

DESARROLLO DE UN DISPOSITIVO DE MEDICIÓN EN TRES
DIMENSIONES UTILIZANDO TÉCNICAS DE METROLOGÍA
ÓPTICA

NICOLÁS EMANUEL ULRICH



Ingeniería en Informática
Facultad de Ingeniería y Ciencias Hídricas
Universidad Nacional del Litoral

Director: Nicolás Bonadeo
Co-director: Nestor Calvo

ÍNDICE GENERAL

1	INTRODUCCIÓN	1
1.1	Motivación	1
1.2	Objetivos	4
1.3	Organización de la tesis	5
2	METROLOGÍA ÓPTICA	7
2.1	Introducción	7
2.2	Principio de triangulación	11
2.2.1	Intersección línea-plano	12
2.2.2	Intersección línea-línea	13
2.3	Visión estéreo	15
2.4	Escáner de línea	16
2.5	Luz estructurada	17
2.5.1	Patrones de luz estructurada	18
3	CALIBRACIÓN DE LA CÁMARA	23
3.1	Introducción	23
3.2	Sistema óptico de la cámara	24
3.2.1	Óptica de primer orden	25
3.2.2	El círculo de confusión	27
3.2.3	Óptica de tercer orden	28
3.3	Adquisición de imágenes	30
3.4	Modelo de cámara pinhole	31
3.5	Modelo pinhole con distorsión	34
3.5.1	Distorsión radial	34
3.5.2	Distorsión tangencial	35
3.6	Parámetros extrínsecos	36
3.7	Calibración de la cámara	37
3.7.1	Patrón de calibración	38
3.7.2	Homografía	39
3.7.3	Obtención de los parámetros de la matriz de la cámara	40
3.7.4	Obtención de los parámetros extrínsecos	43
3.7.5	Obtención de los parámetros de distorsión	43
3.8	Calibración multi-cámara	44
4	DESARROLLO	47
4.1	Prototipo	47
4.1.1	Componentes	48
4.2	Software	50
4.3	Proyección y decodificación de los patrones	51
4.4	Reconstrucción	55
4.5	Calibración del dispositivo	60
4.6	Limitaciones	62

5	RESULTADOS	63
5.1	Preparación	63
5.2	Objetos de prueba	67
5.3	Resultados	67
5.3.1	Objeto plano	67
5.3.2	Objeto cilíndrico	71
6	CONCLUSIONES Y TRABAJOS FUTUROS	77
6.1	Conclusiones	77
6.2	Trabajos futuros	78
A	FUENTES EN RUIDO EN CÁMARAS DIGITALES	81
A.1	Ruido fotónico	81
A.2	Ruido de lectura	81
A.3	Ruido térmico	82
A.4	Respuesta no uniforme del pixel	82
A.5	Error de cuantización	82
B	EJEMPLO COMPLETO DE UNA MEDICIÓN	83
C	SUPERFICIES BRILLOSAS	89
	BIBLIOGRAFÍA	91

ÍNDICE DE FIGURAS

Figura 1	Ejemplos de usos de metroología óptica	2
Figura 2	Ejemplos de usos de metroología óptica	3
Figura 3	Ejemplos de usos de metroología óptica	4
Figura 4	Ejemplos de usos de metroología óptica	5
Figura 5	Esquema genérico de un sensor 3D óptico	8
Figura 6	Triangulación láser	9
Figura 7	Interferómetro de luz blanca	10
Figura 8	Principio de funcionamiento de un sensor de deflectometría	11
Figura 9	Intersección línea-plano	12
Figura 10	Intersección línea-línea	13
Figura 11	Visión estéreo	16
Figura 12	Escáner de línea	16
Figura 13	Sistema de luz estructurada típico	17
Figura 14	Secuencia de pasos un sistema de luz estructurada	18
Figura 15	Patrón infrarrojo usado por el sensor Microsoft Kinect	19
Figura 16	Diferencias entre patrones usando código Gray y código binario estándar	20
Figura 17	Codificación con tres patrones sinusoidales desfasados	21
Figura 18	Rayo de luz atravesando un medio	25
Figura 19	Lente esférica	25
Figura 20	Rayo de luz atravesando un medio	27
Figura 21	Aberraciones	28
Figura 22	Astigmatismo	29
Figura 23	Curvatura de campo	30
Figura 24	Distorsión	30
Figura 25	Esquema del sensor de una cámara	31
Figura 26	Modelo de cámara pinhole	32
Figura 27	Modelo de proyección de la cámara	33
Figura 28	Possible causante de distorsión tangencial	36
Figura 29	Parámetros extrínsecos	37
Figura 30	Distorsión	39
Figura 31	Rotación y traslación entre dos cámaras	45
Figura 32	Esquema del prototipo	47
Figura 33	Dispositivos comerciales	49
Figura 34	Diagrama de funcionamiento de la tecnología DLP	50
Figura 35	Software de medición	51

- Figura 36 Patrones de luz estructurada usando código gray 52
- Figura 37 Ejemplo de la proyección de un patrón y el patrón inverso correspondiente 53
- Figura 38 Ejemplo de un patrón proyectado 53
- Figura 39 Resultado de la decodificación de las imágenes. El color representa el código: en este caso el código es el valor de la columna del proyector 54
- Figura 40 Bordes entre franjas decodificadas 55
- Figura 41 Geometría epipolar 57
- Figura 42 Visualización de la nube de puntos 3D 59
- Figura 43 Software de calibración de la cámara 61
- Figura 44 Imágenes de calibración para la cámara izquierda 64
- Figura 45 Imágenes de calibración para la cámara derecha 65
- Figura 46 Imágenes para la calibración extrínseca entre cámaras. Izquierda: cámara 1, derecha: cámara 2 66
- Figura 47 Medición del patrón de planicidad 67
- Figura 48 Objeto plano: desvío estándar del error en función de la distancia 68
- Figura 49 Diversas vistas de las mediciones del objeto plano 69
- Figura 50 Diversas vistas de las mediciones del objeto plano (continuación) 70
- Figura 51 Medición de una cupla 71
- Figura 52 Objeto cilíndrico: error en la estimación del radio en función de la distancia 72
- Figura 53 Objeto cilíndrico: distribución del error 73
- Figura 54 Objeto cilíndrico: desvío estándar del error en función de la distancia 73
- Figura 55 Objeto cilíndrico: error en la orientación del cilindro en función de la distancia 75
- Figura 56 Objeto cilíndrico: error en la ubicación espacial del cilindro en función de la distancia 75
- Figura 57 Diversas vistas de las mediciones del objeto cilíndrico 76
- Figura 58 Ejemplo de una medición: nube de puntos 83
- Figura 59 Ejemplo de una medición: superficie reconstruida 83
- Figura 60 Cámara izquierda: 5 primeros patrones proyectados (izq.) y su inverso (der.) 84
- Figura 61 Cámara izquierda: 5 últimos patrones proyectados (izq.) y su inverso (der.) 85

Figura 62	Cámara derecha: 5 primeros patrones proyectados (izq.) y su inverso (der.)	86
Figura 63	Cámara derecha: 5 últimos patrones proyectados (izq.) y su inverso (der.)	87
Figura 64	Ejemplo de un problema causado por superficie brillosa	89
Figura 65	Causas de la decodificación incorrecta	90
Figura 66	Spray para la medición de superficies brillosas	90

ÍNDICE DE TABLAS

Tabla 1	Comparación de los distintos tipos de sensores ópticos	11
Tabla 2	Resultados de la medición del objeto plano	68
Tabla 3	Resultados de la medición del objeto cilíndrico	72
Tabla 4	Orientación y ubicación espacial (punto sobre el eje) del cilindro	74

ACRÓNIMOS

API	Application Programming Interface
CCD	Charge-Coupled Devices
CMOS	Complementary Metal-Oxide Semiconductors
CSI	Coherence Scanning Interferometry
MEMS	Micro Electro-Mechanical Systems
PSF	Point Spread Function
SDK	Software Development Kit

1

INTRODUCCION

En la primer sección de este capítulo se explican los motivos que llevan al desarrollo de este trabajo, mientras que en la segunda sección se detallan los principales objetivos buscados. Por último se incluye también una breve descripción de la manera en que está organizada la tesis.

1.1 MOTIVACIÓN

La obtención de modelos tridimensionales precisos y de alta calidad es crucial para un amplio rango de aplicaciones, entre las que se puede nombrar el control de robots inteligentes, la detección de obstáculos para vehículos automáticos, el control de calidad, la ingeniería inversa, entre muchas otras [1]. En la industria existe la necesidad de medir con precisión la geometría de diversos objetos para acelerar el desarrollo de productos y asegurar la calidad del proceso de manufactura. Mediante el relevamiento de la geometría de objetos se pueden automatizar tareas como la inspección y el reconocimiento de defectos en la línea de producción.

Debido a su funcionamiento sin contacto, las técnicas de inspección óptica se presentan como una alternativa muy interesante para este tipo de aplicaciones. Si bien diversos métodos ópticos de medición son conocidos desde hace décadas, la evolución y abaratamiento experimentados en los últimos años en las computadoras, dispositivos de obtención de imágenes digitales, componentes electro-ópticos, láseres y otras fuentes de luz permitieron su aplicación con éxito en diversos ambientes industriales y de consumo masivo.

Existe una gran variedad de técnicas ópticas tridimensionales: triangulación láser, luz estructurada, visión estéreo, fotogrametría, tiempo de vuelo, interferometría, entre otras [2]. Entre ellos se destacan los métodos de luz estructurada y triangulación láser debido a su bajo costo, robustez, flexibilidad, velocidad y principalmente su precisión y resolución [3]. Estos métodos se utilizan con éxito para medir elementos a diferentes escalas, partiendo desde pequeños componentes Micro Electro-Mechanical Systems (**MEMS**) hasta grandes paneles de hasta 2 metros de largo, chequeo de la geometría de paneles de estampado, medición de deformaciones, análisis de corrosión, etc, existien-

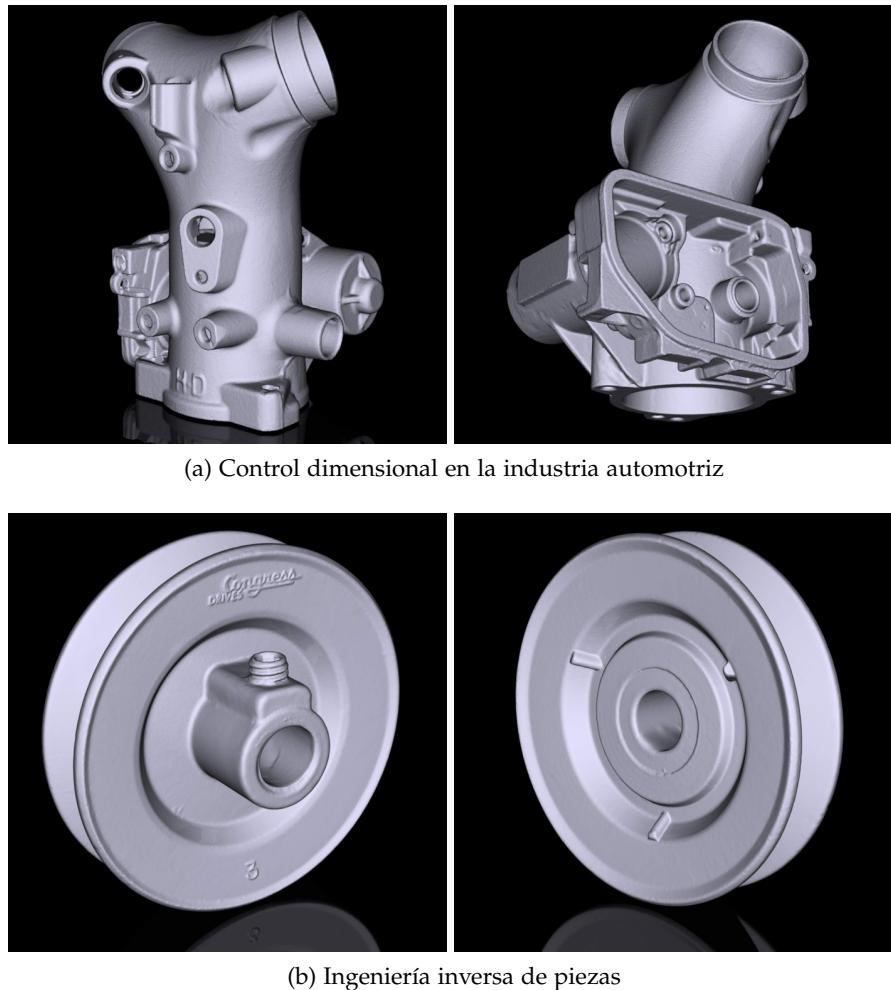


Figura 1: Ejemplos de usos de metroología óptica

do diversas empresas que ofrecen dispositivos e incluso soluciones completas de manera comercial¹²³⁴.

Los dispositivos disponibles comercialmente suelen ser relativamente fáciles de utilizar y muchas veces vienen calibrados desde fábrica, pero este mismo hecho limita su flexibilidad respecto al rango de medición y la distancia mínima de trabajo (comúnmente conocida como *standoff*). Al tratarse de soluciones cerradas, se limitan considerablemente las posibilidades de adaptación para su uso en situaciones particulares. Como ejemplo, un producto⁵ del cual se dispone⁶ actualmente basado en tiempo de vuelo, no logra la precisión deseada (menor a 1 mm) y tampoco resulta suficiente su resolución (genera

¹ <http://www.gom.com/metrology-systems/3d-scanner.html>

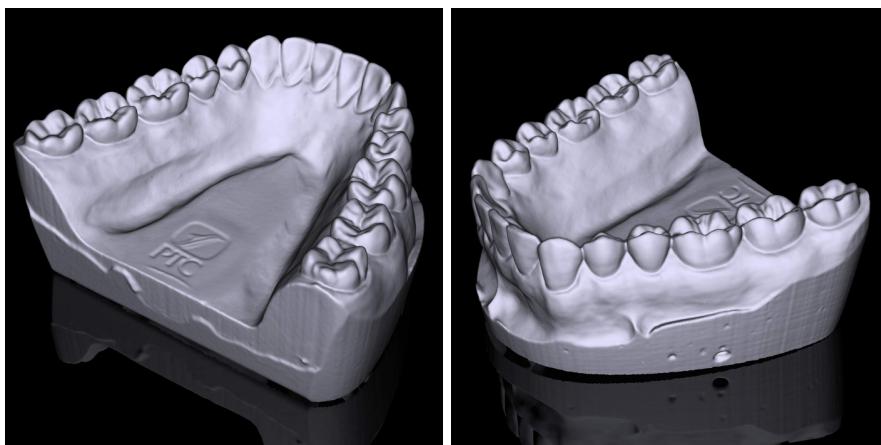
² <http://lmi3d.com/products/hdi/>

³ <http://lmi3d.com/products/gocator/snapshot-sensor/>

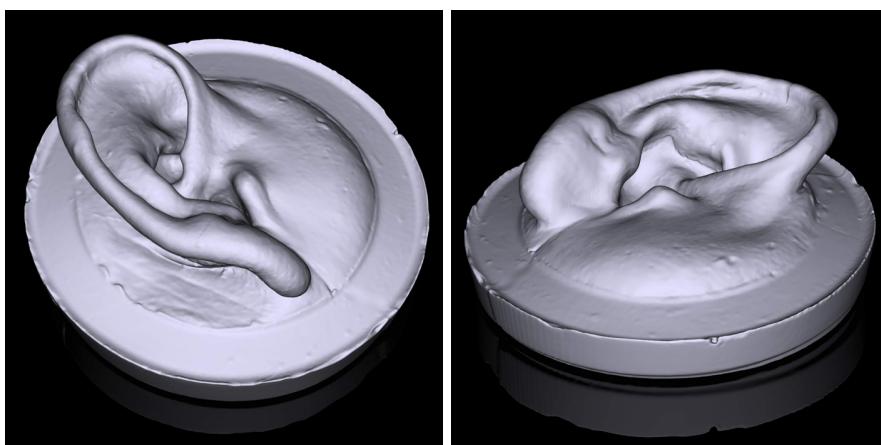
⁴ <http://geomagic.com/en/products/capture/overview/>

⁵ Cámara 3D Fotonic TOF-C70E, más detalles disponibles en http://www.fotonic.com/assets/documents/products/Fotonic_C-series.pdf

⁶ En el laboratorio de Física Aplicada del Centro de Investigación Industrial de Tenaris Siderca



(a) Odontología: digitalización de modelos dentales



(b) Medicina

Figura 2: Ejemplos de usos de metrología óptica

una imagen en profundidad de 160x120 pixels, lo que equivale a sólo 19200 puntos, con una distancia mínima entre ellos de aproximadamente 4mm).

En este proyecto se propone el desarrollo de un dispositivo de medición en tres dimensiones utilizando técnicas de metrología óptica, y evaluar su aplicación para la ubicación y el dimensionamiento de objetos cilíndricos en un ambiente industrial.

El desarrollo de un dispositivo a medida brindará la flexibilidad requerida para optimizar el balance entre precisión y rango de medición necesarios y permitirá obtener la resolución deseada. Otros beneficios importantes son el completo control sobre las diversas etapas de la reconstrucción de la geometría en tres dimensiones y la generación de una base tecnológica sólida sobre la cual continuar evolucionando la tecnología. Los desarrollos, el conocimiento y la experiencia adquiridos podrán ser utilizados luego como punto de partida para el desarrollo de nuevas soluciones a otros problemas similares.

