es parte de las piernas y dorso de un paciente; con el modelo construido, se puede imprimir el asiento y respaldo del paciente.

### 2.12.2 Industria

El campo industrial, también tiene un amplio uso de la tecnología de reconstrucción 3d. Una de las aplicaciones más usadas es para la ingeniería inversa para fabricación de piezas a partir de modelos obtenidos.

Durante más de 14 años, STYL & TECH ofrece servicios completos de ingeniería inversa. Este servicio comienza con el proceso de escaneo 3d; Luego se producen informes de inspección y mapas de color, maquetas virtuales, superficies de

reconstrucción en 3d y una nueva integración de diseño. Después de una ingeniería inversa, también se ofrecen servicios de creación de prototipos para fabricar nuevas piezas y ensamblajes (Styl&Tech, 2020).

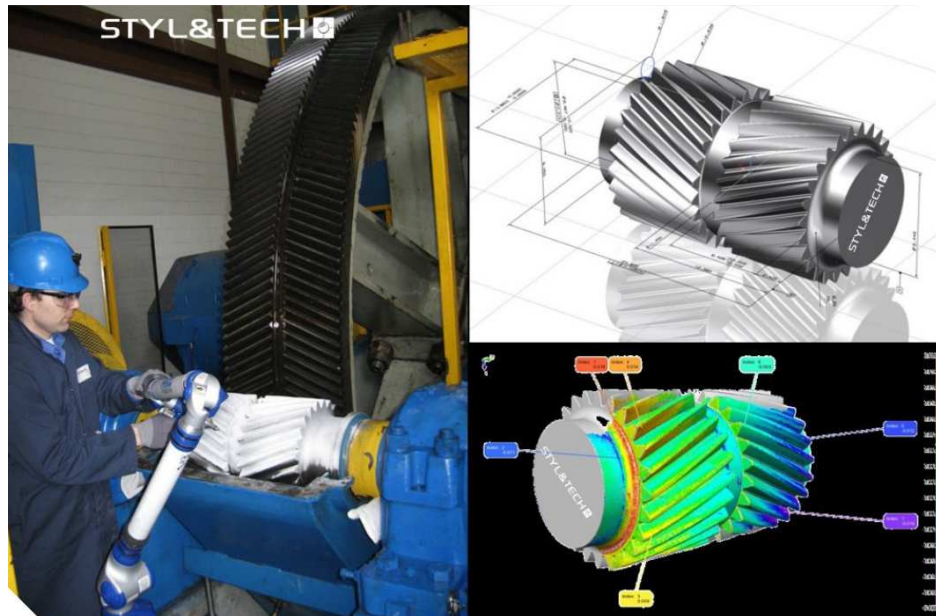


Figura 37 Ingeniería inversa en la industria (Styl&tech).

Dependiendo de la aplicación, también se pueden obtener modelos paramétricos en superficies. El control de calidad y las pruebas son una parte importante e integral del proceso de fabricación de superficies. Las superficies se estudian para buscar signos de corrosión, grietas, impurezas, abolladuras y poros.



Figura 38 Detección de errores en líneas de producción (Keyence, 2020).

### 2.12.3 Robótica

En el campo de la robótica, existen diversas aplicaciones, solo se mencionaran dos de ellas. La primera, tiene que ver con el posicionamiento y reconocimiento de objetos.

FANUC 3d Vision permite a los robots reconocer y realizar el picking de objetos situados de manera aleatoria directamente desde un contenedor. Automatizar los procesos por medio de esta tecnología de reconocimiento de objetos representa una manera eficiente de aumentar la productividad y reducir los costes en muy diversas aplicaciones generales de manipulación de materiales. Ya se trate de objetos sueltos, mezclados e irregulares, o de bolsas y sacos, equipar a un robot con un 3d Vision constituye una solución de manipulación eficaz que se puede adaptar rápidamente para gestionar distintos productos de acuerdo con los requisitos de su cartera de pedidos. Con garras adecuadas para todas las posibles eventualidades de manipulación, ofrecen la misma flexibilidad para aplicaciones de despaletizado, bin picking, clasificación y alimentación de máquinas, con todo tipo de objetos (Fanuc, 2020).



Figura 39 Manipulación de materiales automatizada mediante un sistema de visión (FANUC).

La segunda tiene que ver con la creación de mapas tridimensionales creados a partir de la ruta que va siguiendo un robot.

En (Gunatilake et al., 2019), se utiliza un proyector de anillo láser infrarrojo estructurado (IR) y un sistema de cámara estéreo para generar el perfil 3d de la tubería a medida que el robot se mueve dentro de la tubería.

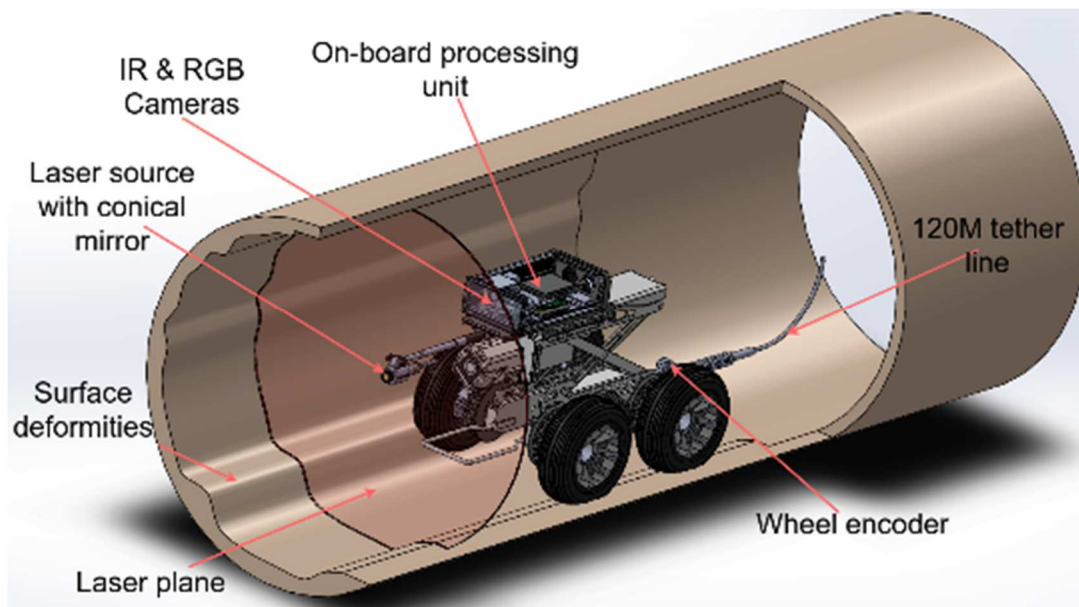


Figura 40 Generacion de un mapa 3d de una tuberia.

El sistema propuesto tiene la capacidad de generar una nube de puntos 3d en tiempo real mientras atraviesa la tubería. El mapa 3d en tiempo real de la tubería se genera utilizando el framework ROS (ROS, 2020), realizando pruebas para determinar la efectividad de los mapas 3d.

## 2.12.4 Reconstrucción de edificios y reliquias

Entre las diferentes aplicaciones, está la reconstrucción de edificios, generando modelos 3d con las proporciones reales de la construcción. Se ha empleado en ésta área para la reconstrucción de edificios históricos o ciudades enteras.



Figura 41 Reconstrucción 3d de edificios.

También, una de las aplicaciones más usadas, es para la conservación de reliquias culturales en formato digital, pues es obvio que, con el paso del tiempo, el material de las reliquias se va degradando y perdiendo su forma original (Figura 42).



Figura 42 Reconstrucción 3d de reliquia (Faro).



Actualmente, con las nuevas tecnologías informáticas, se pueden crear museos virtuales como sitios web, donde se puede visitar y ver a detalle piezas históricas de todo el mundo (Figura 43).



Figura 43 Museo 3d en línea de reliquias de Vietnam.



Figura 44 Detalle de pieza histórica de museo 3d.

En la Figura 44, se puede observar la calidad del detalle en la reconstrucción 3d.

## 2.12.5 Entretenimiento

3d FASCINATION, es una empresa que ofrece la reconstrucción 3d a través de un escáner basado en la proyección de luz estructurada con cámara de visión en 360°. Ofrecen tanto la reconstrucción y la impresión de los modelos 3d obtenidos para diferentes usos comerciales (3dFASCINATION, s/f).

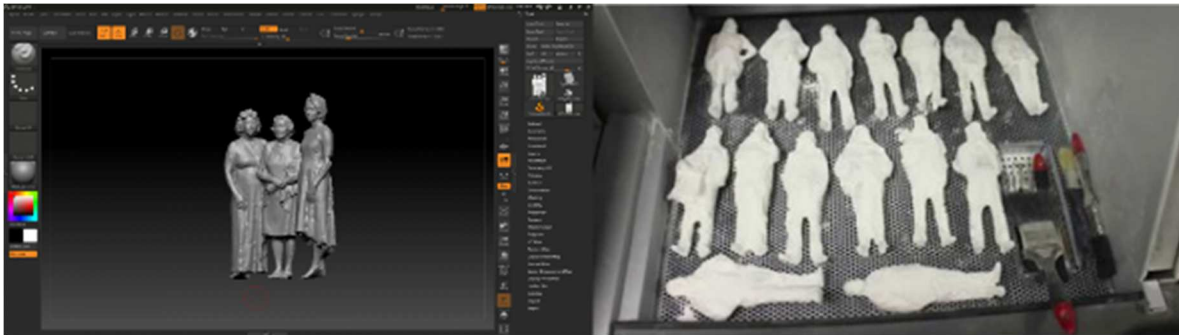


Figura 45 Reconstrucción 3d de personas e impresión 3d.

## CAPÍTULO 3 DESARROLLO TEÓRICO/EXPERIMENTAL

### 3.1 Componentes del sistema

#### 3.1.1 Módulos del sistema

El funcionamiento del sistema consta, básicamente, de 3 módulos: 1) Procesamiento local de imágenes; 2) Comunicación entre escáner 3d y computadora de usuario y 3) Interfaz de usuario.