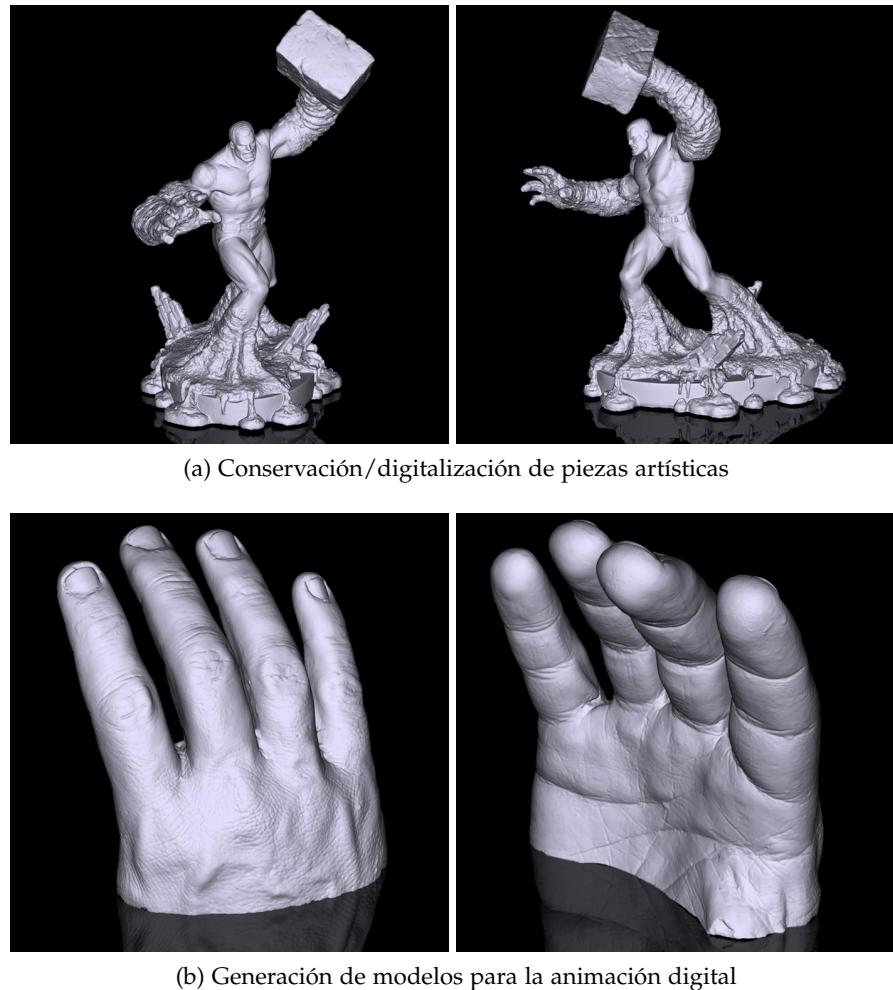


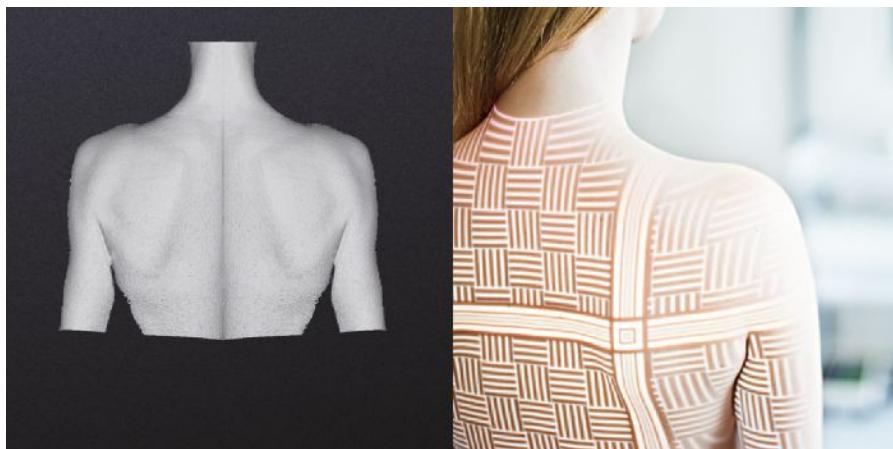
Figura 3: Ejemplos de usos de metrología óptica

1.2 OBJETIVOS

- Generales:
 - Desarrollar un dispositivo para detectar la presencia de un objeto cilíndrico en el espacio y determinar con precisión su ubicación y dimensiones.
- Específicos:
 - Obtener conocimiento sobre técnicas y algoritmos de metrología óptica, luz estructurada, triangulación.
 - Adquirir conocimientos en los fundamentos de la óptica, la fotografía digital y la geometría computacional aplicables a estos métodos de medición.
 - Desarrollar un método para facilitar la calibración del sistema.
 - Estimar la precisión que se puede lograr bajo distintas configuraciones del sistema.

1.3 ORGANIZACIÓN DE LA TESIS

En el primer capítulo se expuso la motivación y los principales objetivos del proyecto. En el segundo capítulo se realizará una introducción a los sistemas ópticos de medición con especial énfasis en los métodos basados en triangulación. En el tercer capítulo se explicará en detalle todo lo relacionado a la óptica del sistema y la calibración del principal componente, la cámara. En el cuarto capítulo se explican los detalles del prototipo desarrollado, tanto respecto al hardware como software. En el quinto capítulo se muestran algunos resultados obtenidos y finalmente la conclusiones se presentan en el sexto capítulo.



(a) Uso medicinal: corrección de la postura



(b) Preservación de evidencias para uso forense

Figura 4: Ejemplos de usos de metrología óptica

2

METROLOGÍA ÓPTICA

En este capítulo se brinda una mirada general de la metrología óptica. En la primer sección se realiza una introducción a los sistemas ópticos de medición, sus ventajas y desventajas y se realiza una clasificación general en base a su principio de funcionamiento. En la segunda sección se trata en detalle el principio de triangulación, la base de los sensores más conocidos y utilizados. En las secciones a continuación se explican los principales métodos de medición basados en triangulación: en la tercera sección se explican los sensores basados en visión estéreo, en la cuarta sección se describen los sensores conocidos como escaneres de línea, y finalmente en la quinta sección se explicará en detalle el funcionamiento de las distintas variantes de los métodos conocidos como luz estructurada, una de las cuales será utilizada para el desarrollo de este proyecto.

2.1 INTRODUCCIÓN

En la [Figura 5](#) se observa el esquema de un sensor 3D óptico genérico. Básicamente se observa una fuente de luz, el objeto y el sensor que observa la interacción entre el objeto y la luz. La iluminación puede ser implementada de diversas formas y suele tener un efecto importante en la performance del sensor. Por ejemplo, la iluminación puede ser coherente o incoherente, directa o difusa, estructurada o homogénea, monocromática o multicromática, polarizada o no, y temporalmente continua o pulsada.

La luz incide sobre el objeto observado e interactúa con los átomos que lo conforman. Generalmente se produce scattering coherente, por lo que la luz reflejada está en fase con la luz proveniente de la fuente. El scattering coherente genera una limitación fundamental, conocida como ruido speckle, que será explicado más adelante. Para evitarlo se puede utilizar una excitación térmica o se pueden aprovechar los efectos de la fluorescencia, pero ambos temas exceden el alcance de este trabajo y no serán tratados.

Por otro lado, la interacción de la luz con la superficie puede causar reflexión difusa o especular (como un espejo), y se puede producir en la superficie o dentro del volumen del objeto (como se produce por ejemplo en la piel humana, o en plásticos).

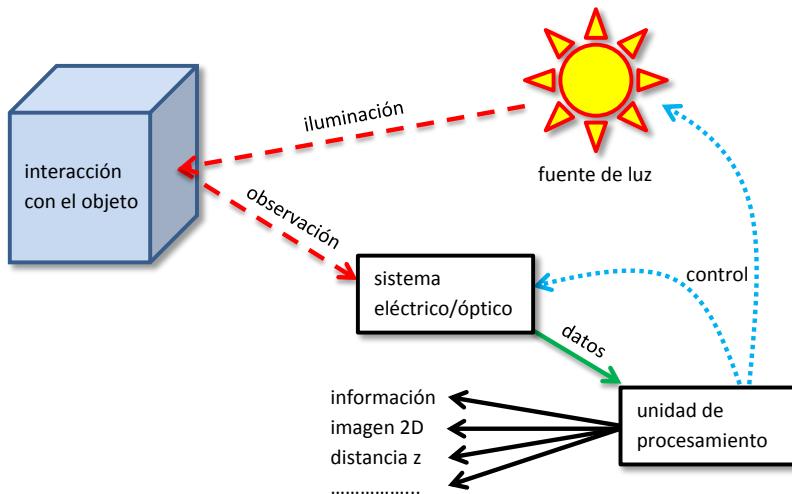


Figura 5: Esquema genérico de un sensor 3D óptico

La luz reflejada (o generada) por el objeto contiene mucha información, principalmente intensidad y color (como en fotografía), amplitud compleja (como en interferometría y holografía), polarización, coherencia y tiempo de vuelo. Estas características pueden ser combinadas y utilizadas para obtener la forma 3D de la superficie.

Los sistemas ópticos de medición ofrecen varias ventajas respecto a los tradicionales métodos de medición por contacto, sobre todo en ambientes y procesos industriales. Las que más nos interesan son su velocidad y el hecho de que no entran en contacto con la pieza, lo que permite por ejemplo realizar mediciones sin comprometer la integridad del propio dispositivo de medición.

Los sensores ópticos pueden clasificarse en cuatro grandes categorías a partir de las características de la incertezza en su medición [4]:

1. Triangulación
2. Interferometría de escaneado coherente y tiempo de vuelo
3. Interferometría clásica
4. Deflectometría

El tipo 1 incluye todos los sensores basados en triangulación: triangulación laser, luz estructurada, visión estereo, microscopía confocal, *shape from focus*, *shape from defocus*, etc. Los sensores de triangulación generalmente miden a partir de la deformación lateral provocada por la perspectiva. La incertezza de la medición queda determinada principalmente por la incertezza de la medición de este desplazamiento. A partir de un simple análisis geométrico se puede observar que la incertezza de la medición escala cuadráticamente con la distancia, como se explica en [5]. En el caso de la medición de objetos rugosos la fuente de ruido dominante es el producido por los speckles, de los cuales hablaremos más adelante. Vamos a denominar a este tipo

de sensores como sensores del tipo 1a. Para superficies suaves o para triangulación con fluorescencia, la fuente de ruido dominante es el ruido fotónico (ver [Apéndice A](#)). Denominaremos a este tipo de sensores como tipo 1b.

En la [Figura 6](#) se muestra una representación de un sensor de triangulación. En este caso se trata de un sensor del tipo 1a, con una fuente de luz laser puntual. El haz de luz se enfoca sobre un eje de proyección hacia la superficie de un objeto rugoso. El punto es observado mediante un sistema óptico, sobre el eje de observación. El eje de proyección y el eje de observación determinan el ángulo de triangulación θ . A partir de la ubicación del punto sobre la imagen y de la geometría del sensor se puede calcular la distancia z . En la parte derecha de la figura se puede observar un ejemplo de cómo se ve el punto laser en la imagen obtenida por la cámara. Se nota que se produce interferencia constructiva y destructiva (debido a la rugosidad del objeto sumado al uso de una fuente de luz laser, es decir, luz coherente). Este efecto hace difícil determinar con precisión la ubicación exacta del punto, y es la fuente de ruido que mencionamos previamente, conocido como ruido speckle¹.

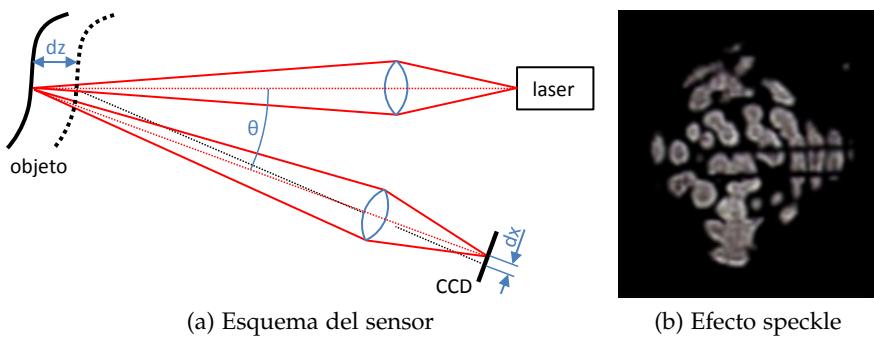


Figura 6: Triangulación láser

Los sensores del tipo 2 incluyen los sensores de tiempo de vuelo y los denominados Coherence Scanning Interferometry ([CSI](#)) en superficies rugosas (también conocidos como radar de coherencia). En la [Figura 7](#) se puede observar el esquema de uno de ellos. Cada haz de luz proveniente de la fuente de luz pasa por una lente colimadora, luego el beam splitter lo divide en 2, de los cuales uno viaja hacia el objeto y el otro hacia un espejo de referencia. El reflejo de ambos vuelve al dispositivo, pasa nuevamente por el beam splitter, luego por otra lente y finalmente llegan al mismo punto del sensor de la cámara. Si la distancia recorrida por cada uno es exactamente igual, se producirá interferencia constructiva, y en caso contrario se produce interferencia destructiva. De esta forma, para cada ubicación del espejo de referencia se obtiene un valor para cada pixel. Evaluando esta función temporal de cada pixel se puede obtener con gran precisión

¹ http://en.wikipedia.org/wiki/Speckle_noise

el valor z correspondiente. Para superficies rugosas la principal fuente de ruido es la micro-topografía de la superficie misma. Un aspecto muy interesante de este tipo de sensores es que la incertezza de la medición no escala con la distancia o con la apertura de observación. Esto permite medir agujeros profundos o realizar mediciones a larga distancia sin incrementar la incertezza de la medición. Sin embargo el método no es muy eficiente ya que requiere tomar una gran cantidad de imágenes variando la posición del espejo de referencia.

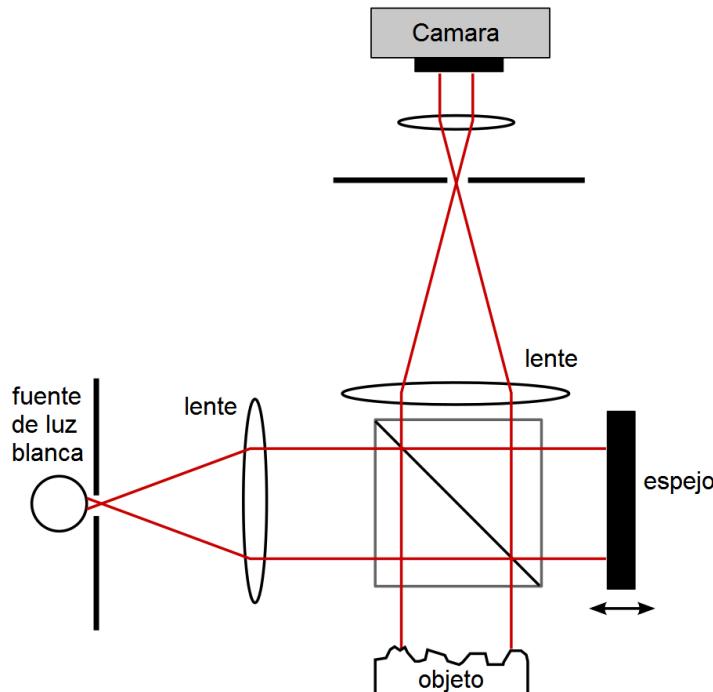


Figura 7: Interferómetro de luz blanca

Los sensores del tipo 3 son los interferómetros clásicos para medir superficies suaves y especulares. La mayor fuente de ruido es el ruido fotónico. La interferometría clásica puede obtener una incertezza de medición de niveles sub-nanómetro, en ambientes adecuados. Los interferómetros clásicos, a diferencia del radar de coherencia, realizan un promediado lateral de la topografía.

Por último, los sensores del tipo 4 incluyen deflectometría y deflectometría de micro-escala. Los sensores del tipo 4 realizan una medición implícita a partir de la variación de la forma de la superficie en vez de medir la altura directamente. La derivada espacial se genera ópticamente, por lo que la topografía obtenida mediante integración numérica puede obtener un nivel de ruido muy bajo, del orden de los nanómetros. En la [Figura 8](#) se puede observar un esquema de un sensor de este tipo.

En la [Tabla 1](#) se muestra una comparación de los distintos tipos de sensores para diversos tipos de superficie. Como se puede observar

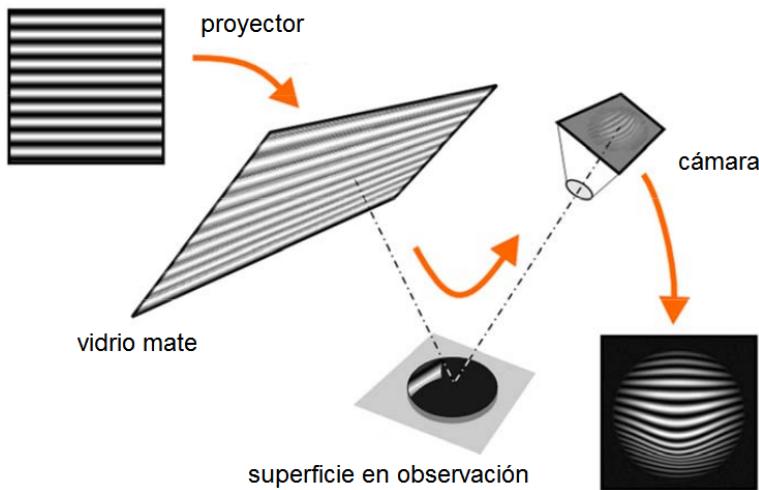


Figura 8: Principio de funcionamiento de un sensor de deflectometría

no existe el sensor ideal, y es de suma importancia la elección del tipo de sensor adecuado para cada situación particular.

	Tipo de sensor				
	Triangulación laser (tipo 1a)	Luz estructurada (tipo 1a)	CSI (tipo 2)	Interferometría clásica (tipo 3)	Deflectometría (tipo 4)
Especular, plana	--	--	++	++	++
Especular, curva	--	--	0	0	++
Mate / Lambertiana	+	++	++	--	--
Maquinada	0	0	++	--	-
Maquinada, inclinada	-	0	+	--	-
Agujeros profundos	--	--	+	--	--
Piel / plástico	-	0	+	--	--

-- inaplicable; - no recomendable; 0 neutro; + aplicable; ++ específico

Tabla 1: Comparación de los distintos tipos de sensores ópticos

2.2 PRINCIPIO DE TRIANGULACIÓN

El principio de triangulación se basa en la premisa de que la luz se mueve en línea recta (en medios homogéneos), lo que permite derivar las ecuaciones de reconstrucción 3D a partir de conceptos básicos de

geometría tridimensional como intersecciones entre líneas y planos, o intersecciones aproximadas entre pares de líneas².

Las técnicas basadas en triangulación más utilizadas son visión estéreo, escáner de línea y luz estructurada. Las tres se basan en el mismo principio de triangulación pero funcionan de manera diferente, y cada una tiene sus ventajas y desventajas. Más adelante se analizará en detalle cada una de ellas, pero antes de avanzar veremos una breve introducción a los principios matemáticos involucrados.

2.2.1 Intersección línea-plano

El cálculo de la intersección entre una línea y un plano es sencillo cuando usamos la representación parámetrica de la línea [6]

$$L = \{p = q_L + \lambda v : \lambda \in \mathbb{R}\}$$

y la representación implícita del plano

$$P = \{p : n^T(p - q_P) = 0\}$$

con q_L un punto perteneciente a la línea y v su vector dirección, siendo n el vector normal al plano y q_P un punto perteneciente a éste.

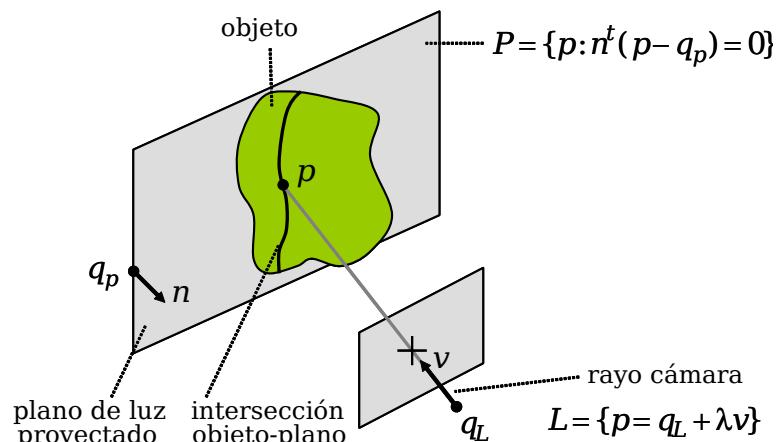


Figura 9: Intersección línea-plano

Como se puede notar, puede que la línea y el plano no se intersecten, en cuyo caso decimos que la línea y el plano son paralelos. Este caso sucede cuando el vector dirección de la línea v y el vector normal al plano n son ortogonales ($n^T v = 0$). También puede suceder que la línea esté contenida en el plano. Si los vectores v y n no son ortogonales, entonces la línea y el plano intersectan en un sólo punto p . Como

² En tres dimensiones dos líneas pueden no intersectarse, aún cuando no son paralelas. El punto de intersección se reemplaza por el punto que minimiza las distancias a las rectas

este punto pertenece a la línea, podemos escribirlo como $p = q_L + \lambda v$, donde λ es el valor que debemos determinar. Como el punto también pertenece al plano, λ debe satisfacer también la ecuación

$$n^T(p - q_P) = n^T(\lambda v + q_L - q_P) = 0$$

o de manera equivalente

$$\lambda = \frac{n^T(q_P - q_L)}{n^T v}$$

2.2.2 Intersección línea-línea

Ahora vamos a considerar la intersección entre dos líneas arbitrarias L_1 y L_2

$$L_1 = \{p = q_1 + \lambda_1 v_1 : \lambda_1 \in \mathbb{R}\}$$

y

$$L_2 = \{p = q_2 + \lambda_2 v_2 : \lambda_2 \in \mathbb{R}\}$$

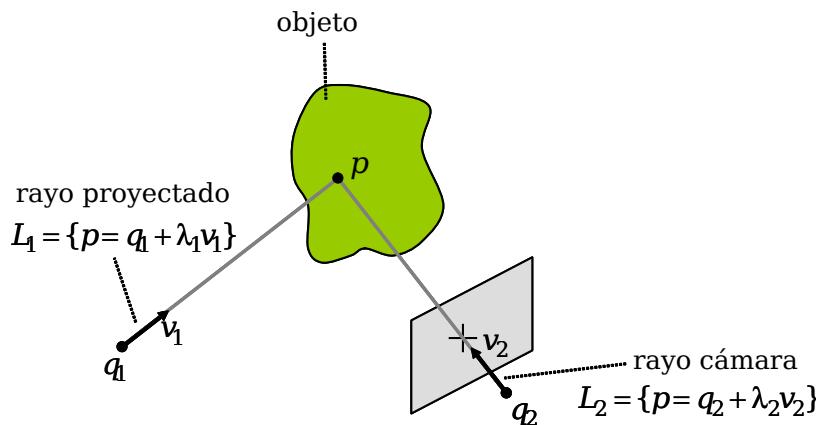


Figura 10: Intersección línea-línea

Primero analizaremos los casos especiales. Los vectores v_1 y v_2 pueden ser linealmente dependientes (es decir uno es un múltiplo escalar del otro) o independientes. Las dos líneas son paralelas si los vectores v_1 y v_2 son linealmente dependientes. Si además el vector $q_2 - q_1$ también es un múltiplo de v_1 o v_2 , entonces las líneas son idénticas. Por otro lado, en el caso que las líneas son paralelas pero no identicas, las líneas no se intersectan. Si v_1 y v_2 son linealmente independientes, las líneas pueden intersectar o no. Si las líneas intersectan, la intersección contiene un sólo punto. En este caso, las condiciones necesarias y suficientes para que dos líneas se intersecten es que los valores escalares λ_1 y λ_2 existan tal que

$$q_1 + \lambda_1 v_1 = q_2 + \lambda_2 v_2$$

o de manera equivalente que el vector $q_2 - q_1$ sea linealmente dependiente de v_1 y v_2 .

Como dos líneas pueden no intersectar (lo que sucede muchas veces), vamos a definir la intersección aproximada como el punto que esté más cerca de las dos líneas. Para ser más precisos, vamos a definir la intersección aproximada como el punto p que minimiza la suma del cuadrado de las distancias a ambas líneas, sin importar si las líneas intersectan o no

$$\phi(p, \lambda_1, \lambda_2) = \|q_1 + \lambda_1 v_1 - p\|^2 + \|q_2 + \lambda_2 v_2 - p\|^2$$

Al igual que anteriormente, asumimos que v_1 y v_2 son linealmente independientes de manera tal que la intersección es un punto único.

Para determinar el valor de p vamos a utilizar un enfoque algebraico. La función $\phi(p, \lambda_1, \lambda_2)$ es una función cuadrática y no-negativa de cinco variables, las tres coordenadas del punto p y los dos escalares λ_1 y λ_2 .

Primero procedemos a reducir el problema a la minimización de una función cuadrática no-negativa de sólo dos variables λ_1 y λ_2 . Para ello comenzaremos definiendo los puntos $p_1 = q_1 + \lambda_1 v_1$ y $p_2 = q_2 + \lambda_2 v_2$, y el punto p_{12} como el punto medio entre ellos:

$$p_{12} = p_1 + \frac{1}{2}(p_2 - p_1) = p_2 + \frac{1}{2}(p_1 - p_2)$$

Una condición necesaria para el mínimo $(p, \lambda_1, \lambda_2)$ de ϕ es que las derivadas parciales de ϕ respecto a las cinco variables, todas se hagan cero en el mínimo. En particular, las tres derivadas respecto a las coordenadas del punto p deben cumplir:

$$\frac{\partial \phi}{\partial p} = (p - p_1) + (p - p_2) = 0$$

Por lo tanto es necesario que el punto mínimo p sea el punto medio p_{12} del segmento que une p_1 con p_2 .

Con este resultado el problema se reduce a la minimización del cuadrado de la distancia de un punto p_1 en la línea L_1 hacia un punto p_2 en la línea L_2 , es decir, sólo debemos minimizar la función cuadrática no-negativa de dos variables:

$$\psi(\lambda_1, \lambda_2) = 2\phi(p_{12}, \lambda_1, \lambda_2) = \|(q_2 + \lambda_2 v_2) - (q_1 + \lambda_1 v_1)\|^2$$

Como se mencionó previamente, también es necesario que las dos derivadas parciales de ψ respecto a λ_1 y λ_2 sean iguales a cero en el mínimo:

$$\frac{\partial \psi}{\partial \lambda_1} = v_1^T (\lambda_1 v_1 - \lambda_2 v_2 + q_1 - q_2) = \lambda_1 \|v_1\|^2 - \lambda_2 v_1^T v_2 + v_1^T (q_1 - q_2) = 0$$

$$\frac{\partial \psi}{\partial \lambda_2} = v_2^T (\lambda_2 v_2 - \lambda_1 v_1 + q_2 - q_1) = \lambda_2 \|v_2\|^2 - \lambda_1 v_2^T v_1 + v_2^T (q_2 - q_1) = 0$$

Estas ecuaciones lineales en λ_1 y λ_2 pueden ser expresadas en forma matricial como:

$$\begin{bmatrix} \|\mathbf{v}_1\|^2 & -\mathbf{v}_1^T \mathbf{v}_2 \\ -\mathbf{v}_2^T \mathbf{v}_1 & \|\mathbf{v}_2\|^2 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} \mathbf{v}_1^T (\mathbf{q}_2 - \mathbf{q}_1) \\ \mathbf{v}_2^T (\mathbf{q}_1 - \mathbf{q}_2) \end{bmatrix}$$

Debido a la independencia lineal entre \mathbf{v}_1 y \mathbf{v}_2 la matriz de 2×2 de la izquierda es no-singular. De esta forma la solución única queda dada por:

$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} \|\mathbf{v}_1\|^2 & -\mathbf{v}_1^T \mathbf{v}_2 \\ -\mathbf{v}_2^T \mathbf{v}_1 & \|\mathbf{v}_2\|^2 \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{v}_1^T (\mathbf{q}_2 - \mathbf{q}_1) \\ \mathbf{v}_2^T (\mathbf{q}_1 - \mathbf{q}_2) \end{bmatrix}$$

o de manera equivalente:

$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \frac{1}{\|\mathbf{v}_1\|^2 \|\mathbf{v}_2\|^2 - (\mathbf{v}_1^T \mathbf{v}_2)^2} \begin{bmatrix} \|\mathbf{v}_2\|^2 & \mathbf{v}_1^T \mathbf{v}_2 \\ \mathbf{v}_2^T \mathbf{v}_1 & \|\mathbf{v}_1\|^2 \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^T (\mathbf{q}_2 - \mathbf{q}_1) \\ \mathbf{v}_2^T (\mathbf{q}_1 - \mathbf{q}_2) \end{bmatrix}$$

Finalmente la intersección aproximada puede obtenerse a partir del valor de λ_1 o λ_2 .

2.3 VISIÓN ESTÉREO

Visión estéreo consiste en dos o más cámaras que observan el mismo objeto desde diferentes ubicaciones, y la triangulación se obtiene a partir de la identificación de los pixels en cada imagen que corresponden a un mismo punto del objeto. Esta técnica tiene la ventaja de ser rápida (sólo requiere una imagen de cada cámara) pero la dificultad en establecer las correspondencias entre las imágenes hace que su precisión no sea buena. Por otro lado su aplicación resulta sumamente limitada en el caso de objetos homogéneos ya que se torna muy difícil (en ocasiones imposible) encontrar las correspondencias. En la [Figura 11](#) se puede observar un ejemplo para ayudar a entender el principio de funcionamiento de este tipo de sensores.

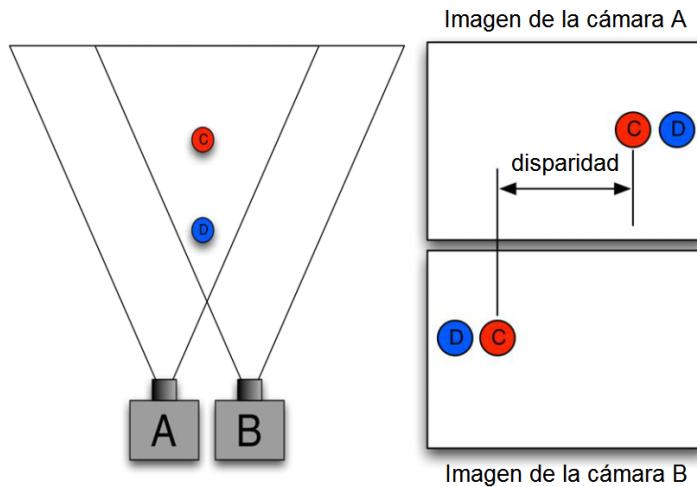


Figura 11: Visión estéreo

2.4 ESCÁNER DE LÍNEA

Un escáner de línea consiste en un generador de una línea de luz (generalmente un láser) y una cámara que observa la incidencia de la línea de luz sobre el objeto, como se observa en la [Figura 12](#). La reflexión de la luz que incide sobre el objeto es captada por la cámara y, conociendo la distancia y el ángulo entre el plano de luz y la cámara, a partir de una triangulación se puede obtener el perfil del objeto en la zona iluminada. Estos sistemas permiten obtener muy buena precisión, pero se debe notar que sólo se obtiene la geometría del objeto sobre una línea. Para obtener la geometría del objeto completo (o de una zona de interés) hace falta mover una de las partes (el objeto o el escáner), lo que implica un incremento considerable en el tiempo de medición y agrega una nueva fuente de error debido al sistema de desplazamiento, entre otras cosas.

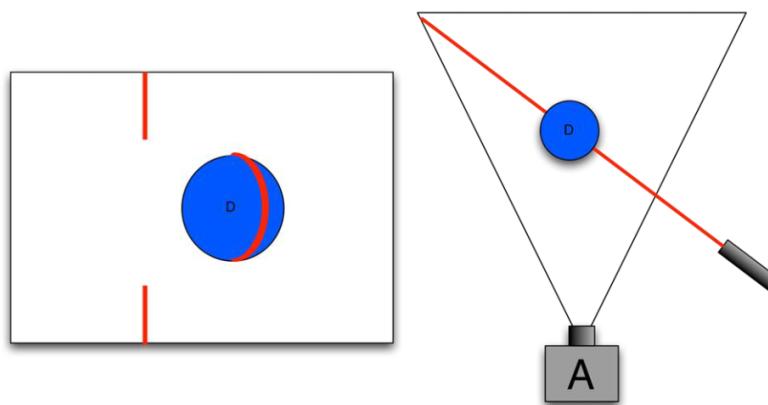


Figura 12: Escáner de línea

2.5 LUZ ESTRUCTURADA

Los sistemas de luz estructurada aceleran el proceso de medición a partir de la proyección de patrones de luz estructurada sobre el objeto. Una cámara observa la distorsión de estos patrones al ser reflejados y, al igual que los métodos anteriores, mediante triangulación se obtiene información tridimensional del objeto. Su principal ventaja es la flexibilidad, ya que la precisión, resolución y velocidad dependerán en gran medida del tipo de patrones que se proyectan (entre otras cosas).

Los sistemas de luz estructurada consisten de un generador de patrones de luz y al menos una cámara. Un patrón de luz 1D (línea) o 2D (grilla) es proyectado sobre el objeto y la cámara se utiliza para determinar la deformación producida. En la [Figura 13](#) se muestra el esquema de un sistema de luz estructurada típico.