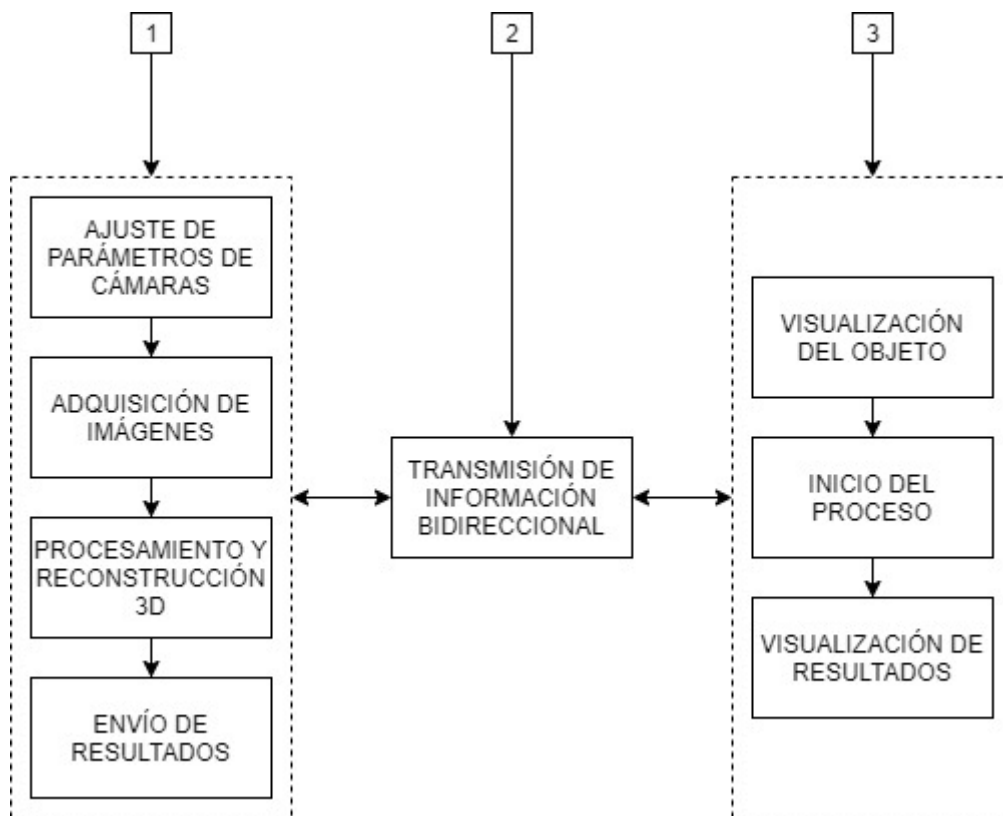


Figura 46 Diagrama de bloques del funcionamiento general del escáner 3d.

Se ha diseñado un funcionamiento de comunicación bidireccional que permite al usuario interactuar con el escáner 3d de modo local y remoto. De tal forma que vaya acompañando el proceso de reconstrucción. En caso de repetir procesos, el usuario, tiene la opción de regresar a cualquier sub proceso del módulo 1 (Figura 46 ).

En la Figura 47, se puede observar diagrama de flujo global del proceso de reconstrucción 3d. Muestra la interacción del usuario con el escáner 3d. Nótese que, solo se presenta la secuencia de pasos como un sistema integral, la arquitectura del sistema se presenta más adelante.

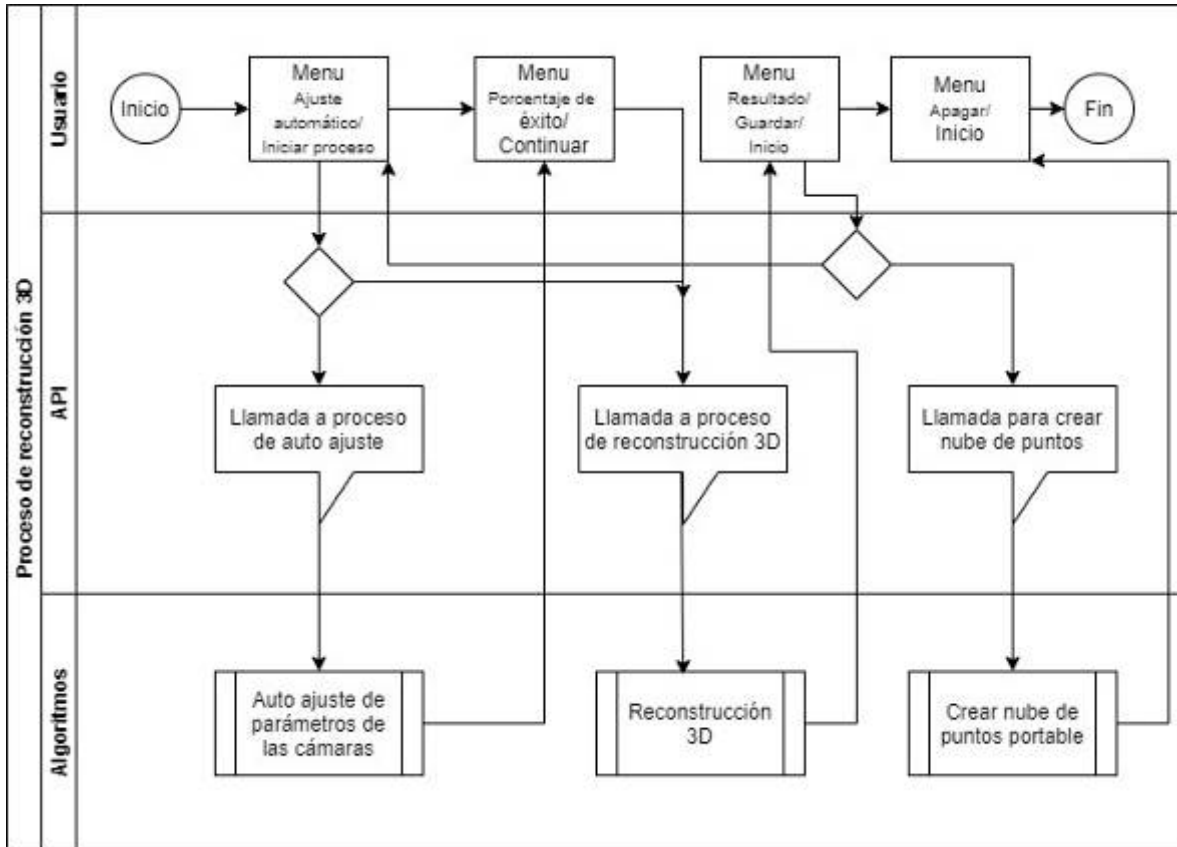


Figura 47 Proceso de reconstrucción 3d.

### 3.1.2 Comunicación escáner 3d - usuario

La comunicación se realiza a través del protocolo de aplicaciones HTTP, que corre sobre el protocolo de capa TCP/IP, que se establece mediante el estándar IEEE 802.11.b/g/n/ac wireless LAN. Permitiendo de este modo una velocidad de conexión, en modo local, de hasta 433 Mbps entre dispositivos con el estándar 802.11ac con un rango de operación de hasta 50 metros. En modo remoto, la velocidad de transferencia de datos dependerá del proveedor de servicios ISP (Internet Service Provider), pero la distancia es desde el sitio local donde se encuentra el escáner 3d, hasta cualquier punto del planeta con acceso a internet (Figura 48).

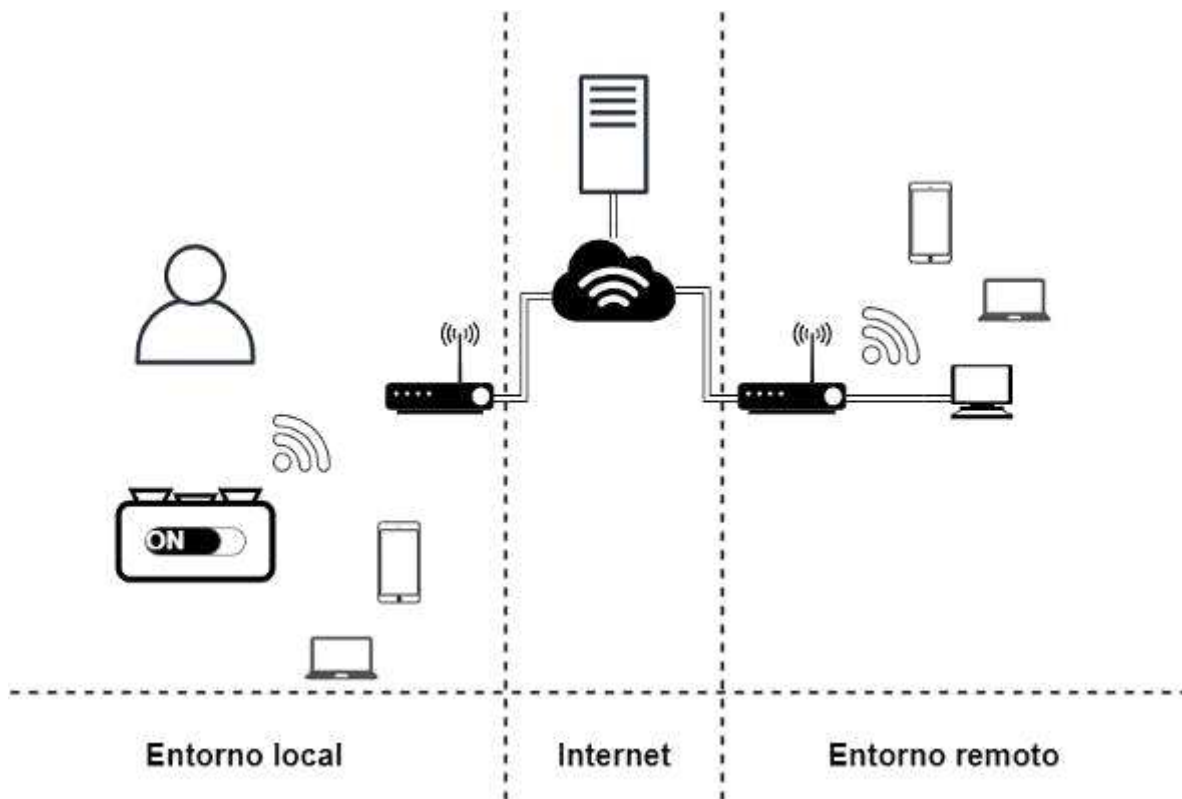


Figura 48 Comunicación en entorno local y remoto.

La comunicación para el modo remoto, hace uso de una aplicación desarrollada en la nube, conocida también como Cloud Computing, lo que permite enlazar dispositivos desde diferentes puntos de todo en todo el planeta a través de internet (Figura 48).

### 3.1.3 Configuración modo local

El escáner 3d, permite una configuración local o remota. El modo local, se utiliza cuando se va a realizar una reconstrucción en sitio teniendo en cuenta la distancia máxima de operabilidad del escáner 3d (50 metros máximos, 25 recomendados), no es necesaria la conexión remota o no hay disponibles puntos de acceso a internet.

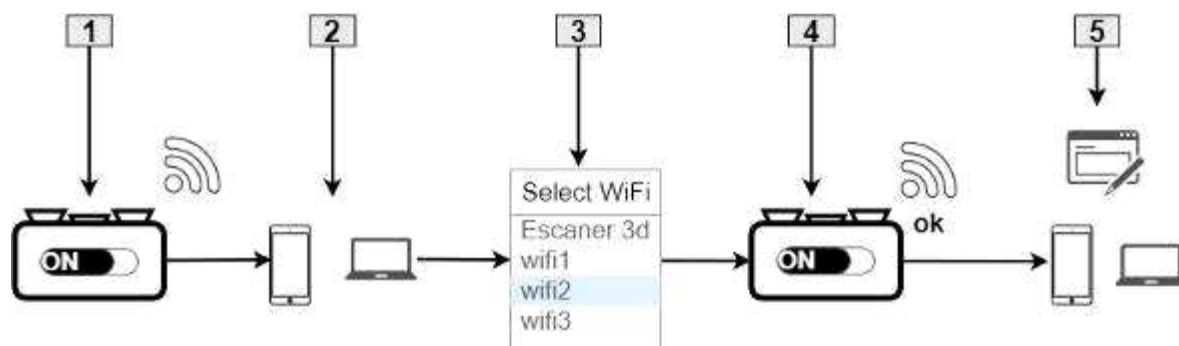


Figura 49 Configuración local del escaner 3d.

Consiste en 5 pasos:

1. Se enciende el escáner 3d desde su switch de encendido.
2. Con el sistema de cómputo, se detecta la red inalámbrica que genera el escáner 3d.
3. Se selecciona el nombre de la red o SSID Escaner 3d.
4. El escáner 3d se conecta con el sistema de cómputo del usuario.
5. Ofrece al usuario trabajar en modo local o remoto, para este caso, se escoge el modo local. Al seleccionar el modo local, el sistema re direcciona a la interfaz de usuario del escáner 3d para empezar a realizar reconstrucciones.

### 3.1.4 Configuración modo remoto

El modo remoto del escáner, permite al usuario tener un monitoreo y control del escáner 3d a través de internet. Éste concepto es ya conocido como el “Internet de las cosas” IoT (Internet of Things, por sus siglas en inglés), propuesto en 1999 por (Ashton, 2009), en el Auto-ID Center del MIT, en donde se realizaban investigaciones en el campo de la identificación por radiofrecuencia en red (RFID) y tecnologías de sensores. En (Oracle, 2020), se puede encontrar información actualizada sobre las aplicaciones de IoT.

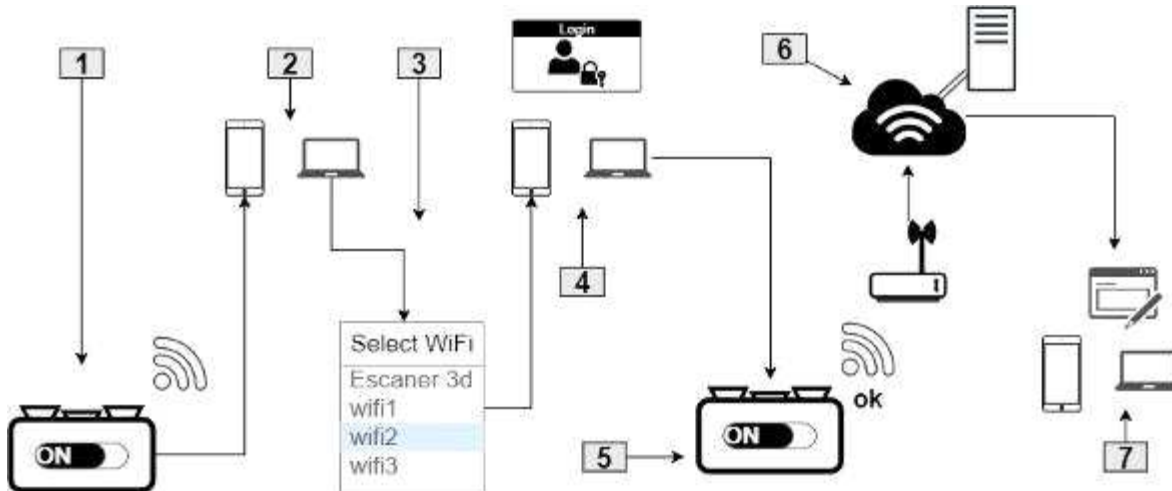


Figura 50 Configuración de acceso remoto.

Consiste en 7 pasos:

1, 2 y 3, son los pasos para la configuración local.

4. Al establecer la conexión entre el escáner 3d y el sistema de cómputo del usuario, se selecciona ahora el modo remoto. En este paso, el sistema pregunta a que red inalámbrica con acceso a internet se va a conectar, se introducen los datos de conexión y se finaliza la configuración en sitio.

5. Ahora, el escáner 3d, se ha conectado a internet y ha establecido una conexión con el servidor en la nube.
6. El servidor en la nube, tiene una interfaz de usuario, la cual se puede acceder mediante un nombre de dominio de internet, por ejemplo [www.ecaner3dCiicap.com](http://www.ecaner3dCiicap.com).
7. El usuario, ahora conecta su sistema de cómputo a un punto de acceso a internet y puede abrir la página en su navegador, se identifica con el servidor en la nube y una vez identificado como usuario válido, lo re direcciona a la interfaz de usuario para monitorear y controlar el escáner 3d.

### 3.1.5 Arquitectura de software del sistema

En la Figura 51, se muestra la comunicación en entorno local y remoto con el escáner 3d ya descritas. Dentro del escáner, se instalaron diversas tecnologías que permiten los dos tipos de comunicaciones. Para el modo local, se hace uso de una aplicación desarrollada con los lenguajes de programación HTML, CSS Y Javascript, con el fin de que el usuario a través del navegador de un sistema de cómputo personal, pueda interactuar con los algoritmos de reconstrucción.

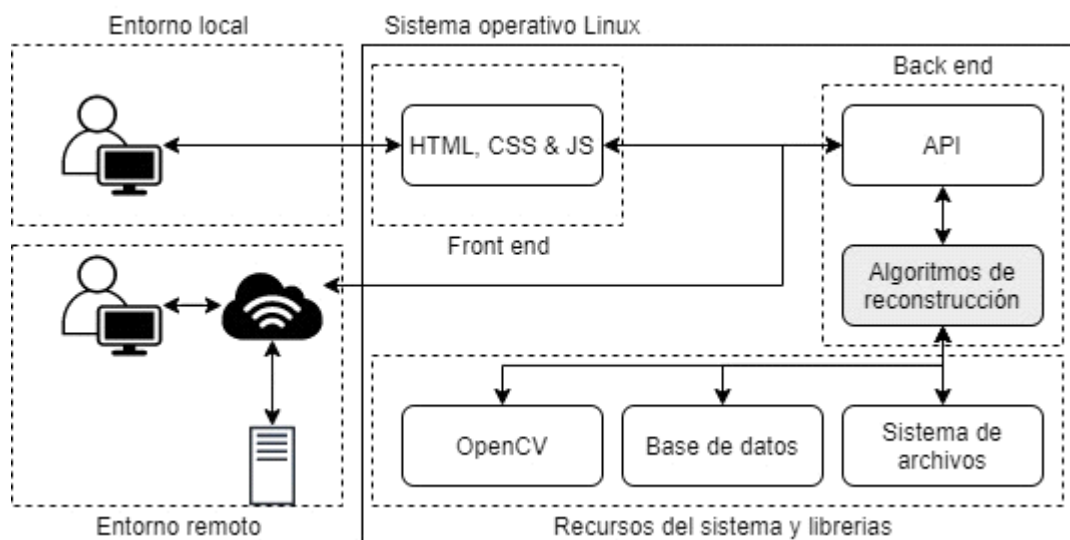


Figura 51 Arquitectura del software.



Para el entorno remoto, se ha diseñado una aplicación “Front end” en la nube. Ambas aplicaciones “Front end”, se comunican con una Interfaz de programación abierta o API (Application Programming Interface, por sus siglas en ingles). La API, se comunica con los algoritmos de reconstrucción. Tanto la API, como los algoritmos de reconstrucción, se han desarrollado en un Framework (Conjunto de funciones TI para un fin específico), basado en el lenguaje de programación Python, y ambos, componen la parte del “Back end”. Finalmente, el “Back end”, se comunica con las librerías para visión por computadora OpenCV, la base de datos SQLite y con el sistema de archivos del sistema operativo.

### 3.1.6 Componentes de hardware

Los componentes utilizados para la construcción del escáner 3d, son los siguientes:

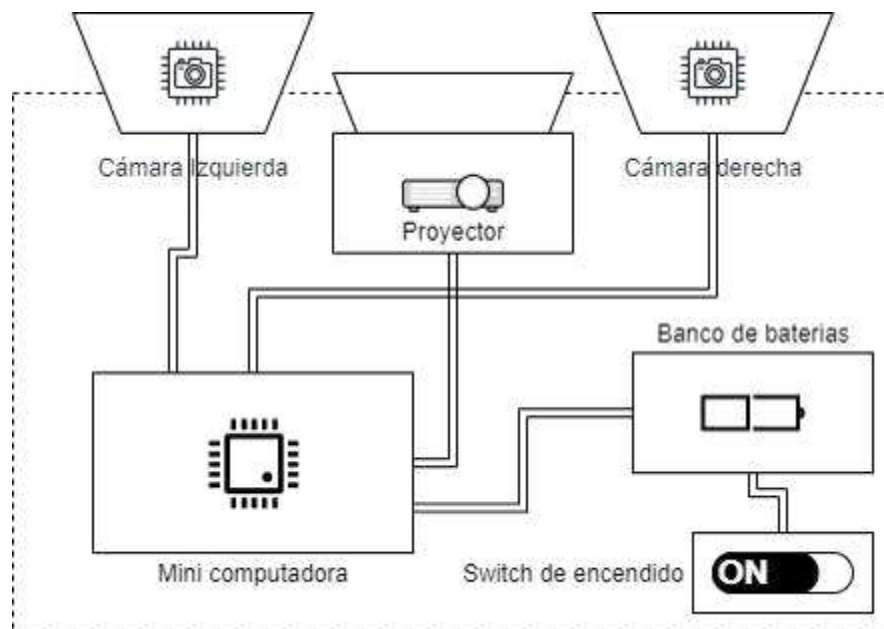


Figura 52 Componentes de hardware del escáner 3d.