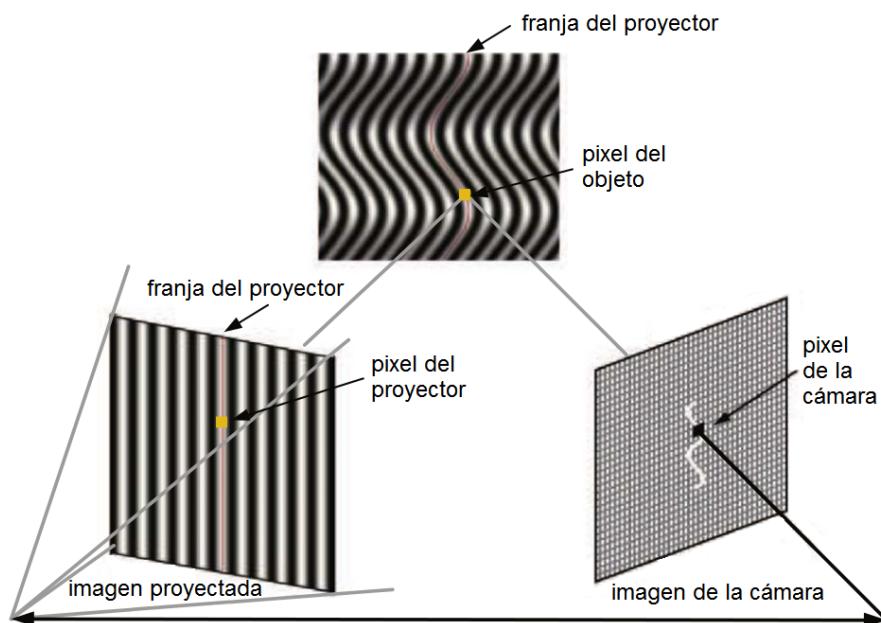


Figura 13: Sistema de luz estructurada típico

Los patrones de luz estructurada se diseñan de manera tal que cada pixel o linea de pixeles tenga un código único que pueda ser capturado por la cámara. Solamente se puede medir áreas iluminadas por el proyector y a la vez observadas por la cámara, por lo que si el objeto es muy grande o se quiere medir toda la vuelta se deberá mover el objeto (o el sistema de medición).

En la [Figura 14](#) se muestra la secuencia de pasos estándar de un sistema de medición basado en luz estructurada.

El primer paso es la calibración del sistema para obtener la posición relativa del proyector de patrones y de la/s cámara/s, y para mejorar la exactitud y precisión considerando la distorsión producida por la óptica de la/s cámara/s y, en el caso que sea necesario, del proyector.

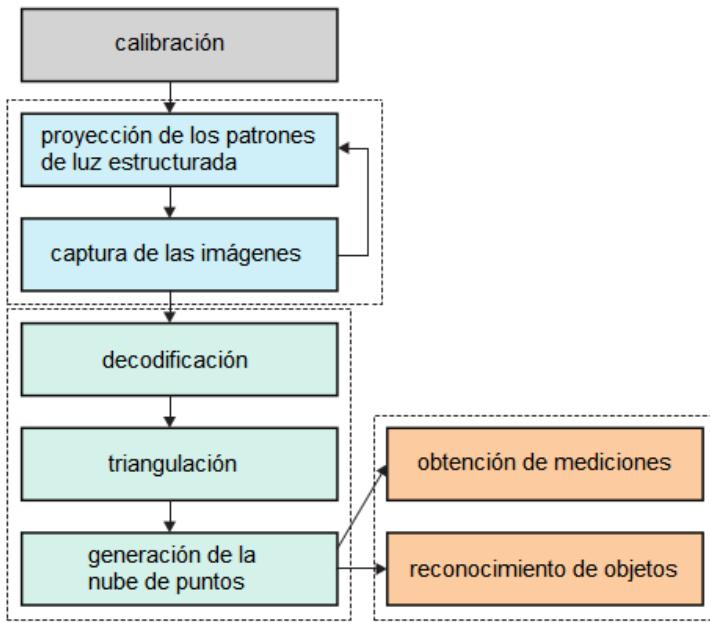


Figura 14: Secuencia de pasos un sistema de luz estructurada

El paso siguiente es la proyección de el/los patron/es de luz estructurada y la obtención de la imagen correspondiente a cada uno de ellos. Una vez obtenidas las imágenes, estas deben ser procesadas para obtener la información requerida. La primer tarea es decodificar los patrones, es decir obtener la correspondencia entre cada pixel de la imagen y la línea o pixel del patrón de luz estructurada. Al igual que un escáner de línea (en el caso de patrones de líneas -1D-) o un sistema de visión estéreo (para patrones 2D), a partir de las correspondencias se utiliza la triangulación para obtener información sobre los puntos en tres dimensiones. Este conjunto de puntos 3D es conocido típicamente como nube de puntos. La nube de puntos no contiene información sobre la conexión entre pares de puntos, por lo que en ciertos casos se debe utilizar otro método para reconstruir la superficie a partir de la nube de puntos.

2.5.1 Patrones de luz estructurada

A través de los años se desarrollaron diversos tipos de patrones de luz estructurada. La elección del tipo apropiado depende de los requerimientos de resolución espacial, de las características del objeto a medir, de la robustez de la decodificación, de la velocidad, etc. A continuación se brinda un pequeño resumen de los más utilizados.

2.5.1.1 Patrones 2D

El caso más conocido de la utilización de patrones de luz estructurada 2D es el sensor 3D Microsoft Kinect³, desarrollado originalmente para la consola de videojuegos XBOX 360⁴. Este sistema utiliza luz infrarroja para proyectar un patrón de puntos sobre la escena, los cuales son observados por una cámara ubicada a un costado. Los puntos establecen una codificación 2D. La calibración del sistema permite obtener la correspondencia entre un grupo de puntos y una zona de la imagen. Este tipo de patrones permite obtener la información requerida para la triangulación a partir de una sola imagen, como es el caso de visión estéreo, pero el uso del patrón de luz estructurada facilita y hace más robusta la detección de correspondencias. Esto permite lograr altas velocidades y observar escenas dinámicas (como es el caso requerido para el control de una consola de videojuegos), pero tienen como desventaja una pobre resolución espacial/lateral.



Figura 15: Patrón infrarrojo usado por el sensor Microsoft Kinect

2.5.1.2 Patrones 1D binarios

Los patrones binarios utilizan solamente dos valores, 0 o 1, los cuales son proyectados secuencialmente de manera que cada pixel es asociado a una cadena única de valores binarios. Generalmente se utilizan en líneas, tanto verticales como horizontales, de manera que $\log_2(m)$ patrones son requeridos para codificar m filas (o columnas) de pixeles.

³ <http://en.wikipedia.org/wiki/Kinect>

⁴ http://en.wikipedia.org/wiki/Xbox_360

La decodificación de este tipo de patrones es muy robusta debido a que cada pixel corresponde a una única fila (o columna). Sin embargo la gran cantidad de patrones requeridos imposibilita su uso en escenas dinámicas (donde el objeto está en movimiento), o bien requieren que tanto la cámara como el proyector sean capaces de trabajar a altas velocidades de adquisición/proyección.

Para generar la cadena única de valores binarios que identifican las filas (o columnas) de los patrones se utiliza generalmente un código Gray⁵. El uso de este tipo de código binario reduce las consecuencias en caso de producirse un error de decodificación. Esto se debe a que la distancia de Hamming entre códigos de Gray sucesivos es de uno. Para entender mejor supongamos un pixel que cae en el límite entre una columna y otra (ver pixel indicado con recuadro rojo en la Figura 16). En el caso de un código Gray, este límite será observado en la imagen de un solo patrón, mientras que si se utiliza un código binario estándar es probable que aparezca en más de una imagen. Si se produce un error de decodificación en este pixel, en una imagen, en el caso del código Gray solamente tendremos un error de ± 1 columna, mientras que en el otro caso la diferencia puede llegar a ser mucho mayor.

Gray	Binario
000	000
001	001
011	010
010	011
110	100
111	101
101	110
100	111

Figura 16: Diferencias entre patrones usando código Gray y código binario estándar

2.5.1.3 Patrones 1D phase shift

Estos patrones de luz consisten en franjas donde la intensidad de la luz es modulada de acuerdo a una función, por ejemplo sinusoidal, como puede verse en Figura 17.

Los diversos patrones son generados a partir de un cambio de fase de la función. A partir de los valores detectados en cada pixel se puede conocer la fase relativa. Debido a que los patrones son periódicos se necesita un paso más, conocido comúnmente como *phase-*

⁵ http://en.wikipedia.org/wiki/Gray_code

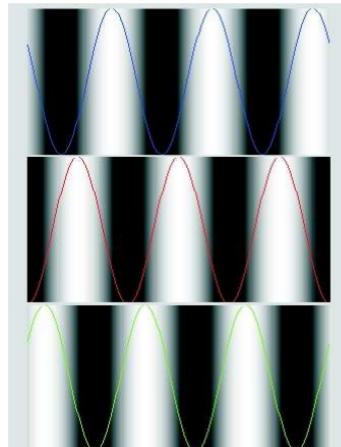


Figura 17: Codificación con tres patrones sinusoidales desfasados

unwrapping, para obtener la fase absoluta. Esta etapa utiliza información de píxeles vecinos para conocer cuando se pasó de una franja a la siguiente, en otras palabras cuando estamos en el pixel con fase 359° y pasamos a fase 0° , podemos suponer que en realidad la fase absoluta es 360° , al siguiente 361° (en vez de 1°), y así sucesivamente.

En principio alcanza con sólo 3 patrones desfasados 120 para obtener la correspondencia entre cada pixel de la imagen y su fase absoluta dentro del patrón, por lo que este método resulta muy rápido y permite ser utilizado para medir escenas dinámicas. Otra ventaja es la mejora en resolución debido a la naturaleza misma de la luz y las lentes utilizadas. lo que hace que el patrón sea más suave de lo que se proyecta (por ejemplo aprovechando el efecto de desenfoque). Esto significa que cada pixel de la cámara brinda información útil, a diferencia de los patrones binarios donde la única información válida son los límites entre una franja y la siguiente.

La principal desventaja viene dada por la periodicidad de los patrones, por lo que cambios bruscos en profundidad (por ejemplo bordes de un objeto) hacen imposible obtener correctamente la fase absoluta. Para compensar esta situación se suelen utilizar más patrones, o incluso se combinan con patrones binarios, lo que reduce su ventaja en velocidad.

3

CALIBRACIÓN DE LA CÁMARA

Este capítulo se dedica exclusivamente a la calibración de la cámara. Este es uno de los aspectos más importantes del proyecto, considerando que la cámara es el principal elemento que utilizaremos para realizar mediciones.

En la primer sección se realiza una introducción para entender mejor a qué nos referimos cuando hablamos de la calibración de la cámara. En la segunda sección se explicarán de manera sencilla algunos conceptos de la óptica relacionados con el funcionamiento de una cámara fotográfica. En la tercera sección se explica el funcionamiento de una cámara de fotos digital y la manera en la que se obtienen las imágenes. En la cuarta sección se introduce un modelo de cámara simple conocido como *cámara pinhole*. En la quinta sección se extiende el modelo de cámara pinhole para incluir las distorsiones ópticas más frecuentes. En la sexta sección se explica el concepto de parámetros extrínsecos de una cámara. La séptima sección puede considerarse la más importante, en la cual se explica un método para obtener los parámetros del modelo de cámara, incluyendo los parámetros extrínsecos y los que modelan la distorsión. Finalmente en la última sección se trata brevemente la calibración conjunta de más de una cámara.

3.1 INTRODUCCIÓN

Una cámara fotográfica realiza un mapeo desde el mundo en tres dimensiones (3D) hacia una imagen (dos dimensiones, 2D). A lo largo de los años diversos modelos matemáticos han sido propuestos para describir este mapeo. En palabras simples, un modelo matemático de una cámara describe cómo se proyectan los objetos en tres dimensiones sobre la imagen y viceversa, cómo proyectar desde la imagen hacia el espacio tridimensional. La calibración de una cámara consiste en determinar los parámetros de este modelo. Este problema incluye el modelado y la parametrización del proceso completo de obtención de las imágenes.

Los modelos de cámaras pueden ser clasificados de acuerdo a diferentes criterios, como por ejemplo en el uso de la suposición (o no) de que existe un único punto por el que pasan todos los rayos de luz. Los modelos que utilizan esta suposición son conocidos como

modelos de cámara centrales, y el resto como no centrales. Otra clasificación puede hacerse entre modelos globales, locales o discretos. En los modelos globales el cambio de un parámetro afecta a todo el campo de visión, en cambio en los modelos locales el cambio de un parámetro afecta sólo a una zona. Por otro lado un modelo discreto es aquel que tiene un grupo de parámetros para cada punto de la imagen [7].

En este trabajo vamos a utilizar un modelo de cámara central y global, conocido generalmente como modelo de cámara pinhole. El uso de un modelo de cámara central nos permite formular el mapeo como una proyección. Esto nos posibilita utilizar todos los conceptos de la geometría proyectiva [8].

Muchas aplicaciones de visión por computadora no requieren alta precisión en la reconstrucción. Por ejemplo en robótica muchas veces solamente es necesario obtener información sobre un entorno desconocido. En estos casos puede ser más importante estimar la orientación del sensor que medir distancias con exactitud. En otros casos donde se busca reconstruir dimensiones con exactitud -como el planteado en este trabajo-, el modelo de la cámara pinhole debe ser extendido con un modelo de distorsión que generalmente se define en el plano de la imagen. Este modelo, aún siendo sólo una aproximación del comportamiento real de la cámara, nos permitirá incrementar considerablemente la calidad de nuestros resultados.

3.2 SISTEMA ÓPTICO DE LA CÁMARA

Las cámaras funcionan detectando rayos de luz. Los rayos provenientes de alguna fuente de luz viajan por el espacio hasta los objetos. El objeto absorbe parte de esa luz y el resto es reflejado. Parte de estos rayos pueden ser reflejados hacia la cámara, donde se produce una imagen. Para poder obtener información del mundo real a partir de la imagen necesitamos conocer en detalle lo que sucede con el rayo de luz al pasar por la lente y en qué lugar es capturado por el sensor de la cámara.

En nuestro caso vamos a analizar el sistema óptico con la ayuda de la óptica geométrica. Se define la óptica geométrica como aquella que abarca el estudio de los fenómenos relativos a la propagación de la luz sin incluir los efectos de interferencia ni de difracción, considerando los objetos compuestos por un conjunto de fuentes radiantes puntuales independientes. Esta descripción, basada en el análisis de las trayectorias (rayos) de propagación de la energía, es válida siempre que la longitud de onda de la perturbación que se desplaza sea mucho menor que las dimensiones características de los objetos con los que se encuentra [9].

También asumiremos que la lente es simétrica alrededor de un eje de rotación que denominaremos eje óptico de la lente.

3.2.1 Óptica de primer orden

La óptica geométrica se basa en la ley de refracción, también conocida como ley de Snell. Dados dos medios con índices de refracción n_1 y n_2 y un rayo de luz que pasa desde el medio 1 al medio 2, como se observa en la [Figura 18](#), la ley de refracción establece que los ángulos θ_1 y θ_2 del rayo respecto a la normal de la interfaz entre los medios es:

$$n_1 \sin(\theta_1) = n_2 \sin(\theta_2)$$

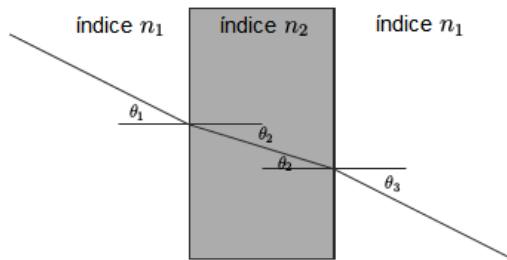


Figura 18: Rayo de luz atravesando un medio

En el ámbito de la visión por computadora generalmente se acepta que el índice de refracción del aire es tan cercano a 1 que se trata como si fuera igual a 1 [8]. Una lente esférica tiene dos interfaces entre medios (aire-lente y lente-aire) los cuales se describen como dos esferas con el mismo radio r , como se observa en la [Figura 19](#).

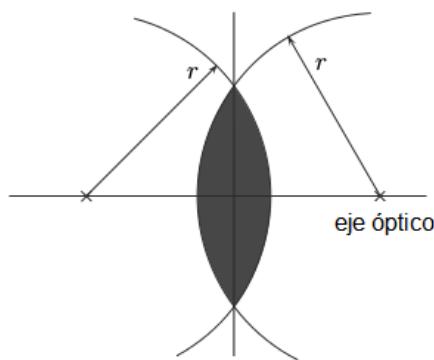


Figura 19: Lente esférica

Mediante la ley de Snell podemos reconstruir la refracción de cada rayo proveniente de un objeto que pasa por la lente.

En óptica de primer orden se asume que todos los rayos de luz considerados en el modelo de la cámara son casi paralelos al eje óptico de la lente. En este caso todos los ángulos respecto a la normal de la superficie de la lente son pequeños, por lo que podemos aproximar

el seno como $\sin(x) \approx x$. Debido a que esta aproximación es el primer término de la aproximación de Taylor para el seno, las derivaciones que siguen esta suposición son conocidas como *óptica de primer orden*.

Otra suposición que se realiza comúnmente en visión por computadora es llamada suposición de lente delgada: se asume que la lente es infinitesimalmente fina. Desde el punto de vista de la refracción una lente delgada se comporta como una lente esférica, pero la distancia recorrida por el rayo en el interior de la lente es infinitesimalmente pequeña. Por lo tanto podemos asumir que la lente es un plano. Este plano es generalmente llamado *plano principal*. Un rayo de luz pasando por una lente de estas características es afectado por la ley de Snell en la interfaz aire-lente e inmediatamente afectado por la ley de Snell en la interfaz lente-aire. Una consecuencia de ésto es que cualquier rayo de luz que pase por el eje óptico de la lente no será afectado por la refracción, debido a que la superficie de ambas interfaces en este punto es paralela y la distancia entre ellas es nula. Otra consecuencia de este modelo de lente combinado con la simplificación de que todos los rayos son casi paralelos es que todos los rayos emitidos en un punto p en un lado de la lente, al pasar por la lente se vuelven a encontrar en un punto p_i al otro lado de la lente de manera que:

$$\frac{1}{d_p} + \frac{1}{d_i} = \frac{1}{f}$$

donde

$$f = \frac{r}{2(n_l - 1)}$$

siendo n_l el índice de refracción de la lente, y d_p y d_i la distancia del punto p y p_i al plano principal. Esto significa que hay una relación entre los objetos y la imagen, y que esta relación depende solamente de la distancia entre el objeto y el plano principal, y que no depende de la distancia hacia el eje óptico. Esta ecuación es conocida como *ecuación del fabricante de lentes* (*lens' maker equation*).