1) Mini computadora raspberry pi 3 Modelo B+ con las siguientes especificaciones:

- Procesador Broadcom BCM2837B0, Cortex-A53 (ARMv8) 64-bit SoC @ 1.4GHz

- Memoria RAM 1GB LPDDR2 SDRAM
- Tarjeta inalámbrica 2.4GHz and 5GHz IEEE 802.11.b/g/n/ac
- USB 2.0 ports
- Puerto Micro SD con soporte clase 10 como disco duro.
- Entrada de energía 5V/2.5A en corriente continua

2) Proyector LG PH550G con las siguientes especificaciones:

- Resolución nativa de 1280x720 píxeles.
- Brillo de 550 lúmenes
- Enfoque manual
- Tamaño de imagen de 25" a 100"
- Fuente de luz tipo RGB LED
- Salida HDMI

3) Dos cámaras Logitech c920 con las siguientes especificaciones

- Resolución nativa de 1280x720 píxeles.
- Tamaño del píxel de 72 DPI.
- Enfoque automático controlado por software

- Conexión USB 2.0
- Objetivo Car Zeiss lens

4) Banco de baterías con las siguientes especificaciones:

- Baterías de ion litio
- Arreglo calculado para una autonomía del 2 hrs.

5) Circuito de conexión a CA y switch de encendido

### 3.2 Geometría del arreglo óptico

Para el diseño de la geometría del arreglo óptico del escaner3d, se consideraron las especificaciones técnicas del proyector y las cámaras.

Con el dato del proyector que puede realizar proyecciones en un plano en un rango de distancia que va desde los 63.5 cm a 254 cm. Se realizaron pruebas en las vistas de las cámaras para que pudiesen capturar imágenes en ese rango de distancia. Al mismo tiempo de que se fue comprobando que los centros de las cámaras se pudieran intersectar. Se desarrolló un programa para capturar en paralelo lo que ven las cámaras, se puede apreciar en la Figura 53, se trazaron líneas verticales y paralelas para obtener el centro de la imagen y así poder posicionar las cámaras para que todos los píxeles converjan en un punto  $P_i$ .

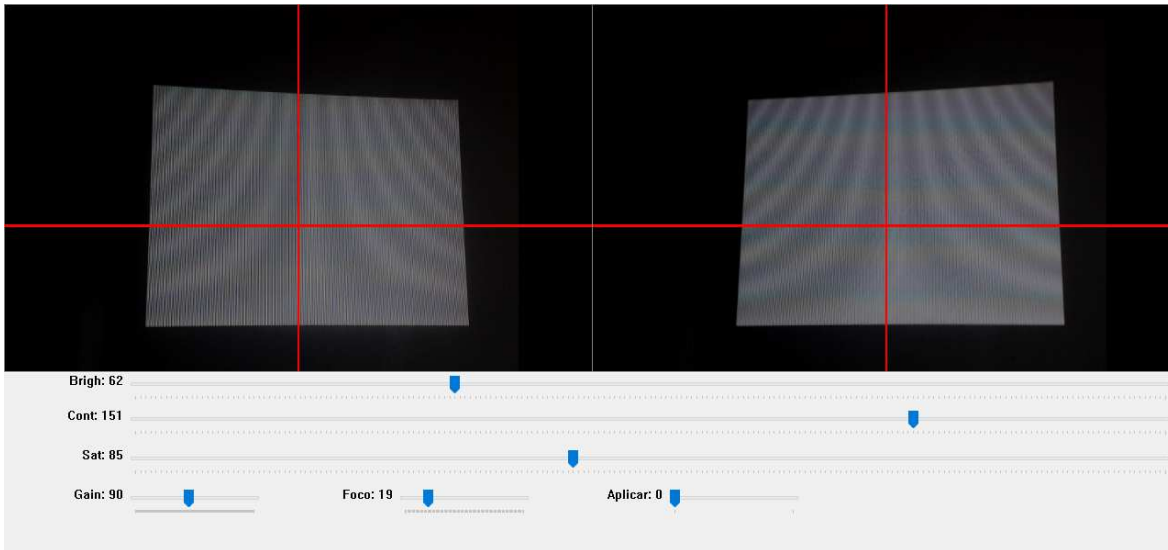


Figura 53 Pruebas para obtener geometría de visión estéreo.

Una vez realizadas las pruebas de visión de ambas cámaras para cubrir el rango de distancia del proyector. Se obtuvo un arreglo como en la Figura 54. Con el fin de entender cómo obtener la información de profundidad de un determinado pixel, se trazan con líneas rojas, a modo de representación teórica, las líneas del centro óptico. También se puede observar que en la geometría de visión estéreo como la de la Figura 29, están presentes los ángulos de rotación en cada cámara, una distancia entre ambas cámaras o la traslación, y la distancia máxima en que pueden intersectar los centros ópticos de ambas cámaras.

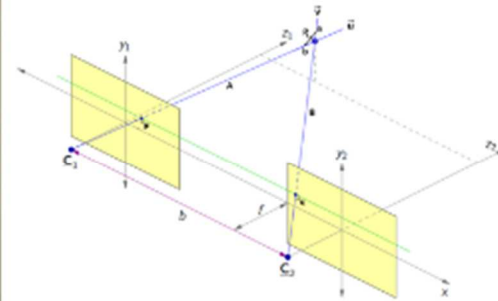
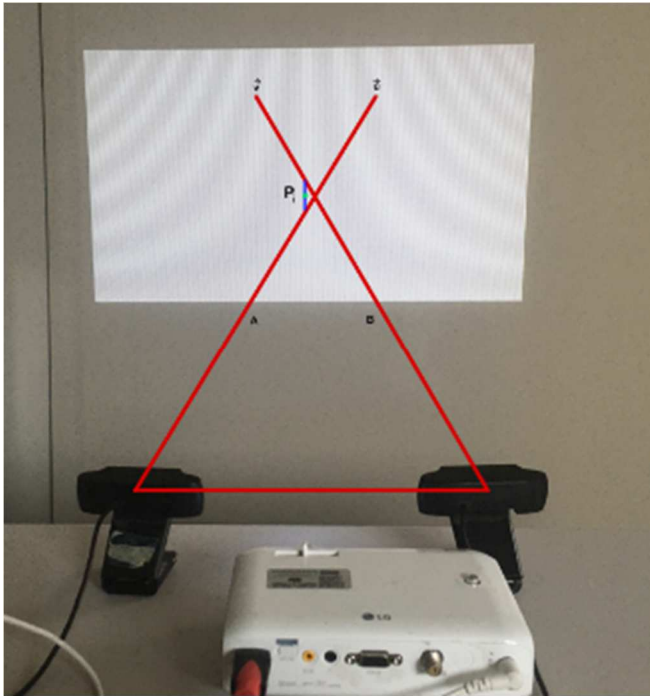


Figura 54 Geometría estéreo en arreglo de cámaras.

Teniendo este arreglo, ahora podemos ajustar los parámetros de las cámaras.

### 3.3 Ajuste automático de parámetros de cámaras

La mayoría de cámaras para propósitos de estudio científico, vienen acompañadas con sus controladores para el sistema operativo con el que se esté trabajando. Estos controladores, contienen funciones para cambiar los parámetros de la cámara.

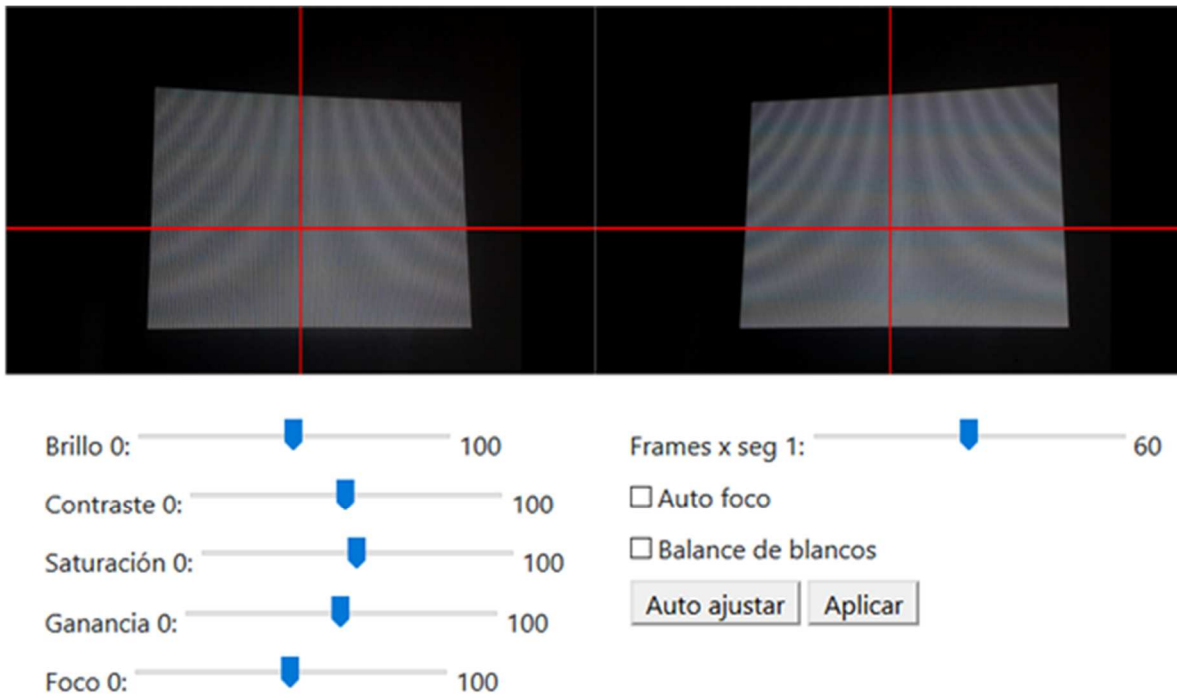


Figura 55 Ajuste de parámetros de cámaras.

Ya que el desarrollo del escáner 3d, se realizó sobre el sistema operativo Linux, los controladores que vienen con Linux, permiten acceder a varios de los parámetros que necesitamos para obtener los mejores resultados que nos permitan los ajustes de parámetros.

Se desarrolló un algoritmo para ajustar los parámetros de las cámaras de manera automática o manual. El cual, aplica valores de configuración haciendo un análisis a imágenes capturadas por las cámaras, haciendo un barrido aleatorio a diferentes

líneas en el eje de las x de cada imagen. Obteniendo un porcentaje de píxeles en blanco y negros. Entre mayor es el porcentaje, mejor será la reconstrucción 3d.

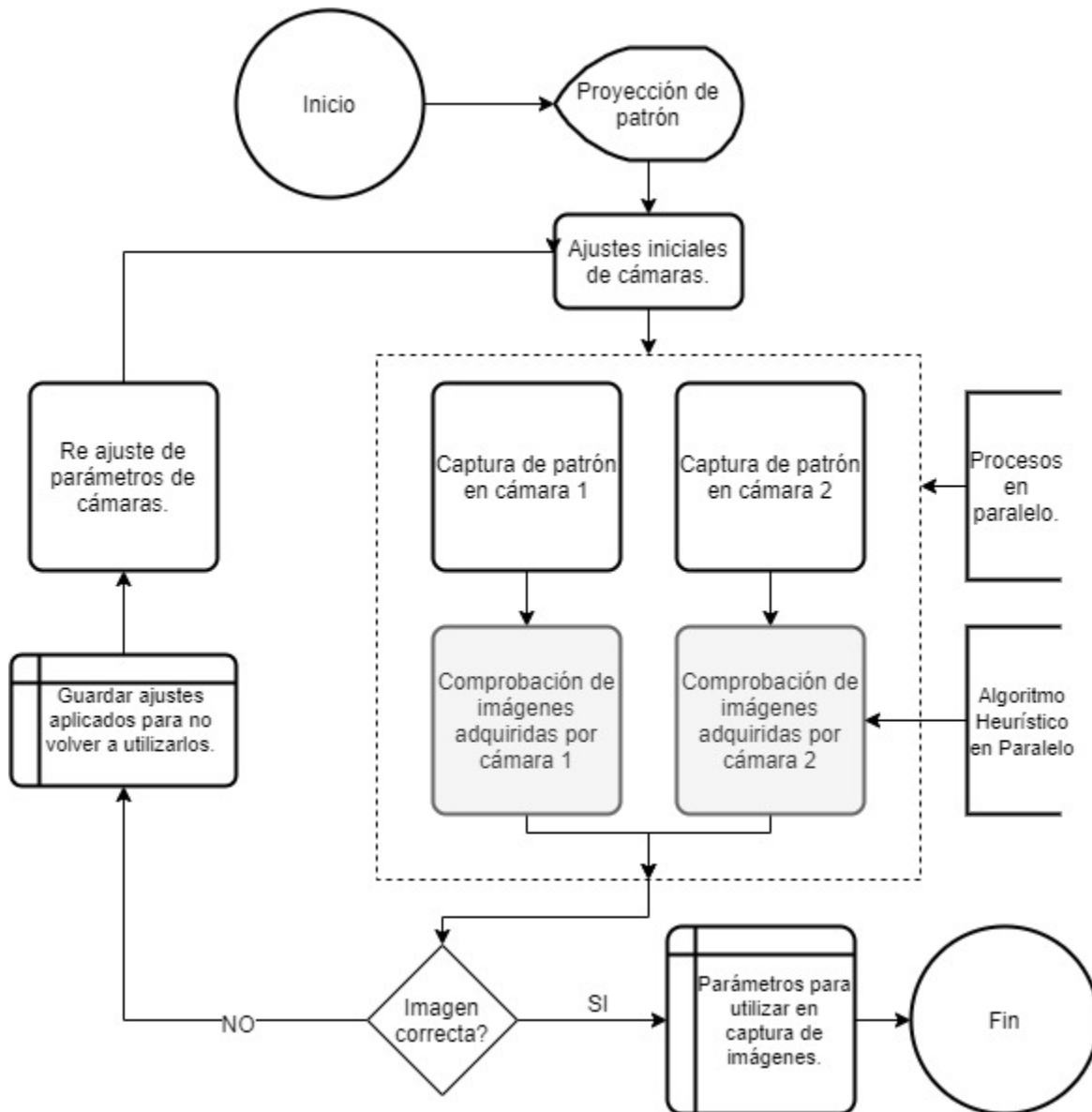
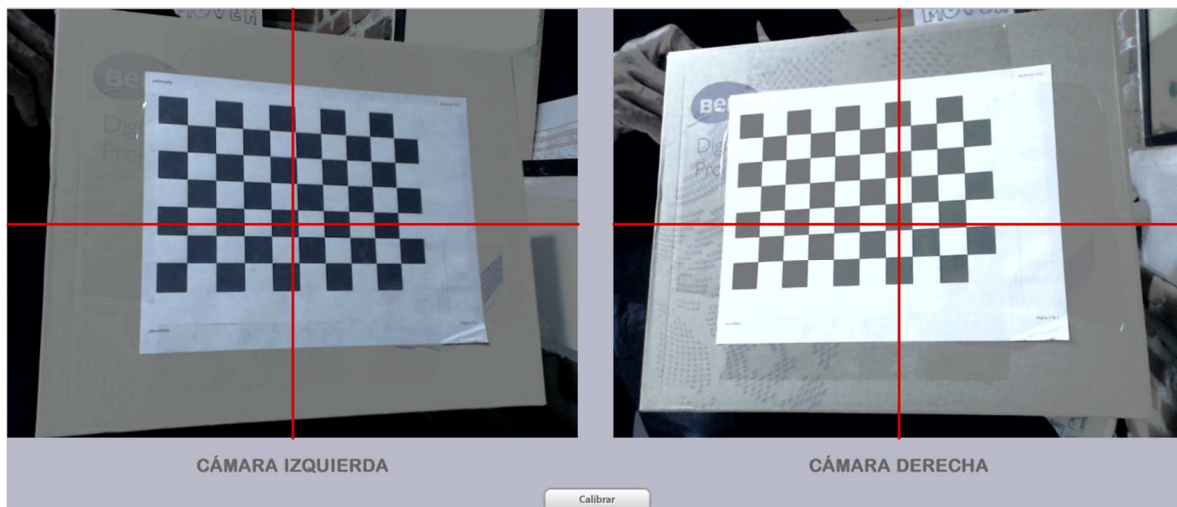


Figura 56 Algoritmo de auto ajuste de cámaras.

### 3.4 Calibración de cámaras

Para la calibración de cámaras, se tomó como base el algoritmo para generar la geometría de la visión estéreo del paso anterior. Se añadió el algoritmo para ir haciendo capturas en paralelo cada 2 segundos, con el fin de que la persona que está sosteniendo el tablero de ajedrez, pueda cambiar de la rotación y traslación del tablero.



Al finalizar las capturas, se ejecuta el código de la Figura 57, donde se pasan parámetros obtenidos en la captura de 10 imágenes del tablero en ambas cámaras, a la función `stereoCalibrate()`, de la librería OpenCV.



```

1. stereocalib_criteria = (cv2.TERM_CRITERIA_MAX_ITER +
2.                         cv2.TERM_CRITERIA_EPS, 100, 1e-5)
3.     ret, M1, d1, M2, d2, R, T, E, F = cv2.stereoCalibrate(
4.         self.objpoints, self.imgpoints_l,
5.         self.imgpoints_r, self.M1, self.d1, self.M2,
6.         self.d2, dims,
7.         criteria=stereocalib_criteria, flags=flags)
8.
9.     print('Intrinsic_mtx_1', M1)
10.    print('dist_1', d1)
11.    print('Intrinsic_mtx_2', M2)
12.    print('dist_2', d2)
13.    print('R', R)
14.    print('T', T)
15.    print('E', E)
16.    print('F', F)

```

Figura 57 Función de calibración estéreo.

Obteniendo como resultado los parámetros intrínsecos de cada cámara M1 y M2, los coeficientes de distorsión d1, d2, las matrices de rotación R, traslación T, esencial E y Fundamental F, para el arreglo estereoscópico.

### 3.5 Proyección y adquisición de patrones

El código generado para la adquisición de imágenes es un ciclo donde se captura de manera paralela las imágenes proyectadas de cada patrón.

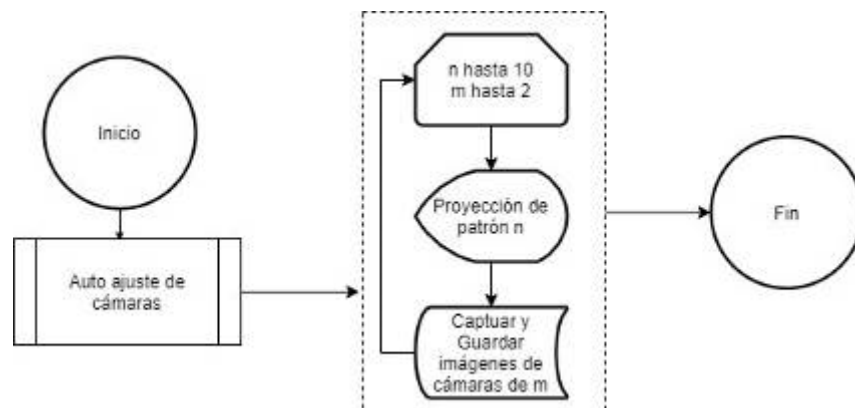


Figura 58 Algoritmo de adquisición de patrones.

Los resultados obtenidos, son 20 imágenes para el patrón de líneas verticales y 20 imágenes para el patrón de líneas horizontales.

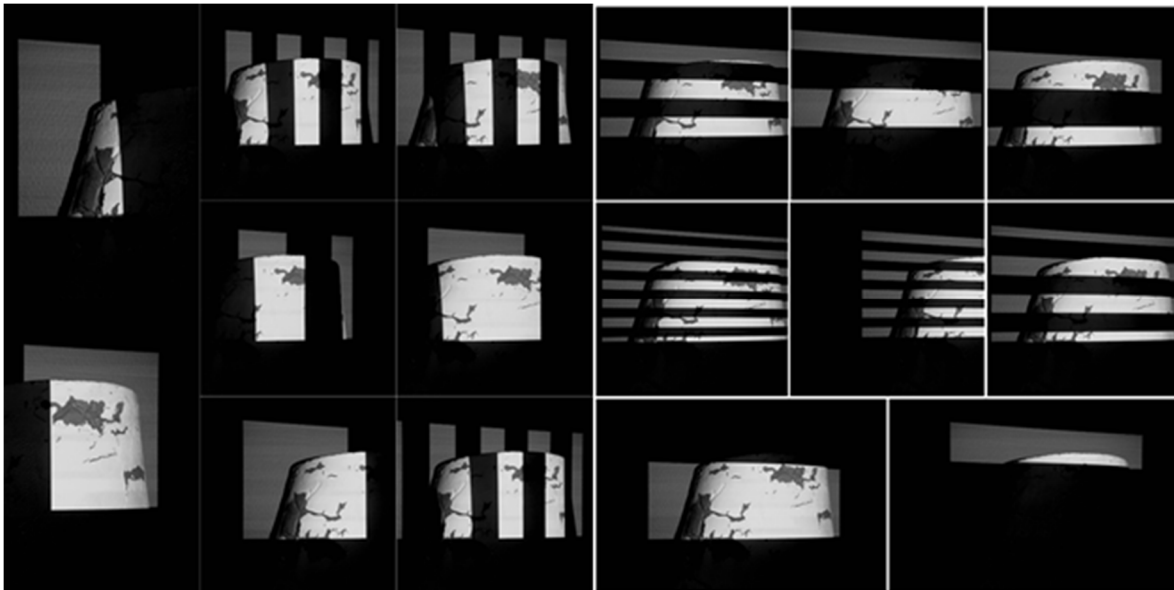


Figura 59 Proyección y captura de patrones sobre objeto a reconstruir.