Existe una distancia detrás del plano principal (d_i en la [Figura 20](#)) para la cual la observación de un punto p se transforma en una imagen enfocada. Se puede decir que cada plano paralelo al plano principal determina un plano focal detrás de la lente donde los puntos se ven enfocados.

En la [Figura 20](#) se pueden observar tres rayos de luz que pasan por el modelo de lente delgada:

- El rayo que proviene de un lado en dirección paralela al eje óptico de la lente pasa por un punto F al otro lado de la lente, ubicado a una distancia f del plano principal. F es denominado punto focal. La distancia f es llamada *distancia focal posterior*.
- Un rayo que pasa por el centro de la lente no cambia su dirección. Este rayo es denominado rayo central. El centro de la lente es llamado *centro óptico de la lente*.

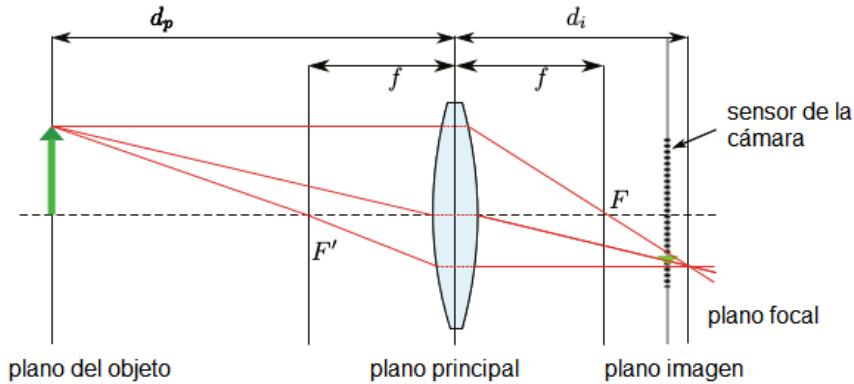


Figura 20: Rayo de luz atravesando un medio

- Un rayo que arriba a la lente desde el punto F' (ubicado a una distancia f' del plano principal) sale de manera paralela al eje óptico del otro lado. La distancia de F' al plano principal es llamada *distancia focal anterior*.

A lo largo de este trabajo consideraremos una lente esférica con óptica de primer orden y con distancia focal posterior igual a la distancia focal anterior. Para una lente delgada con óptica de primer orden la distancia focal es independiente de la distancia del punto/objeto al eje óptico.

3.2.2 *El círculo de confusión*

Generalmente la adquisición de la imagen se produce mediante un sensor plano. Llamaremos a este plano donde se obtiene la imagen *plano imagen*. Generalmente el plano imagen no coincide con el plano focal, por lo tanto no todos los rayos provenientes de los objetos en observación se cruzan en un punto sobre el plano imagen.

Si consideramos una fuente de luz puntual en el lado del objeto, todos los rayos provenientes de este punto que pasan la lente forman un cono en el otro lado de la lente. La intersección de este cono con el plano imagen se denomina *círculo de confusión*.

Si el círculo de confusión es más pequeño que el tamaño de un elemento del sensor, el objeto aparecerá nítido, bien definido. El rango de posiciones del objeto (distancias a la cámara en dirección del eje focal) en el cual esto es cierto se denomina *profundidad de campo* (*depth of field*).

Esta dependencia entre nitidez y profundidad puede ser usada para estimar la distancia a un objeto (la distancia del objeto al plano principal), y es utilizada por las técnicas denominadas *depth from focus* ó *depth from de-focus*.

Este efecto borroso de la óptica de primer orden también puede ser tenido en cuenta en el modelado de la cámara. La manera más común de manejar este efecto es mediante la convolución de la *imagen ideal*

con un kernel. La *imagen ideal* es la imagen obtenida a partir de los rayos centrales únicamente, lo cual en la vida real es impracticable. En visión por computadora este kernel es también conocido como Point Spread Function ([PSF](#)). El [PSF](#) puede ser visto como la respuesta al impulso del sistema óptico. Existen diferentes maneras de estimar el [PSF](#), aunque en este trabajo no lo tendremos en cuenta.

3.2.3 Óptica de tercer orden

Como se dijo previamente, la simplificación $\sin(x) \approx x$ solamente es válida para ángulos pequeños. Un modelo de refracción más realista se obtiene truncando la serie de Taylor del seno en el término de tercer orden: $\sin(x) \approx x - \frac{1}{3!}x^3$. Esto nos lleva a la denominada óptica de tercer orden. Se puede comprobar fácilmente que en óptica de tercer orden la refracción de un rayo de luz si depende de la distancia al eje óptico.

También podemos expandir el modelo de lente al denominado modelo de lente gruesa, que puede ser modelado como dos lentes delgadas paralelas. Este modelo sumado a la óptica de tercer orden permiten modelar las aberraciones monocromáticas de un sistema óptico. A continuación describiremos rápidamente las denominadas cinco aberraciones de Seidel, las cuales pueden ser modeladas bajo las suposiciones de este modelo [8].

3.2.3.1 Aberraciones esféricas

Rayos que inciden sobre la superficie de un lente esférico a mayor distancia del eje óptico enfocarán más cerca que rayos a menor distancia. Esto lleva a la aparición de un círculo de confusión incluso en el plano focal. Esta aberración puede ser eliminada usando una lente oval en vez de una esférica, pero estas lentes tienen un costo más elevado. Se puede observar un ejemplo de esta aberración en la [Figura 21](#).

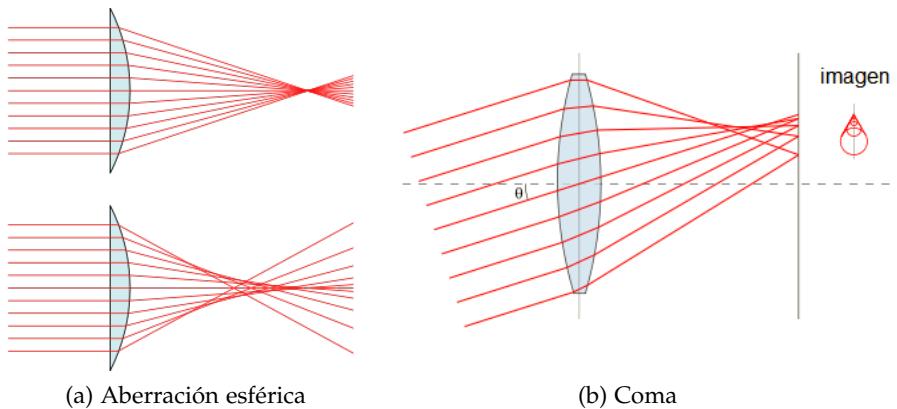


Figura 21: Aberraciones

3.2.3.2 Coma

Rayos provenientes de un objeto que no está en el eje óptico serán enfocados en diferentes puntos. En sistemas no óptimos el enfoque puede ser asimétrico. La imagen de un punto aparece como gota en vez de como un círculo. Esta aberración afecta en mayor medida a los telescopios pero igualmente debemos tener en cuenta que transforma el círculo de confusión en una figura asimétrica. Se puede observar un ejemplo de esta aberración en la [Figura 21](#).

3.2.3.3 Astigmatismo

Rayos provenientes de un objeto alejado del eje óptico no incidirán de manera simétrica sobre la superficie de la lente. Esto hace que los rayos no enfoquen en un mismo punto, provocando una deformación elíptica del círculo de confusión. Se puede observar un ejemplo de esta aberración en la [Figura 22](#).

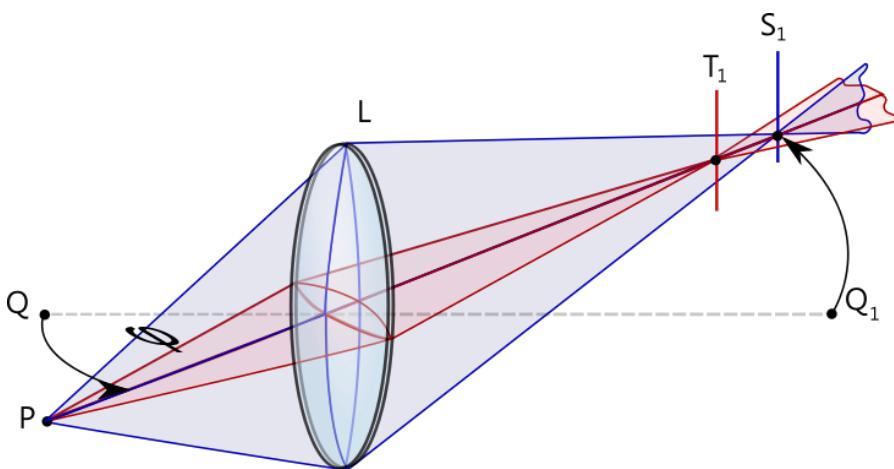


Figura 22: Astigmatismo

3.2.3.4 Curvatura de campo

El comportamiento de la refracción en la óptica de tercer orden también impacta la relación entre objetos y puntos en la imagen. Para un plano objeto paralelo al plano principal de la lente, el área donde los puntos estarán enfocados no será un plano sino una superficie curva¹. Sólo se puede aproximar con un plano para puntos cercanos al eje óptico. Esto significa que no es teóricamente posible que todos los pixels de una imagen estén en foco si usamos una cámara con un array plano de sensores. Este efecto sin embargo no es significativo debido a que el tamaño del sensor es usualmente muy pequeño en comparación al objeto, aunque puede ocurrir por ejemplo en las

¹ Esta aberración también es conocida como *Petzval field curvature*, en honor a Josef Petzval, que fue el primero en analizar este efecto

fotografías en modo *macro*. Se puede observar un ejemplo de esta aberración en la [Figura 23](#).

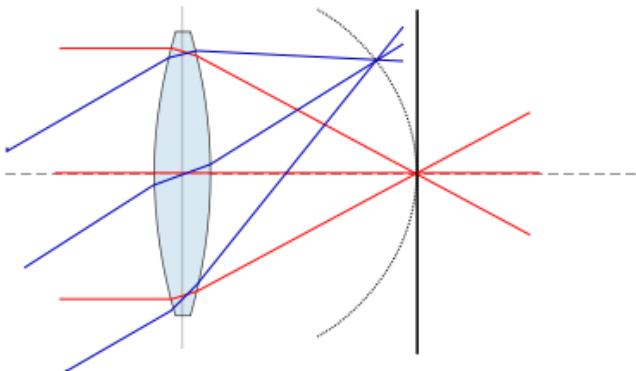


Figura 23: Curvatura de campo

3.2.3.5 Distorsión

En óptica de tercer orden la magnificación transversal en el plano de la imagen se transforma en una función de la distancia al eje óptico, es decir radial. A diferencia de las aberraciones anteriores, la distorsión también afecta los rayos centrales. La distorsión (como la describe Seidel) está determinada completamente en el plano de la imagen. Se puede observar ejemplos de este tipo de aberración en la [Figura 24](#).

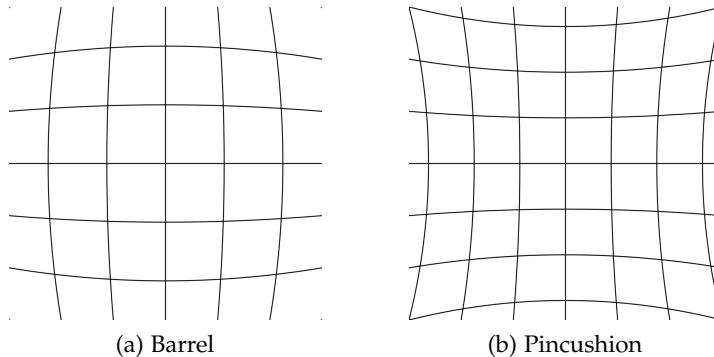


Figura 24: Ejemplos de distorsión

3.3 ADQUISICIÓN DE IMÁGENES

El sensor de la cámara es un dispositivo electrónico que convierte la intensidad de los rayos de luz en una señal eléctrica. Generalmente se trata de una grilla rectangular de elementos sensibles a la luz, y los más usados son del tipo Charge-Coupled Devices ([CCD](#)) ó Complementary Metal-Oxide Semiconductors ([CMOS](#)). Un esquema de la estructura de un sensor puede observarse en la [Figura 25](#). Esta

grilla rectangular establece el sistema de coordenadas canónico para la imagen.

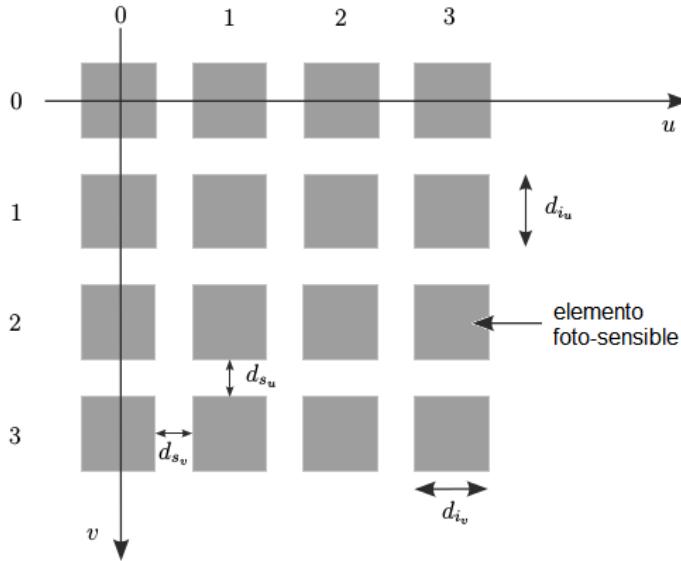


Figura 25: Esquema del sensor de una cámara

Cada sensor de la grilla determina el valor correspondiente a un elemento de la imagen (pixel), es decir que cada pixel representa un área rectangular del plano de la imagen.

El modelo simplificado del sensor de la cámara no considera:

- la superficie no foto-sensitiva presente entre cada elemento foto-sensitivo del array
- la interacción entre los elementos foto sensitivos y sus vecinos, por ejemplo el ruido causado durante la lectura de los valores de cada elemento en cámaras de bajo costo,
- las diferencias entre los elementos foto-sensitivos. Se asume que todos tienen el mismo tamaño y las mismas características

3.4 MODELO DE CÁMARA PINHOLE

Generalmente denominamos lente a lo que en realidad es un sistema de lentes, pero a fines prácticos lo que nos interesa es que las propiedades ópticas del sistema de lentes se aproximan a las propiedades ópticas de una lente delgada *virtual*. El modelo pinhole se basa en el rayo central del modelo de lente delgada. El rayo central pasa por el centro de la lente para cada posición del objeto. Esto significa que podemos modelar la cámara como una *cámara oscura*, con el centro óptico del lente como el *pinhole*, como se observa en la [Figura 26](#).

Para explicar el modelo de cámara pinhole vamos a definir dos sistemas de coordenadas: el sistema de coordenadas de la cámara

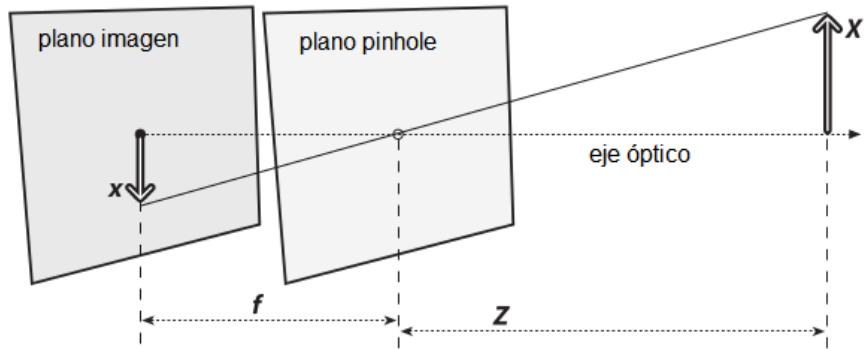


Figura 26: Modelo de cámara pinhole

(CCS) y el sistema de coordenadas de la imagen (ICS). El sistema de coordenadas de la cámara es un sistema de coordenadas cartesiano definido por el plano principal: los ejes x e y del CCS determinan el plano principal mientras que el eje z viene dado por el eje óptico. El centro óptico del lente determina el origen $(0,0,0)$ del CCS. El plano principal queda definido entonces por $z = 0$.

El sistema de coordenadas de la imagen es el sistema de coordenadas canónico del array de sensores de la cámara. En nuestro caso vamos a asumir que el primer eje del sistema de coordenadas de la cámara es paralelo al primer eje del sistema de coordenadas de la imagen. La intersección del plano imagen con el eje óptico se denomina punto principal.

En este momento vamos a modificar el modelo pinhole a una forma equivalente pero que nos va a facilitar la comprensión: vamos a pensar que el plano de la imagen está delante del plano principal, como se muestra en la [Figura 27](#), es decir consideraremos que la imagen se forma en un plano virtual delante del centro óptico. El centro óptico se puede ver entonces como un centro de proyección.

El hecho de que en el modelo pinhole todos los rayos desde el objeto hacia la imagen se intersectan en un punto (el centro óptico) nos permite utilizar una transformación proyectiva: la relación que asigna un punto Q en el mundo físico con coordenadas (X, Y, Z) a un punto q del plano de proyección con coordenadas (x, y) se llama transformación proyectiva. Cuando se trabaja con transformaciones de este tipo es conveniente usar lo que se conoce como coordenadas homogéneas. La coordenada homogénea asociada a un punto en un espacio proyectivo de dimensión n se expresa típicamente como un vector de dimensión $n + 1$, con la restricción adicional de que dos puntos cualesquiera cuyos valores son proporcionales serán equivalentes. En nuestro caso, el espacio proyectivo es el plano de la imagen, que posee dos dimensiones, por lo que vamos a representar los puntos que en él se encuentran como tridimensionales: $q = (x, y, w)$. Es posible recuperar las coordenadas en el plano de la imagen dividiendo las dos primeras componentes del vector por la tercera. De esta manera,

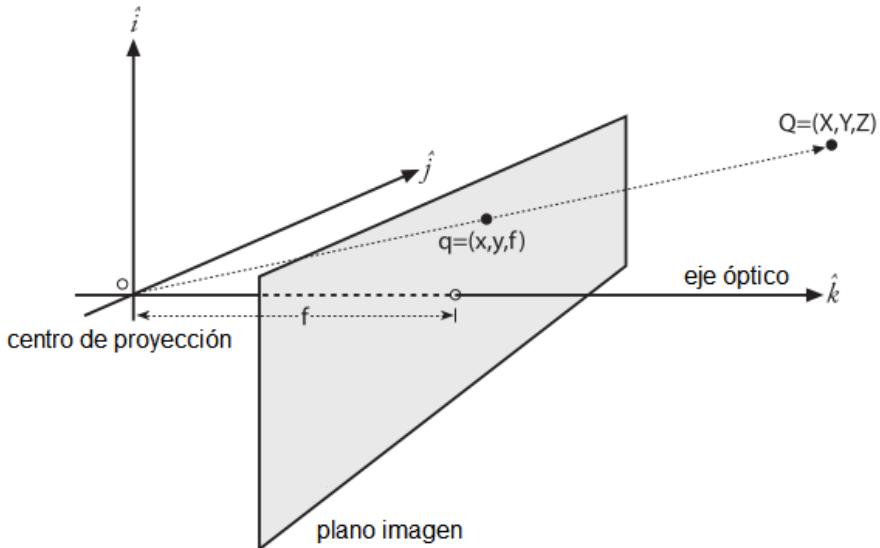


Figura 27: Modelo de proyección de la cámara

se podrá disponer de los parámetros de la proyección en un matriz M de 3×3 , que será denominada matriz intrínseca.

$$M = \begin{bmatrix} fx & s & cx \\ 0 & fy & cy \\ 0 & 0 & 1 \end{bmatrix}$$

donde fx y fy se pueden interpretar como la escala para los ejes x e y respectivamente, cx y cy es el punto principal respecto al ICS, y s describe el skewness en el sistema de coordenadas de la imagen. Si s es cero, los ejes x e y del sistema de coordenadas de la imagen son perpendiculares.

Como consideramos que el lente era simétrico alrededor del eje de proyección podríamos llegar a creer que $fx = fy$, sin embargo no siempre podemos usar esta simplificación ya que el array de sensores de la cámara puede tener diferente espaciado en cada dirección, o bien el elemento foto sensible puede ser rectangular en vez de cuadrado. Es por esto que lo dejaremos modelado por separado.

La proyección de los puntos del mundo físico (en coordenadas del CCS) hacia la cámara (en coordenadas del ICS) puede ahora escribirse en forma matricial:

$$[u, v, 1]^T = M[X, Y, Z]^T$$

donde $u = x/z$ y $v = y/z$ son las coordenadas para los ejes x e y en el ICS.

3.5 MODELO PINHOLE CON DISTORSIÓN

Como se dijo previamente, la única aberración que afecta los rayos centrales es la distorsión. Por lo tanto, para el modelo de cámara pinhole, con solo considerar los efectos de la distorsión podemos lograr una aproximación a una óptica de tercer orden. Con este fin se introduce un modelo para corregir la distorsión en el plano de la imagen $z = 1$. Este modelo es una función de dos variables, x e y .

3.5.1 Distorsión radial

La mayoría de los lentes utilizados comúnmente son invariantes a la rotación. Se puede decir que existe un eje de rotación de la lente, y vamos a aceptar la suposición de que éste eje de rotación corresponde al eje óptico de nuestro modelo de cámara. De esta forma podemos ver que la distorsión producida por la lente depende únicamente de la distancia al eje óptico, el radio r .

Para cada punto $q = (x, y)$ perteneciente a \mathbb{R}^2 se define su dirección n_q como

$$n_q = \frac{q}{\|q\|}$$

Una función de distorsión radial c es una función que mapea:

$$q = \|q\| n_q \mapsto c(\|q\|) n_q$$

Esta función depende únicamente de la distancia de q al origen, es decir, del radio $r = \|q\|$. Se puede modelar la distorsión con cualquier función analítica, pero por comodidad la vamos a formular como una serie de Taylor:

$$c(\|q\|) = c(r) = \sum_{i=0}^{\infty} c_i r^i$$

Como asumimos que la distorsión es nula en el origen, fijamos el coeficiente constante $c_0 = 0$. A su vez fijamos el coeficiente lineal $c_1 = 1$, ya que es simplemente una escala alrededor del origen, y no puede ser 0 ya que estaríamos excluyendo la distorsión nula como posible solución. Nuestra función de distorsión queda entonces definida como:

$$c(r) = r + \sum_{i=2}^{\infty} c_i r^i = r(1 + \sum_{i=1}^{\infty} c_{i+1} r^i)$$

De esta forma:

$$\begin{aligned} q_{\text{corregido}} &= c(\|q\|)n_q = \|q\|(1 + \sum_{i=1}^{\infty} c_{i+1}\|q\|^i)n_q \\ &= (1 + \sum_{i=1}^{\infty} c_{i+1}\|q\|^i)q \\ &= q + q \sum_{i=1}^{\infty} c_{i+1}\|q\|^i \end{aligned}$$

Finalmente, como el radio sólo puede ser positivo, podemos asumir que c puede ser aproximada por una función impar, lo que nos brinda la ventaja de que sólo debemos considerar los exponentes pares de la serie de potencias. Esta ventaja se hace evidente cuando debemos calcular el radio, ya que podemos obviar el cálculo de la raíz cuadrada para obtener $r = \|q\|$. De esta forma podemos escribir:

$$\begin{bmatrix} x_{\text{corregido}} \\ y_{\text{corregido}} \end{bmatrix} = \begin{bmatrix} x(1 + k_1r^2 + k_2r^4 + k_3r^6 + \dots + k_ir^{2i}) \\ y(1 + k_1r^2 + k_2r^4 + k_3r^6 + \dots + k_ir^{2i}) \end{bmatrix}$$

con los k_i como parámetros de distorsión radial.

3.5.2 Distorsión tangencial

La alineación de la lente en un sistema óptico también puede producir distorsiones. Estas distorsiones no pueden ser descritas como funciones que dependan solamente de la distancia al eje óptico.

En nuestro modelo de cámara asumimos que el plano de la imagen y el plano definido por el array de sensores eran paralelos. Por lo tanto cualquier desalineación entre ellos también debe ser modelado por la función de distorsión. Por ejemplo para lentes que están inclinados respecto al plano de la imagen (ver [Figura 28](#)) la distorsión radial se transforma en una distorsión elíptica.

Por otro lado, como se mencionó previamente, los sistemas de lentes reales generalmente son una combinación de dos o más lentes. Idealmente todas las lentes compartirán el mismo eje óptico. Sin embargo durante la construcción del sistema de lentes se pueden producir pequeñas diferencias, lo que también puede generar una aberración en la imagen observada.

Una característica de este tipo de aberraciones es que el eje óptico puede no ser el centro de la distorsión radial. Para modelar este tipo de casos se puede introducir una distorsión tangencial definida como [\[10\]](#):

$$\begin{bmatrix} x_{\text{corregido}} \\ y_{\text{corregido}} \end{bmatrix} = \begin{bmatrix} x + [2p_1y + p_2(r^2 + 2x^2)] \\ y + [p_1(r^2 + 2y^2) + 2p_2x] \end{bmatrix}$$

con p_1 y p_2 como parámetros de distorsión tangencial.

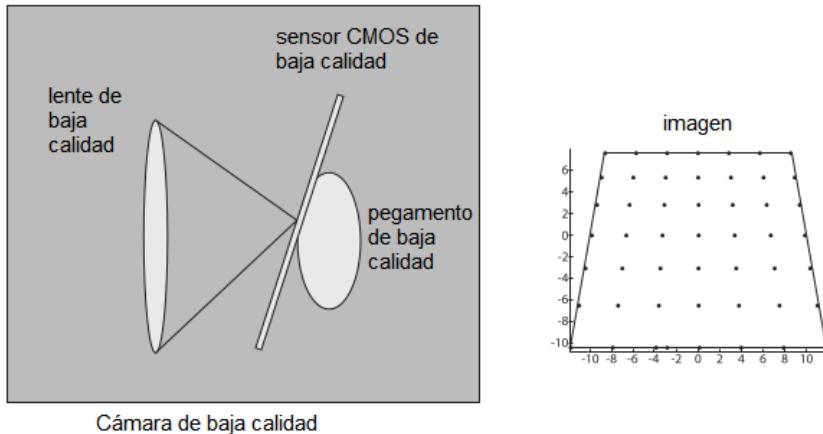


Figura 28: Posible causante de distorsión tangencial

Se pueden seguir añadiendo muchos otros modelos de distorsión, pero para nuestros fines la distorsión radial y tangencial resultan suficientes. Por otro lado se debe ser cuidadoso de que los parámetros que definen la distorsión completa sean independientes entre sí.

Como se comentó previamente, la transición de un sistema de coordenadas euclíadiano a un sistema de coordenadas proyectivo nos permite utilizar matrices para realizar los cambios de sistemas de coordenadas. En este momento es importante destacar que, si bien el modelo de cámara que utilizamos (excepto la distorsión) consiste simplemente en una transformación del sistema de coordenadas, el modelo de distorsión debe ser tratado de manera separada ya que no se puede describir como una matriz. Debemos recordar que el modelo de distorsión que se presentó está definido y sólo actúa sobre el plano $z = 1$ del sistema de coordenadas de la cámara.

3.6 PARÁMETROS EXTRÍNSECOS

Previamente cuando se introdujeron los parámetros intrínsecos de la cámara se asumió el hecho de que el centro óptico era el origen del sistema de coordenadas global. Sin embargo en diversas aplicaciones -como se verá más adelante en este proyecto- es necesario conocer la ubicación espacial de la cámara respecto a otro sistema de coordenadas.

Para poder modelar el sistema con independencia de su posición o para poder referenciar un objeto respecto a otro origen de coordenadas es necesario introducir un cambio de coordenadas. Este cambio de coordenadas equivale a una transformación rígida compuesta de una rotación R y una traslación t , como se observa en la [Figura 29](#). Gracias a que estamos trabajando con coordenadas homogéneas, po-

demos combinarlas en una sola matriz de 3×4 que denominaremos W :

$$W = [R \ t] \quad (1)$$

La transformación de proyección completa que pasa los puntos 3D a 2D queda definida entonces como:

$$q = M W Q$$

con Q y q en coordenadas homogéneas.

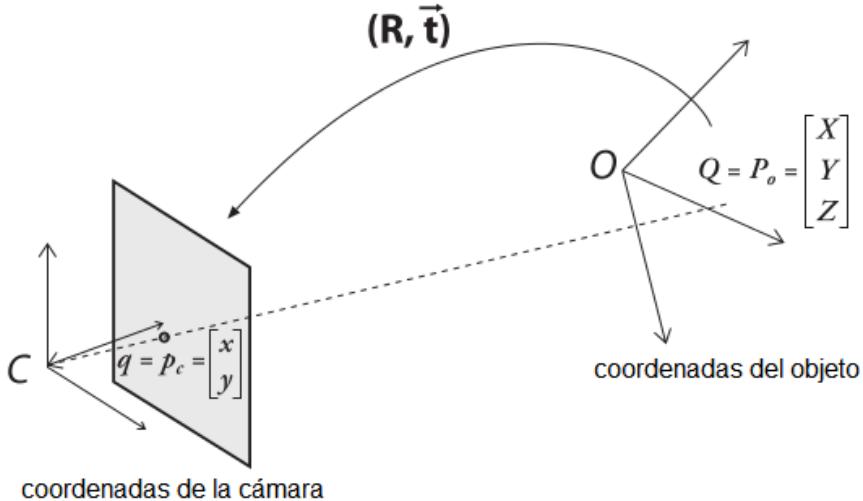


Figura 29: Parámetros extrínsecos

3.7 CALIBRACIÓN DE LA CÁMARA

El proceso de calibración de una cámara consiste en estimar sus parámetros intrínsecos y extrínsecos. Existen distintos tipos de métodos de identificación para este problema. Por un lado están los clásicos que consisten en situar en la escena una cantidad suficiente de puntos 3D cuyas posiciones respecto a un sistema de coordenadas sean conocidas, de modo de observarlos y medirlos en la imagen y así estimar la matriz de proyección del sistema. Por otro lado se encuentran los de autocalibración que, a diferencia de los otros, no necesitan una preparación exhaustiva de la escena ni un conocimiento previo de la misma ya que en su lugar se utiliza la correspondencia de puntos detectados en la escena a lo largo de una secuencia de imágenes.

Uno de los métodos más utilizados en los últimos años es el propuesto por Zhang [11] que, a pesar de no estar dentro del grupo de los de autocalibración, tampoco requiere un conocimiento de la escena en cuestión. Este método consiste en obtener diversas vistas de un patrón de calibración con puntos fácilmente identificables y cuyas coordenadas respecto a un sistema de coordenadas son conocidas.

El patrón de calibración en principio puede tener cualquier forma pero para simplificar la identificación automática de los puntos generalmente se utiliza un patrón plano similar a un tablero de ajedrez o una grilla de círculos.

Para obtener diversas vistas del patrón vamos a rotar y trasladar el patrón de calibración respecto a la ubicación de la cámara. Al observar la proyección de los puntos en la imagen y conociendo su distribución espacial se puede calcular la ubicación y orientación del patrón respecto a la cámara (o viceversa). Esto, sumado al hecho de que contamos con diversas vistas, nos permitirá obtener una estimación de los parámetros intrínsecos de la cámara.

Como vimos previamente, la matriz de proyección completa se definía como:

$$\mathbf{q} = \mathbf{M} \mathbf{W} \mathbf{Q}$$

La solución a este sistema de ecuaciones para puntos \mathbf{Q} y \mathbf{q} conocidos serán los parámetros de calibración que buscamos. Como vimos, la matriz extrínseca \mathbf{W} tiene 12 valores (3×4), pero los componentes independientes son sólo 3 para la rotación² y 3 para la traslación, con lo que nos quedan 6 incógnitas. La matriz intrínseca \mathbf{M} tiene 5 parámetros, por lo tanto este sistema de ecuaciones tiene 11 incógnitas. Sin embargo para cualquier cámara comercial podemos considerar que el sensor de la cámara está bien construido y los ejes x e y son perpendiculares, por lo que vamos a fijar el coeficiente $s = 0$. De esta forma el sistema de ecuaciones que debemos resolver para cada posición del patrón de calibración es de 10 incógnitas (pero para cada posición del patrón la matriz intrínseca es la misma). Como se verá a continuación, al utilizar un patrón plano el sistema a resolver se reduce a 8 incógnitas. De estas 8 incógnitas, 3 corresponden a la rotación y 3 a la traslación, por lo tanto es fácil demostrar que vamos a necesitar al menos dos vistas del patrón con ubicación y orientación diferentes (cada vista nos deja 2 ecuaciones independientes que podemos utilizar para resolver la matriz intrínseca, la cual tiene sólo 4 parámetros ya que fijamos $s = 0$).

3.7.1 Patrón de calibración

En principio cualquier objeto podría ser utilizado como patrón de calibración, pero lo más práctico es utilizar un patrón plano con marcas regulares, por ejemplo una grilla de cuadrados blancos y negros al estilo tablero de ajedrez, o una grilla de círculos negros sobre fondo

² La expresión matricial que utilizamos para la rotación es para hacer más fácil la notación y el cálculo, pero cualquier rotación puede describirse por una rotación de cierto ángulo alrededor de un eje. De esta forma con un vector de 3 componentes podemos indicar la dirección de rotación (para la dirección no influye la magnitud) y con su magnitud se indica el ángulo a rotar.

blanco, como puede observarse en la [Figura 30](#). Además de la reducción del tamaño del sistema de ecuaciones a resolver, la fabricación de un buen patrón de calibración plano resulta mucho más accesible que realizar un objeto tridimensional.

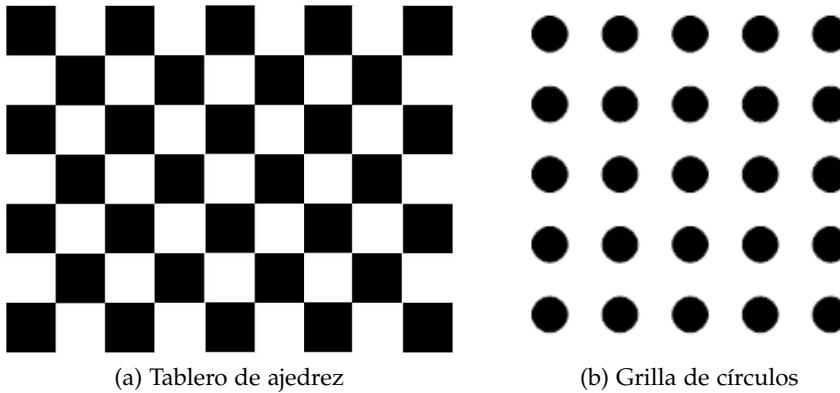


Figura 30: Ejemplos de patrones de calibración

El método de calibración se basa en el hecho de que los puntos en el plano presentarán una transformación perspectiva cuando sean vistos por la cámara. Esta transformación se conoce generalmente con el nombre de homografía plana. A continuación hablaremos acerca de ella.