Las imágenes se guardan en un directorio del sistema de archivos de Linux. Para el siguiente paso.

### 3.6 Codificación de patrones

La codificación de patrones, consiste en generar una sola imagen a partir de las capturas de las líneas verticales y otra para las capturas de las líneas horizontales. De tal modo, que en cada imagen se puedan identificar las coordenadas  $(x,y)$  de los píxeles proyectados.

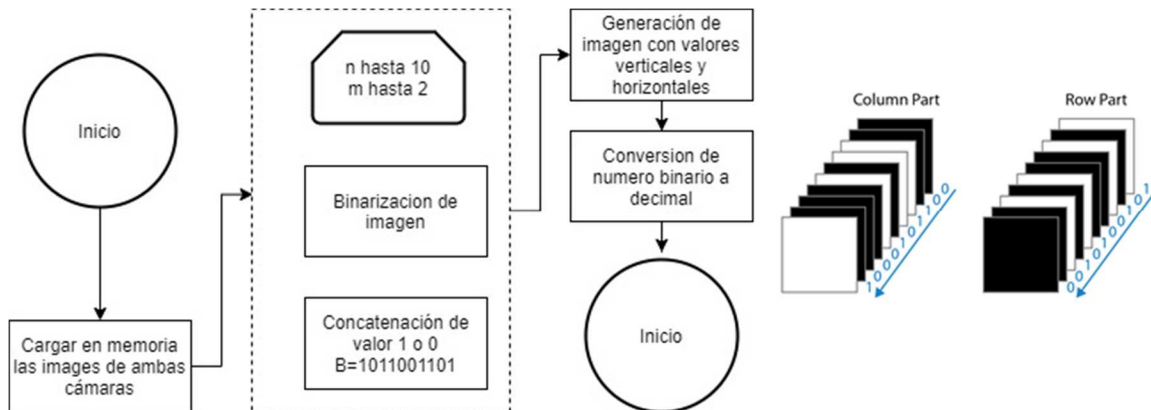


Figura 60 Codificación de patrones capturados.

### 3.7 Segmentación

Con los patrones codificados, es necesario segmentar la región de interés para disminuir el tiempo de procesamiento. A través de la interfaz de usuario, se selecciona el área de ambas cámaras.



Seleccionar

Aplicar

Seleccionar

Figura 61 Segmentación para selección de región de interés.

De este modo, obtenemos coordenadas para ejecutar la decodificación de los patrones, dejando solo las áreas de interés para mejorar el rendimiento del algoritmo.

### 3.8 Decodificación de patrones

El proceso de decodificación de patrones, consiste en generar un arreglo  $(x,y, num)$  de cada imagen de la cámara derecha e izquierda. Donde  $x,y$  es igual a las coordenadas donde se encuentra el pixel número  $num$ , que permite la identificación del pixel del proyector en ambas cámaras.

### 3.9 Método de triangulación

Teniendo los pixeles identificados en cada cámara, el último paso en el proceso de reconstrucción 3d, es encontrar el valor de  $Z$  para el punto  ${}^w p = [X, Y, Z]$ . Se desarrolló el código con base al siguiente algoritmo.

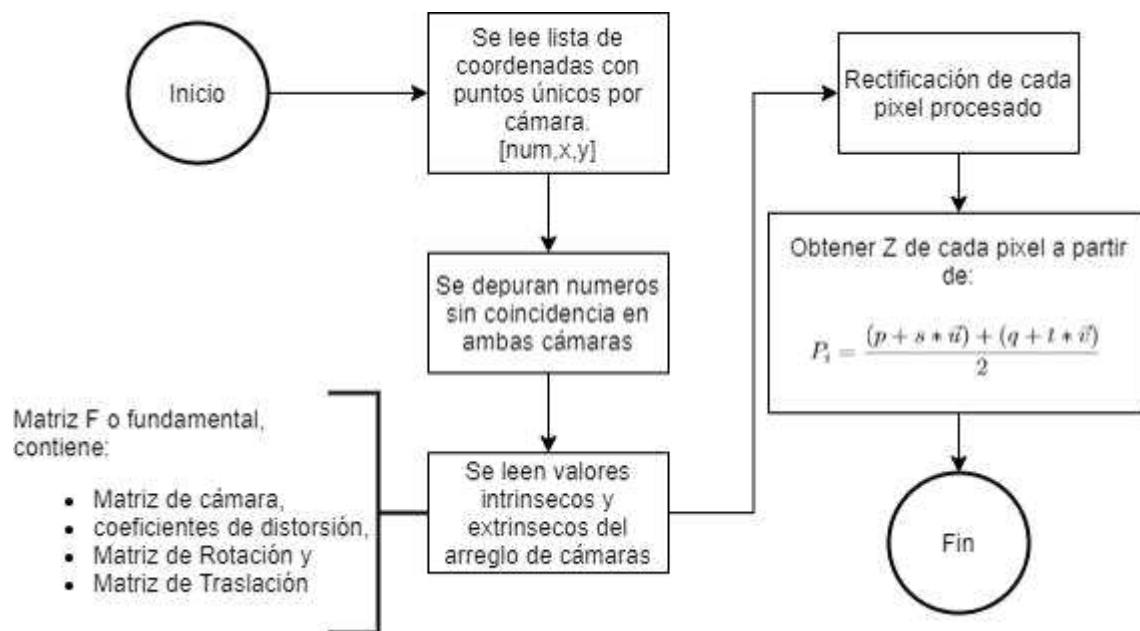


Figura 62 Método de triangulación para obtener Z.

## CAPÍTULO 4 RESULTADOS Y DISCUSIÓN

### 4.1 Pruebas de reconstrucción con visión estéreo y luz estructurada con el patrón de código binario.

Se realizaron diversas pruebas de reconstrucción 3d en diferentes materiales, con el objetivo de analizar los efectos de la reflexión de la luz directa e indirecta en la generación de puntos 3d en el objeto reconstruido.

En la Figura 63 se observa la reconstrucción de un tramo de un prototipo de álabe, elaborado con fibra de vidrio y un recubrimiento de pintura base aceite. Se realizó un pre ajuste de los parámetros de las cámaras, obteniendo los siguientes resultados:

Número de puntos	Tiempo de procesamiento	Resolución del proyector	Porcentaje de puntos reconstruidos
57,872	00:03:32	1280x720	06%

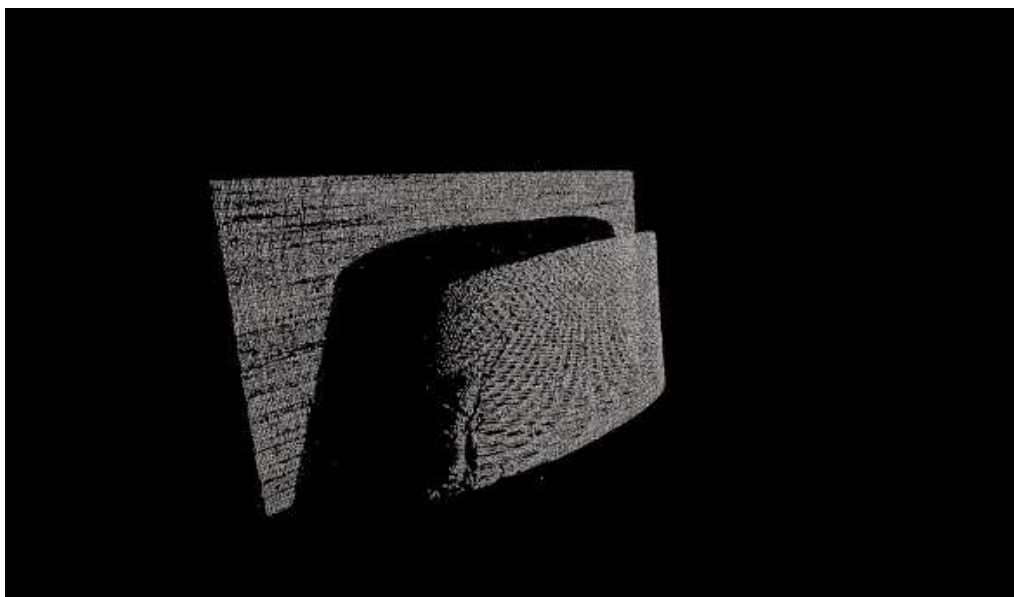


Figura 63 Nube de puntos de un prototipo álabe hecho con fibra de vidrio.