3.7.2 Homografía

Una homografía plana es un mapeo proyectivo de un plano hacia otro plano. La transformación perspectiva que mencionamos previamente es un ejemplo de una homografía plana. Utilizando coordenadas homogéneas podemos describir esta transformación a través de una multiplicación por una matriz. Para un punto Q en el espacio y su respectivo punto q en la imagen, podemos expresar esta transformación H como:

$$q = s H Q$$

con

$$Q = [X, Y, Z, 1]^T \quad q = [x, y, 1]^T$$

El parámetro s se agrega solamente para dejar explícito el hecho de que la homografía se define sólo hasta un factor de escala. Más adelante veremos que este parámetro se elimina del sistema.

Debemos observar que H básicamente consta de dos partes: una transformación física y una proyección (que involucra la matriz intrínseca de la cámara). La transformación física es resultado de una rotación R y una traslación t que relaciona la ubicación del patrón de

calibración con el plano de la imagen, para lo cual podemos utilizar la misma matriz W que definimos en (1). Si agregamos el efecto de esta transformación a la ecuación de proyección de la cámara obtenemos:

$$q = s M W Q \quad (2)$$

Como se mencionó previamente, el uso de un patrón de calibración plano nos permite reducir el tamaño del sistema de ecuaciones a resolver. Esto se logra definiendo el sistema de coordenadas del patrón con el eje z perpendicular al plano, por lo que todos los puntos tendrán coordenada $z = 0$. De esta forma podemos reescribir la ecuación (2) como:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = s M [r_1 \ r_2 \ r_3 \ t] \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} = s M [r_1 \ r_2 \ t] \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}$$

donde r_1 , r_2 y r_3 son las tres columnas de la matriz de rotación R .

La homografía que transforma objetos sobre el plano de calibración hacia el plano de la imagen queda entonces definida como:

$$H = s M [r_1 \ r_2 \ t] \quad (3)$$

Se puede ver que H es una matriz de 3×3 .

El patrón de calibración plano debe tener al menos 4 puntos para describir completamente el mapeo de un rectángulo hacia un cuadrilátero. Si cada punto nos deja 2 ecuaciones, cada imagen nos brinda 8 ecuaciones. Esto significa que el sistema de ecuaciones puede resolverse si contamos con al menos 2 imágenes del patrón de calibración en posiciones distintas. Debemos recordar que para cada imagen tenemos 6 parámetros únicos que describen la rotación y traslación, sin embargo la matriz intrínseca de la cámara es la misma para todas las imágenes. De esta forma, si fijamos el coeficiente de skew (parámetro s de la matriz M) en cero, podemos resolver las 4 incógnitas de la matriz intrínseca de la cámara.

Si bien con 4 puntos sobre el plano de calibración nos alcanza, es importante destacar que en la práctica se utilizan muchos más para reducir el efecto del ruido y de la discretización implícita del sensor de la cámara, aún cuando se utilicen métodos de resolución subpixel.

3.7.3 Obtención de los parámetros de la matriz de la cámara

Si bien existen muchas maneras de resolver el sistema de ecuaciones y obtener los parámetros de la cámara, en este trabajo nos vamos a centrar en un algoritmo basado en el método de Zhang [10][11].

En principio, para obtener los parámetros de la matriz de la cámara, vamos a pretender que la distorsión es nula. Para cada imagen

de nuestro patrón plano de calibración se procede a obtener la homografía H , como se explicó previamente. Ahora vamos a escribir esta matriz H como 3 vectores columna:

$$H = [h_1 \ h_2 \ h_3]$$

Reemplazando en la ecuación (3) nos queda:

$$H = [h_1 \ h_2 \ h_3] = s M W = s M[r_1 \ r_2 \ t]$$

Despejando tenemos:

$$\begin{array}{lll} h_1 = s M r_1 & \text{o} & r_1 = \lambda M^{-1} h_1 \\ h_2 = s M r_2 & \text{o} & r_2 = \lambda M^{-1} h_2 \\ h_3 = s M t & \text{o} & t = \lambda M^{-1} h_3 \end{array} \quad (4)$$

Con $\lambda = 1/s$.

Los vectores de rotación r_1 y r_2 son ortogonales entre sí (por definición), y como la escala se incluye en el parámetro s , podemos decir que también son ortonormales. Esto implica que el producto punto entre r_1 y r_2 es nulo, y que su magnitud es igual. Expresando la ecuación para el producto punto:

$$r_1^T r_2 = 0$$

Por otro lado, para dos vectores a y b sabemos que $(ab)^T = b^T a^T$, por lo tanto si reemplazamos r_1 y r_2 obtenemos la primer restricción:

$$h_1^T M^{-T} M^{-1} h_2 = 0$$

donde usamos $M^{-T} = (M^{-1})^T$. Como dijimos, también sabemos que la magnitud de los vectores de rotación son iguales:

$$\|r_1\| = \|r_2\| \quad \text{o} \quad r_1^T r_1 = r_2^T r_2$$

Sustituyendo nuevamente r_1 y r_2 obtenemos la segunda restricción:

$$h_1^T M^{-T} M^{-1} h_1 = h_2^T M^{-T} M^{-1} h_2$$

Para simplificar los cálculos vamos a definir una nueva matriz B como:

$$B = M^{-T} M^{-1} = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{12} & B_{22} & B_{23} \\ B_{13} & B_{23} & B_{33} \end{bmatrix}$$

Esta matriz B tiene una solución general de la forma:

$$B = \begin{bmatrix} \frac{1}{f_x^2} & 0 & \frac{-c_x}{f_x^2} \\ 0 & \frac{1}{f_y^2} & \frac{-c_y}{f_y^2} \\ \frac{-c_x}{f_x^2} & \frac{-c_y}{f_y^2} & \frac{-c_x^2}{f_x^2} + \frac{-c_y^2}{f_y^2} + 1 \end{bmatrix}$$

Usando esta matriz B, las dos restricciones tienen la forma general $h_i^T B h_j$ en ellas. Como B es simétrica, podemos escribir esta multiplicación como un producto punto de vectores de 6 dimensiones. Si escribimos los elementos de B como un vector b tenemos:

$$h_i^T B h_j = v_{ij}^T b = \begin{bmatrix} h_{i1} & h_{j1} \\ h_{i1} & h_{j2} + h_{i2} h_{j1} \\ & h_{i2} h_{j2} \\ h_{i3} & h_{j1} + h_{i1} h_{j3} \\ h_{i3} & h_{j2} + h_{i2} h_{j3} \\ & h_{i3} h_{j3} \end{bmatrix}^T \begin{bmatrix} B_{11} \\ B_{12} \\ B_{22} \\ B_{13} \\ B_{23} \\ B_{33} \end{bmatrix}^T$$

Usando esta definición de v_{ij}^T las dos restricciones pueden ser escritas como:

$$\begin{bmatrix} v_{12}^T \\ (v_{11} - v_{22})^T \end{bmatrix} b = 0$$

Si tenemos K imágenes del patrón de calibración, podemos apilar K ecuaciones en un sistema:

$$V b = 0$$

donde V es una matriz de $2K \times 6$. Si $K \geq 2$ esta ecuación puede ser resuelta, y los parámetros intrínsecos de la cámara pueden ser obtenidos a partir de la solución general de la matriz B:

$$\begin{aligned} f_x &= \sqrt{\frac{\lambda}{B_{11}}} \\ f_y &= \sqrt{\frac{\lambda B_{11}}{B_{11} B_{22} - B_{12}^2}} \\ c_x &= \frac{-B_{13} f_x^2}{\lambda} \\ c_y &= \frac{B_{12} B_{13} - B_{11} B_{23}}{B_{11} B_{22} - B_{12}^2} \end{aligned}$$

donde:

$$\lambda = B_{33} - \frac{B_{13}^2 + c_y (B_{12} B_{13} - B_{11} B_{23})}{B_{11}}$$

3.7.4 Obtención de los parámetros extrínsecos

Una vez que conocemos la matriz de la cámara M , los parámetros extrínsecos pueden ser obtenidos a partir de la ecuación de la homografía (4):

$$r_1 = \lambda M^{-1} h_1$$

$$r_2 = \lambda M^{-1} h_2$$

$$r_3 = r_1 \times r_2$$

$$t = \lambda M^{-1} h_3$$

donde el parámetro de escala se determina a partir de la condición de ortonormalidad:

$$\lambda = 1/\|M^{-1} h_1\|$$

Sin embargo cuando resolvemos con datos reales y armamos la matriz de rotación $R = [r_1 \ r_2 \ r_3]$, es muy probable que no obtengamos una matriz de rotación correcta que satisfaga la condición $R^T R = R R^T = I$ [10]. Para resolver este problema se utiliza la decomposición en valores singulares (SVD) de R . SVD es un método de factorización de una matriz en dos matrices ortonormales U y V , y una matriz D con valores de escala en su diagonal. Esto nos permite escribir $R = U D V^T$. Como R debe ser ortonormal, la matriz D debe ser la identidad I , de manera que $R = U I V^T$. De esta forma podemos transformar nuestra matriz R en una matriz de rotación correcta obteniendo su descomposición en valores singulares, forzando D a ser la matriz identidad y volviendo a calcular $R = U D V^T$.

3.7.5 Obtención de los parámetros de distorsión

En el modelo de distorsión generalmente se utilizan entre 3 y 6 parámetros para la distorsión radial y los 2 parámetros de la distorsión tangencial.

Para estimar estos parámetros utilizamos la matriz de la cámara obtenida previamente. Debido a la distorsión, los puntos del patrón de calibración se observan en la imagen en lugares incorrectos. Utilizando la rotación y traslación obtenidos para cada imagen, podemos realizar la reproyección de los puntos para obtener donde deberían encontrarse en la imagen. La diferencia entre los puntos reales reproyectados sobre la imagen y la ubicación donde se observan define un error conocido como error de reproyección.

Los parámetros de distorsión que buscamos son aquellos que minimizan este error. Es por ello que podemos obtenerlos mediante cualquier algoritmo de minimización no lineal, utilizando el error de reproyección como función de error a minimizar. El valor inicial de los parámetros de distorsión es cero, a partir de los cuales se itera hasta llegar al criterio de finalización que se establezca.

3.8 CALIBRACIÓN MULTI-CÁMARA

La calibración multi-cámara consiste en encontrar la rotación R y translación t entre uno o más pares de cámaras, como se observa en la [Figura 31](#). Estos parámetros se pueden obtener de manera similar al método utilizado para obtener la calibración intrínseca de la cámara, con la diferencia de que ahora necesitamos imágenes de un mismo patrón de calibración observado desde cada cámara. Como vimos en la sección anterior, para cada imagen del patrón de calibración se podía obtener una rotación y translación. Podemos realizar lo mismo utilizando los modelos de las cámara ya calibradas, con lo cual obtendremos una rotación y una translación para cada cámara por cada ubicación del patrón de calibración. Lo que estamos buscando en realidad es una única rotación y translación entre las cámaras, por lo que deberemos realizar unos pasos extras.

Se puede decir que para cada punto P en coordenadas *del mundo* podemos usar la calibración de la cámara para ubicar a P en coordenadas de la cámara: $P_l = R_l P + t_l$ para una cámara y $P_r = R_r P + t_r$ para la otra. Se puede ver en la [Figura 31](#) que las dos vistas de P están relacionadas por $P_l = R^T (P_r - t)$ donde R y t son la matriz de rotación y el vector de translación entre ambas cámaras. Utilizando estas tres ecuaciones y resolviendo para R y t se obtiene:

$$R = R_r (R_l)^T$$

$$t = t_r - R t_l$$

Utilizando las imágenes del patrón de calibración en diferentes posiciones (con una imagen desde cada cámara) se resuelven las ecuaciones vistas con anterioridad (ver [Sección 3.7.4](#)) para obtener los parámetros de rotación y translación para cada cámara por separado. A partir de estas rotaciones y translaciones se resuelven las ecuaciones recién vistas para obtener la rotación y translación entre las cámaras. Debido a diversos factores como el ruido y errores de redondeo, estas rotaciones y translaciones no son iguales para cada ubicación del patrón de calibración. Para obtener la solución óptima se utiliza la media de estos valores como aproximación inicial y se utiliza el algoritmo iterativo de Levenberg-Marquardt para minimizar el error de reprojeción. De esta forma se obtiene la rotación R y la translación t entre las cámaras, las cuales nos servirán para establecer un único sistema de coordenadas de referencia.

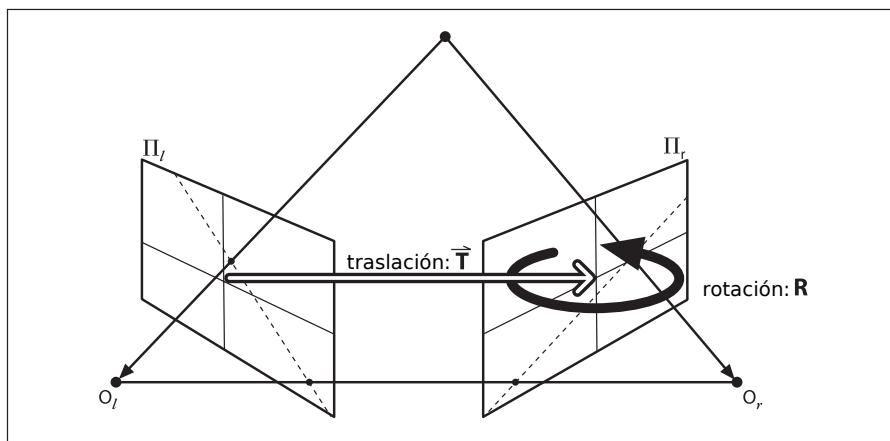


Figura 31: Rotación y traslación entre dos cámaras

4

DESARROLLO

En la primer sección de este capítulo se brindarán detalles del prototipo de dispositivo de medición mediante luz estructurada, junto con las características de sus principales componentes. En la segunda sección se hablará superficialmente de las principales características del software involucrado. En la tercera sección se justificará la elección de los patrones de luz estructurada que se utilizarán y se explicará de manera simple la forma en que se usan, siendo el punto más importante su decodificación de manera robusta. En la cuarta sección se tratará la obtención de la nube de puntos en tres dimensiones a partir del uso de las imágenes de los patrones decodificados. La sección siguiente se dedicará específicamente al proceso de calibración y las herramientas desarrolladas para tal fin. En la última sección se analizarán las más importantes limitaciones del dispositivo.

4.1 PROTOTIPO

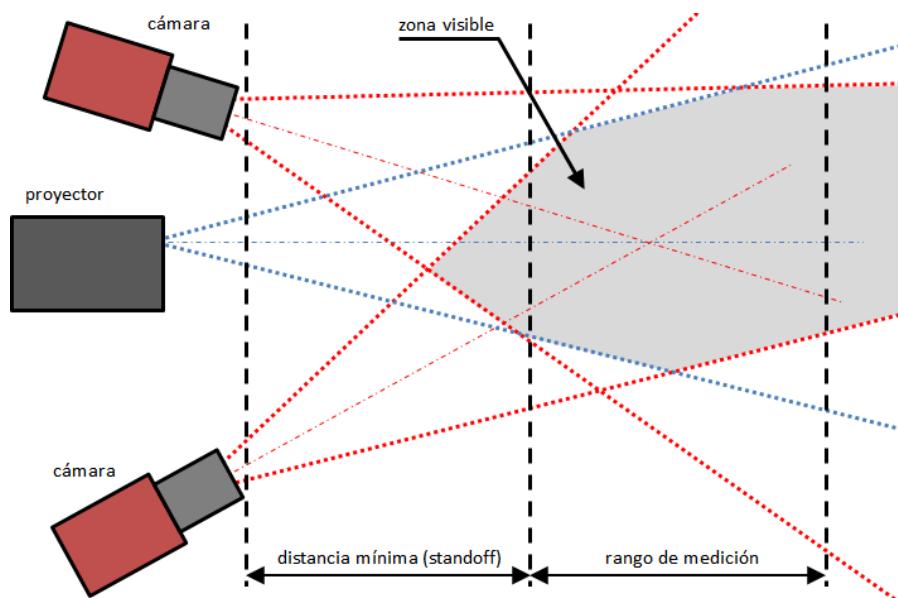


Figura 32: Esquema del prototipo

El prototipo desarrollado cuenta con dos cámaras y un proyector, dispuestos como se indica en la [Figura 32](#). Se puede observar que existe una zona con límites bien definidos, que denominaremos zona visible, que corresponde a la intersección del campo de visión de las cámaras junto con el campo de proyección del proyector. También se observa una distancia mínima y un rango de medición. Estos límites son difusos (a diferencia de la zona visible) y se establecen debido a que las lentes (tanto de las cámaras como del proyector) tienen una profundidad de foco acotada. La calibración se ajusta para realizar mediciones dentro de este rango. Por más que sea posible obtener mediciones en cualquier parte de la zona visible, no se puede brindar ninguna garantía respecto a la cota de error en las mediciones obtenidas fuera del rango de medición.