En la Figura 64 se observa la reconstrucción de un tramo de un prototipo de álabe, elaborado a base de cartón sin recubrimiento. Se realizó un pre ajuste de los parámetros de las cámaras, obteniendo los siguientes resultados:

Número de puntos	Tiempo de procesamiento	Resolución del proyector	Porcentaje de puntos reconstruidos
104,242	00:02:32	1280x720	11%

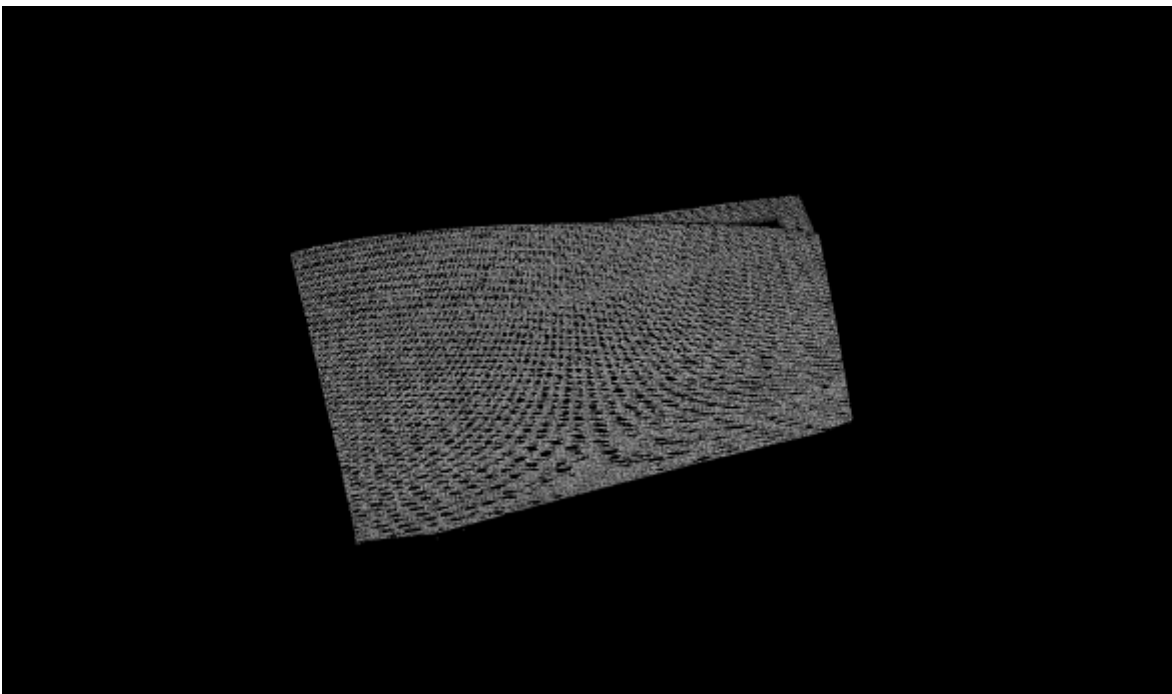


Figura 64 Nube de puntos de un prototipo álabe hecho con casrtón.

El número de punto es mayor al del primer álabe. Se debe a que los parámetros utilizados para el material de fibra de vidrio con recubrimiento con pintura base aceite, hicieron que se captara una imagen con menor exposición para compensar los efectos de la reflexión de la luz directa e indirecta. Por lo que los algoritmos de codificación de imágenes, detectaron menor cantidad de valores 1 y 0.

En el álabe de material de cartón, por las características del material, no hubo una subexposición, no se generaron efectos de reflexión, por lo que aumento el porcentaje de reconstrucción de puntos 3d.

En la Figura 65 se observa la reconstrucción de un rostro humano. Se realizó un pre ajuste de los parámetros de las cámaras, obteniendo los siguientes resultados:

Número de puntos	Tiempo de procesamiento	Resolución del proyector	Porcentaje de puntos reconstruidos
18,762	00:01:56	1280x720	2.0%



Figura 65 Reconstrucción 3d de rostro humano1.

En la reconstrucción del rostro humano1, hay un menor porcentaje de puntos reconstruidos, sin embargo, el área reconstruida es menor respecto a las dos reconstrucciones anteriores. Para comparar, se realizó una segunda reconstrucción con un modelo diferente.

En la Figura 66 se observa la reconstrucción de un rostro humano. Se realizó un pre ajuste de los parámetros de las cámaras, obteniendo los siguientes resultados:

Número de puntos	Tiempo de procesamiento	Resolución del proyector	Porcentaje de puntos reconstruidos
19,397	00:01:43	1280x720	2.1%

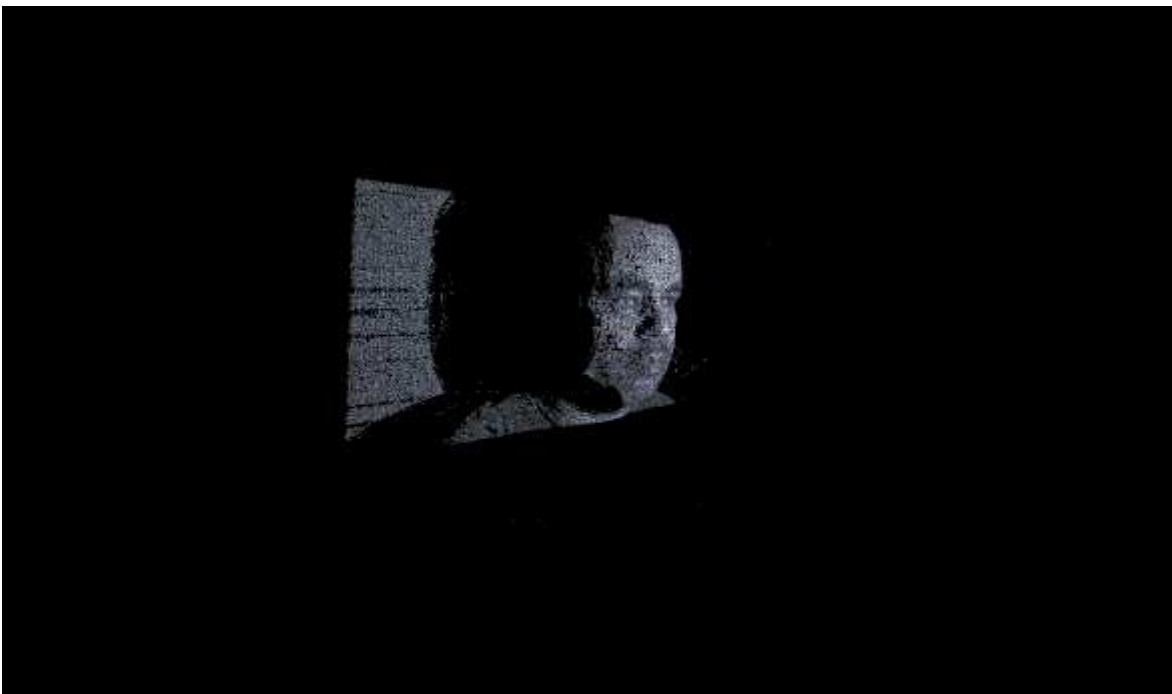


Figura 66 Reconstrucción 3d de rostro humano2.

Respecto a la reconstrucción del rostro de humano1, hay un ligero incremento de puntos 3d reconstruidos. Las características de la piel se diferenciaron por una leve cantidad de grasa cutánea mayor en el rostro humano1. Por lo que también hubo un ligero cambio en los parámetros de las cámaras, creando una subexposición en rostro humano1.



En la Figura 67 se muestra la reconstrucción de un rostro de un maniquí de silicona. Las características de la silicona son, en éste caso, levemente parecidas a la de la piel humana, sin embargo, a diferencia de la piel humana, no hay presencia de grasa. Se realizó un pre ajuste de los parámetros de las cámaras, obteniendo los siguientes resultados:

Número de puntos	Tiempo de procesamiento	Resolución del proyector	Porcentaje de puntos reconstruidos
20,467	00:01:25	1280x720	2.2%



Figura 67 Reconstrucción 3d de rostro maniquí Debra..

## CAPÍTULO 5 CONCLUSIONES, RECOMENDACIONES Y TRABAJO FUTURO

### 6.1 Conclusiones

El objetivo general de la presente investigación, era el desarrollo de algoritmos para la reconstrucción tridimensional mediante imágenes de objetos en ambientes no controlados, clave para el análisis y la medición de las superficies de objetos que pueden poner en riesgo a un usuario o para el caso de que se hagan en un horario específico o bajo las condiciones ideales. Adicionalmente se hizo una comparación de reconstrucciones en algunos distintos tipos de material.

Así pues, la aportación principal de este trabajo de investigación, consiste en el diseño, codificación e implementación de un dispositivo de reconstrucción tridimensional de objetos, llamado escáner 3d, en ambientes no controlados, remoto y en tiempo real, que permita al usuario realizar reconstrucciones de objetos en ambientes donde no hay control de la iluminación, o el lugar donde se encuentra el objeto resulte hostil para la permanencia de una persona. Para llevar a cabo tal objetivo, se han desarrollado algoritmos que trabajan de forma semiautomática, que hacen uso de funciones de procesamiento de imágenes y visión por computadora, implementados en una arquitectura IoT, que permite monitorear y controlar los componentes del hardware del escáner 3d. Como resultado de la investigación de las diferentes técnicas de reconstrucción, se concluye que las técnicas con mayor rapidez y precisión para tales reconstrucciones, son las basadas en luz estructurada.

Debido a que se debía considerar un ambiente hostil, se llevó a cabo el análisis de las técnicas que permitieran llevar el escáner 3d listo para operar, sin configuraciones previas. Esto, dirigió la investigación hacia las técnicas basadas en luz estructurada y con visión estereoscópica. Ya que la geometría del arreglo óptico,

puede ser configurada y queda fija desde el diseño del escáner 3d. Las condiciones a las que estaría el escáner 3d, que no son de un laboratorio de pruebas, hizo que finalmente se decidiera hacer uso de la técnica de luz estructurada con proyección de patrones en código binario. Pues, resulta menos complicado en costo y tiempo computacional, procesar imágenes que pueden ser binarizadas y así obtener, de manera correcta, la información necesaria para las reconstrucciones.

Para el control y monitoreo del escáner 3d, se realizó un estudio de la forma en que se controla actuadores y monitorean sensores en la industria, por ser el ambiente donde se necesitan componentes con el menor porcentaje de fallas, encontrando, que la tendencia va hacia el uso de la arquitectura del internet de las cosas (IoT por sus siglas en inglés). Y que es una arquitectura que ha venido evolucionando y actualmente es conocida como IIoT (Industry Internet of Things) de cuarta generación, y debido a esta tendencia, se tomó la decisión de crear una lógica en el flujo de la información e interacción con el usuario, basada en ésta arquitectura.

Finalmente, en el desarrollo y pruebas de los algoritmos, se detectaron funciones y piezas que pueden mejorar el desempeño del escáner 3d. Por ejemplo, crear un proyector que concentre la luz de los patrones con la luminosidad suficiente para aumentar la relación señal ruido, o hacer uso de técnicas de aprendizaje de máquina para las funciones semi automáticas del escáner 3d, como el ajuste automático de parámetros de cámaras.

No obstante, con los resultados obtenidos, nos permiten tener un mejor conocimiento de la reconstrucción 3d y poder aplicar alguna de las técnicas para resolver otro tipo de problemas.

## 6.2 Recomendaciones

Durante el desarrollo del escáner 3d, se quiso inferir sin un conocimiento sólido, sobre cómo se creaban las reconstrucciones 3d. Se tomó como referencia el

sistema para juegos Kinect de Microsoft, que es conocido como un sensor 3d, el cual, ya genera una nube de puntos mediante la técnica de luz estructurada conocida como time-of-flight. Llevó tiempo darse cuenta que, primero se debía de entender cómo funciona la visión humana, pues es el punto de partida para la reconstrucción 3d.

Para poder seguir con trabajos futuros, se recomienda primero, comprender los temas de visión estéreo, pues son las bases de todas las técnicas de reconstrucción 3d.

También se recomienda hacer uso de librerías como OpenCV, debido a que muchas de las funciones de procesamiento de imagen y visión estéreo, se encuentran desarrolladas y documentadas.

### 6.3 Trabajo futuro

Tomando como base la presente investigación y los algoritmos desarrollados, se pueden mejorar varios puntos:

- Explorar el uso de otras técnicas, encontrar las ventajas de cada una, para realizar pruebas de combinación de técnicas que lleven a mejores resultados.
- Hacer uso de procesamiento en FPGA's, pues si bien los algoritmos generados para la presente investigación resultaron en algoritmos con procesamiento ligero, existen otras técnicas que requieren de mayor poder computacional, que puede ser menor con el uso de estos dispositivos.
- Trabajar con la comunicación entre sistemas desarrollados con procesadores para dispositivos portables como ARM y FPGA's. Ya que la programación de los FPGA's tiene una curva de aprendizaje que puede ser muy larga en comparación con los lenguajes de programación para procesadores, Se puede dividir el desarrollo de los sistemas en dos partes. La parte de interfaz de usuario, puede desarrollarse más fácilmente con tecnologías para sistemas basados en procesador. Las funciones con mayor demanda de procesamiento, pueden programarse en los dispositivos FPGA's. De tal modo que, la mayor inversión del tiempo de desarrollo quede en la creación de funciones consisas, concretas y compactas del lado de los FPGA's y la parte de la interaccion con el el usuario del lado de los sistemas basados en procesador.

## BIBLIOGRAFÍA

- 3DFASCINATION. (2020). *Fascinating World of 3D*. 2020. <https://www.3dfascination.com/>
- Artec. (2020). *Aplicaciones de escaneo 3D de Artec en el cuidado de la salud*.  
<https://www.artec3d.com/es/cases/medical>
- Ashton, K. (2009). That Internet of Things Thing. *RFID Journal*.
- AU - Ekstrand, L., AU - Karpinsky, N., AU - Wang, Y., & AU - Zhang, S. (2013). High-resolution, High-speed, Three-dimensional Video Imaging with Digital Fringe Projection Techniques. *JoVE*, 82, e50421.  
<https://doi.org/doi:10.3791/50421>
- Brooks, M. J., & Horn, B. K. P. (1989). Shape and Source from Shading. En *Shape from Shading*.
- Brown, D. (1966). Decentering Distortion of Lenses. *Photometric Engineering*.
- Courtland, R. (2015). *Gordon Moore: The Man Whose Name Means Progress*. IEEE SPECTRUM.  
<https://spectrum.ieee.org/computing/hardware/gordon-moore-the-man-whose-name-means-progress>
- Fanuc. (2020). *Manipulación de materiales automatizada mediante un sistema de visión*.  
<https://www.fanuc.eu/es/es/robots/accesorios/visión/3d-vision>
- Gonzalez, A. L., Meneses, J. E., & León, J. (2012). Proyección de franjas en metrología óptica facial. *Inge-Cuc*.
- Google. (2019). *Cloud Tensor Processing Unit (TPU)*. Google cloud.  
<https://cloud.google.com/tpu/docs/tpus?hl=es-419>
- Gunatilake, A., Piyathilaka, L., Kodagoda, S., Barclay, S., & Vitanage, D. (2019). Real-time 3D profiling with RGB-D mapping in pipelines using stereo camera vision and structured IR laser ring. *Proceedings of the 14th IEEE Conference on Industrial Electronics and Applications, ICIEA 2019*.  
<https://doi.org/10.1109/ICIEA.2019.8834089>
- Gupta, M., Agrawal, A., Veeraraghavan, A., & Narasimhan, S. G. (2011). Structured light 3D scanning in the presence of global illumination. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. <https://doi.org/10.1109/CVPR.2011.5995321>
- Gupta, M., Yin, Q., & Nayar, S. K. (2013). Structured light in sunlight. *Proceedings of the IEEE International Conference on Computer Vision*. <https://doi.org/10.1109/ICCV.2013.73>
- Haleem, A., & Javaid, M. (2019). 3D scanning applications in medical field: A literature-based review. *Clinical Epidemiology and Global Health*. <https://doi.org/10.1016/j.cegh.2018.05.006>
- Halfacree, G. (2019). *Benchmarking the Raspberry Pi 4*. <https://medium.com/@ghalfacree/benchmarking-the-raspberry-pi-4-73e5afbcd54b>
- Horn, B. K. P. (1970). Shape from Shading: A Method for Obtaining the Shape of a Smooth Opaque Object from One View. *Doctor*.
- Intel, & Microsoft. (2020). *Unleash Your Data Center*. <https://www.intel.la/content/www/xl/es/data-center/products/programmable/overview.html>
- Kanagachidambaresan, G. R., Anand, R., Balasubramanian, E., & Mahima, V. (2020). *Internet of Things for Industry 4.0: Design, Challenges and Solutions*. Springer.
- Keyence. (2020). *Mediciones de forma, contorno y rugosidad en tan sólo 1 segundo*.  
[https://www.keyence.com.mx/landing/microscope/lp\\_vr\\_obtain.jsp](https://www.keyence.com.mx/landing/microscope/lp_vr_obtain.jsp)
- Masood, S. Z., Shu, G., Dehghan, A., & Ortiz, E. G. (2017). *License Plate Detection and Recognition Using Deeply Learned Convolutional Neural Networks*. <http://arxiv.org/abs/1703.07330>
- Matlab. (2019). *What Is Camera Calibration?* R2019b. <https://de.mathworks.com/help/vision/ug/camera-calibration.html>
- Medit. (2020). *Medit i500*. 2020. <https://info.medit.com/download-i500-brochure>
- Okutomi, M., & Kanade, T. (1993). A Multiple-Baseline Stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. <https://doi.org/10.1109/34.206955>
- Oracle. (2020). *Internet of Things*. <https://www.oracle.com/mx/internet-of-things/what-is-iot.html>
- Pollefeys, M., Koch, R., & Van Gool, L. (1999). Self-calibration and metric reconstruction inspite of varying and unknown intrinsic camera parameters. *International Journal of Computer Vision*.  
<https://doi.org/10.1023/A:1008109111715>
- Ricofe-Viala, C., & Sánchez-Salmerón, A. J. (2010). Robust metric calibration of non-linear camera lens distortion. *Pattern Recognition*. <https://doi.org/10.1016/j.patcog.2009.10.003>

- Roberts, L. Gi. (1965). Lawrence Roberts Machine Perception of Three Dimensional Solids. En *PhD Thesis*.
- ROS. (2020). *ROS Framework*. <https://www.ros.org/>
- Styl&Tech. (2020). *Styl&Tech Reverse Engineering*. <https://www.stylntech.com/en/3d-scanning-reverse-engineering/>
- Wang, Y. (2010). *Novel Approaches in Structured Light Illumination*.
- Zelinsky, A. (2009). Learning OpenCV—Computer Vision with the OpenCV Library. En *IEEE Robotics and Automation Magazine*. <https://doi.org/10.1109/MRA.2009.933612>
- Zhang, S. (2013). Handbook of 3D Machine Vision: Optical Metrology and Imaging. En *Handbook of 3D Machine Vision: Optical Metrology and Imaging*. <https://doi.org/10.1080/00107514.2013.856944>
- Zhang, Z. (2000). A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. <https://doi.org/10.1109/34.888718>

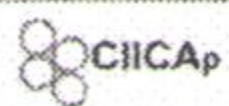
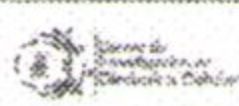


UNIVERSIDAD AUTÓNOMA DEL  
ESTADO DE MORELOS



**INSTITUTO DE INVESTIGACIÓN EN CIENCIAS BÁSICAS Y APLICADAS**

Defatura de Posgrado en Ingeniería y Ciencias Aplicadas



Cuernavaca, Morelos, a 28 de febrero de 2020.

**DR. ROSENBERG JAVIER ROMERO DOMÍNGUEZ**  
**COORDINADOR DEL POSGRADO EN**  
**INGENIERÍA Y CIENCIAS APLICADAS**  
**P R E S E N T E**

Atendiendo a la solicitud para emitir DICTAMEN sobre la revisión de la TESIS titulada “**DISPOSITIVO DE RECONSTRUCCIÓN TRIDIMENSIONAL REMOTO MEDIANTE LUZ ESTRUCTURADA Y VISIÓN ESTEREOSCÓPICA PARA AMBIENTES NO CONTROLADOS**” que presenta el alumno **ROBERTO ALAN BELTRAN VARGAS**, para obtener el título de **DOCTORADO EN INGENIERÍA Y CIENCIAS APLICADAS** con opción terminal en **TECNOLOGÍA ELÉCTRICA**.

Nos permitimos informarle que nuestro voto es:

NOMBRE	DICTAMEN	FIRMA
DR. ÁLVARO ZAMUDIO LARA	Aprobado	
DR. JOSÉ ANTONIO MARBÁN SALGADO	Aprobado	
DRA. MARGARITA TECPOYOTL TORRES	Aprobado	
DR. OMAR PALILLERO SANDOVAL	Aprobado	
DR. J. GUADALUPE VELÁSQUEZ AGUILAR (FCQel)	Aprobado	
DR. JOSÉ ALBERTO HERNÁNDEZ AGUILAR (FCAel)	Aprobado	
DR. J JESÚS ESCOBEDO ALATORRE	Aprobado	

**PLAZO PARA LA REVISIÓN 20 DÍAS HÁBILES (A PARTIR DE LA FECHA DE RECEPCIÓN DEL DOCUMENTO)**

**NOTA. POR CUESTION DE REGLAMENTACIÓN LE SOLICITAMOS NO EXCEDER EL PLAZO SEÑALADO, DE LO CONTRARIO LE AGRADECEMOS SU ATENCIÓN Y NUESTRA INVITACIÓN SERÁ CANCELADA.**