El esquema propuesto es el utilizado en sensores comerciales como los que ofrece Geomagic (ahora parte de 3D Systems) con su producto Geomagic Capture¹, LMI3D con su producto Gocator Snapshot Sensor², GOM a través de sus diversos escáneres 3D³, entre otros (ver [Figura 33](#)). La mayoría de estos productos vienen calibrados de fábrica y no permiten al usuario realizar modificaciones para adaptarlos a requerimientos particulares. El prototipo armado está compuesto por piezas de laboratorio que permiten gran flexibilidad para realizar ajustes y modificaciones, y se desarrolló un procedimiento relativamente sencillo para facilitar su calibración.

4.1.1 Componentes

Los principales componentes del prototipo son:

- cámara Allied Vision Technologies Prosilica GE1910⁴
- cámara Allied Vision Technologies Prosilica GC2450⁵
- 2 lentes de 16 mm de distancia focal
- proyector Texas Instruments DLP Pico Projector Development Kit⁶

Se elige la utilización de cámaras en escala de grises (no color) ya que estas son más eficientes respecto a la adquisición de la luz

¹ <http://geomagic.com/en/products/capture/overview/>

² <http://lmi3d.com/products/gocator/snapshot-sensor/>

³ <http://www.gom.com/metrology-systems/3d-scanner.html>

⁴ <http://www.alliedvisiontec.com/us/products/cameras/gigabit-ethernet/prosilica-ge/ge1910.html>

⁵ <http://www.alliedvisiontec.com/us/products/cameras/gigabit-ethernet/prosilica-gc/gc2450.html>

⁶ <http://www.ti.com/tool/dlp1picokit>



Figura 33: Dispositivos comerciales

y tienen menor ruido⁷. Las cámaras brindan la posibilidad de ser controladas por una señal de *trigger* (por hardware), sin embargo para las pruebas no se las utiliza de esta manera sino que son controladas y sincronizadas completamente por software.

El proyector es de bajo consumo y reducidas dimensiones, pero también de baja resolución: solamente 480x320 pixeles. La ventaja de este proyector está en la utilización de tecnología DLP de Texas Instruments. Se trata básicamente de un proyector que genera la imagen a partir de una fuente de luz LED y una grilla de espejos en miniatura. El diagrama del funcionamiento de un espejo puede verse en la Figura 34. Este tipo de tecnología posibilita generar patrones binarios con frecuencias mayores a 2KHz. Si bien en el prototipo no estamos utilizando esta ventaja (ya que se lo utiliza de igual manera que un proyector común de computadora), se debe tener en cuenta que el dispositivo brinda esta posibilidad, pudiendo ser aprovechada a futu-

⁷ Esto se debe en gran medida a que las cámaras monocromáticas no incorporan lo que se conoce como filtro de bayer (ver http://en.wikipedia.org/wiki/Bayer_filter), el cual es necesario para obtener imágenes color. Existen otros métodos para obtener imágenes color pero éste es el más utilizado.

ro mediante el uso de cámaras de alta velocidad y sincronización por hardware.

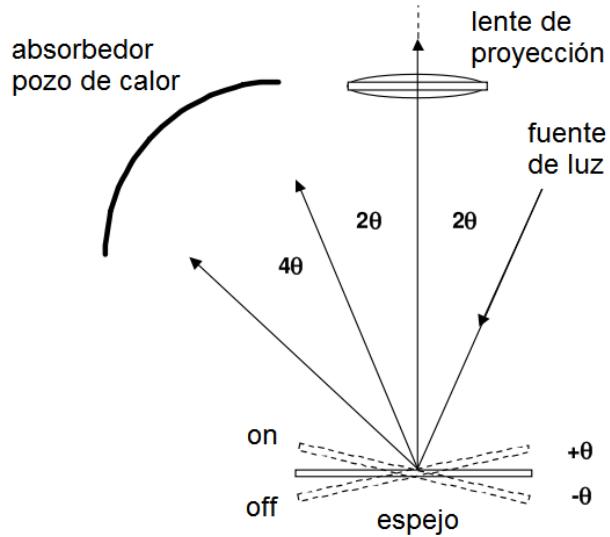


Figura 34: Diagrama de funcionamiento de la tecnología DLP

4.2 SOFTWARE

El software se implementó en lenguaje C++ utilizando el framework open-source multiplataforma Qt⁸ y la biblioteca OpenCV⁹. El principal objetivo del software es facilitar el uso del dispositivo (fundamental en el caso de un usuario inexperto) y ayudar al desarrollador a implementar nuevos tipos de patrones de luz estructurada y visualizar resultados para depurar los algoritmos. Para lograr una buena interacción con el usuario se brinda una interfaz de usuario simple que permite el acceso a la mayoría de las opciones, pero evitando sobrecargarla demasiado con todos los parámetros internos. En el caso del desarrollador (tanto actual como los posibles futuros interesados en contribuir), se diseñó una estructura completamente modular, con diferentes bibliotecas dedicadas exclusivamente a las principales funcionalidades: adquisición de imágenes, calibración de la cámara, sistemas de luz estructurada, visualización de nubes de puntos en 3D, etc.

Una de las características más destacadas es la implementación de una biblioteca genérica para la adquisición de imágenes de distintas marcas y modelos de cámaras, la cual utiliza plugins para incorporar la funcionalidad de cada Software Development Kit ([SDK](#)) utilizado. El uso de plugins permite incorporar fácilmente nuevos modelos y marcas de cámaras o [SDK](#) mediante el uso de una Application Programming Interface ([API](#)) bien definida y con la ventaja de que no

⁸ <http://qt-project.org/>

⁹ <http://opencv.org/>

requiere la recompilación del software que los utiliza. Por otro lado también permite reducir considerablemente las dependencias en tiempo de ejecución, favoreciendo la utilización (o no) de ciertos plugins en base a los requerimientos de las cámaras utilizadas. Actualmente se pueden utilizar las cámaras de Point Grey soportadas por su FlyCapture SDK¹⁰, las cámaras de Lumenera soportadas por su LuCam SDK¹¹, las cámaras de Allied Vision Technologies soportadas por su Vimba SDK¹², todas las cámaras GigE, 10 GigE, USB 2.0, USB 3.0 y WIFI soportadas por el eBUS SDK de Pleora¹³ y todas las cámaras soportadas por NI-IMAQdx de National Instruments¹⁴.

En la Figura 35 se puede observar la ventana principal de la aplicación y la configuración de una cámara. El menú superior de la ventana principal contiene menús para el acceso a la configuración, vista previa y calibración de las cámaras; como así también acceso a la ventana para visualizar en tres dimensiones la ubicación física de las cámaras ó el resultado de la medición en el sistema de coordenadas global.

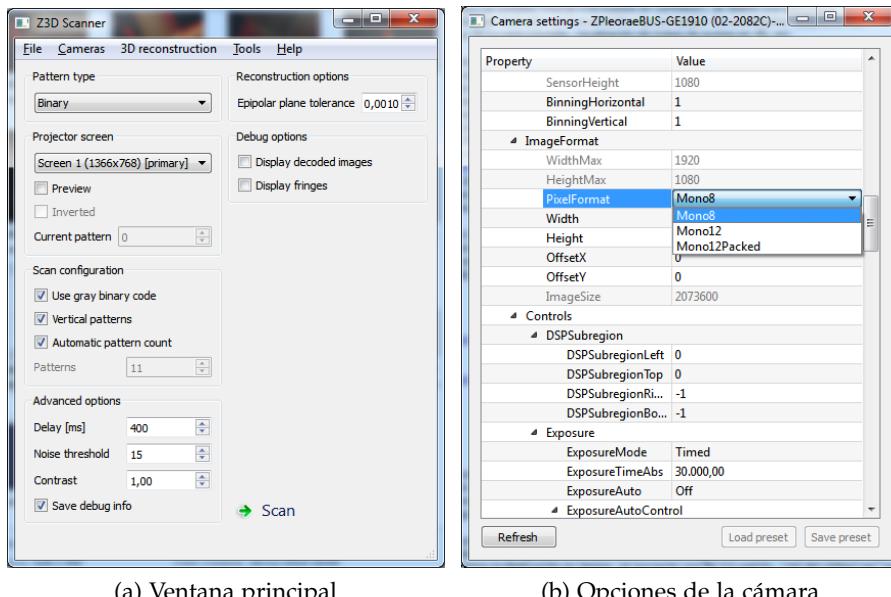


Figura 35: Software de medición

4.3 PROYECCIÓN Y DECODIFICACIÓN DE LOS PATRONES

Para la codificación y decodificación se parte de la suposición de que la escena es estática (o se puede considerar estática durante el período

¹⁰ <http://www.ptgrey.com/flycapture-sdk>

¹¹ <http://www.lumenera.com/support/downloads/industrial-downloads.php>

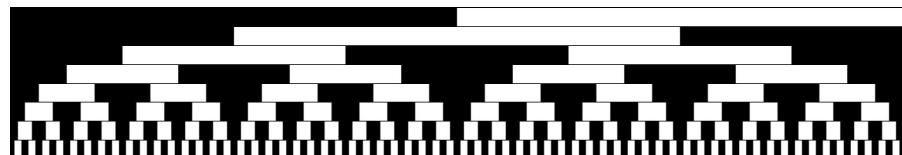
¹² <http://www.alliedvisiontec.com/us/products/software/vimba-sdk.html>

¹³ <http://www.pleora.com/our-products/ebus-sdk>

¹⁴ <http://sine.ni.com/nips/cds/view/p/lang/es/nid/207702>

do de escaneado), ya que utilizamos una codificación temporal. Se proyectan múltiples patrones y la secuencia temporal de valores para cada pixel se utiliza como código.

La codificación que utilizaremos se basa en patrones binarios debido a que es más robusta al ruido. En la parte superior de la [Figura 36](#) se puede observar un ejemplo de la codificación en binario Gray (se muestran como blanco y negro los dos valores posibles, 0 y 1). Como realizamos multiplexación en tiempo, se proyecta una fila (un patrón, 1 bit del código) por vez. Utilizamos patrones con franjas verticales, por lo que el valor de una fila se repite para toda la columna, como se puede ver en la parte inferior de la [Figura 36](#). Se utiliza código Gray por sus ventajas ya mencionadas en la [Sección 2.5.1.2](#).



(a) Código gray de 8 bits. Cada fila representa un bit (un patrón de luz estructurada) y cada columna es un código único



(b) Primeros 4 patrones proyectados, correspondientes a las primeras 4 filas

Figura 36: Patrones de luz estructurada usando código gray

Para facilitar la detección de las zonas donde el patrón está presente y es visible por la cámara, y para determinar un umbral mínimo de la *calidad* del pixel, se utiliza un patrón extra que es completamente blanco, y su inverso completamente negro. Este patrón se utiliza como *máscara* para procesar solamente los pixels que contienen información útil.

Por otro lado, para hacer aún más robusta la decodificación y poder discernir claramente cual es el valor binario correspondiente a cada pixel, se utiliza el patrón en conjunto con su patrón inverso, como se propone en [\[12\]](#). De esta forma la clasificación del valor binario correspondiente se lleva a cabo mediante la comparación de si el valor del pixel es mayor para la imagen del patrón o para la imagen del patrón inverso. Un ejemplo de la proyección de un patrón y el patrón inverso correspondiente puede observarse en la [Figura 37](#).

Para la obtención de las imágenes de los patrones se considera que las cámaras y el proyector están bien configurados (todos los parámetros, incluyendo tanto la configuración del sensor de la cámara como los ajustes manuales de apertura y foco del lente). No entraremos en detalle en estas cuestiones pero la calidad de las imágenes (enfoque,

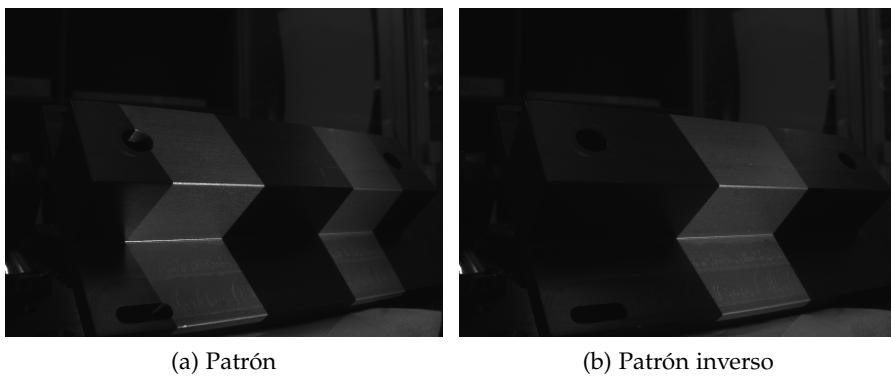


Figura 37: Ejemplo de la proyección de un patrón y el patron inverso correspondiente

nivel de ruido, contraste, etc) juega un papel fundamental en la calidad de los resultados. Un ejemplo de un patrón observado desde ambas cámaras puede verse en la [Figura 38](#). En la [Figura 39](#) se puede observar el resultado de la decodificación para estas imágenes de ejemplo.

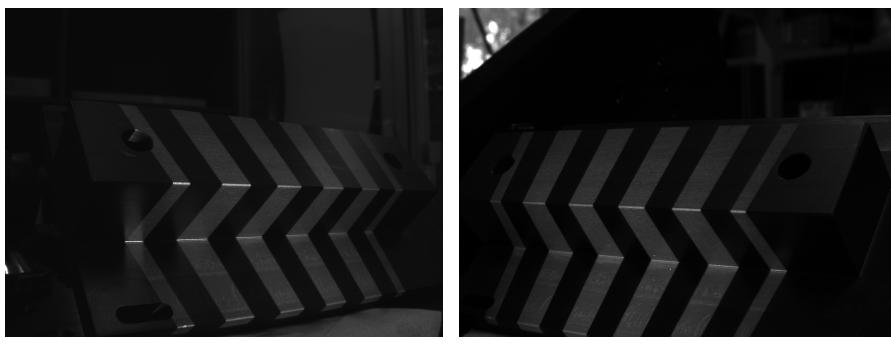


Figura 38: Ejemplo de un patrón proyectado

Al usar patrones binarios, la única información útil es la representada por el límite entre una franja y la siguiente. Es decir, el valor del código de cualquier pixel dentro de una franja de pixeles con el mismo código no nos brinda información relevante¹⁵. Es por esto que lo primero que necesitamos es obtener la ubicación de los *bordes* entre las franjas. Los bordes que más nos interesan son aquellos que se encuentran entre dos franjas consecutivas (con código sucesivo). En la [Figura 40](#) se muestra el resultado de este primer procesamiento de las imágenes decodificadas, donde se eliminan todos los pixels interiores a cada franja y sólo se mantienen los bordes.

¹⁵ Nos sirve para conocer que ese pixel no es un hueco, que corresponde a una parte de la pieza, pero no contiene información relevante para conocer con exactitud su ubicación en el espacio

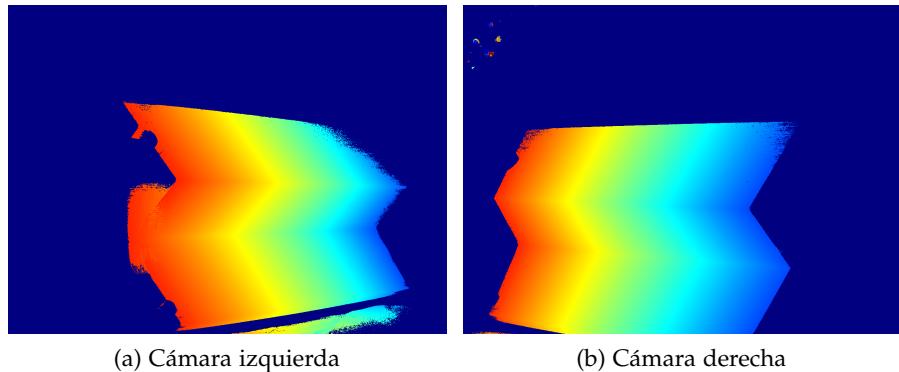


Figura 39: Resultado de la decodificación de las imágenes. El color representa el código: en este caso el código es el valor de la columna del proyector

A partir de este procesamiento se generan listas de pixeles correspondientes a los límites entre franjas sucesivas. Por ejemplo se genera una lista de los pixels correspondientes al límite entre las franjas con código 1 y 2, otra lista para los pixels correspondientes al límite entre las franjas con código 2 y 3, y así sucesivamente. Esto se realiza para las imágenes decodificadas de cada cámara utilizada. Cada lista tiene un identificador para saber a qué borde pertenece, para mas adelante poder establecer las correspondencias entre las distintas cámaras (o respecto a la franja del proyector, en el caso que solamente contemos con una cámara).

En este punto podemos establecer una analogía con un escáner de línea: se puede decir que cada lista de pixels es el equivalente a una línea decodificada en un escáner de línea. De esta manera se hace evidente una de las ventajas de los métodos de luz estructuradas: con el uso de patrones binarios de 9 bits (en nuestro caso, $9 \times 2 + 2 = 20$ imágenes) se puede obtener el equivalente a $2^9 - 1 = 511$ (descartando los límites exteriores de las franjas de los extremos) o $2^9 + 1 = 513$ (utilizando los límites exteriores de las franjas de los extremos) perfiles de un escáner de línea (donde se necesita una imagen para cada perfil).

En el [Apéndice B](#) se muestra una medición de ejemplo, y se incluyen todas las imágenes que fueron utilizadas para obtenerla.

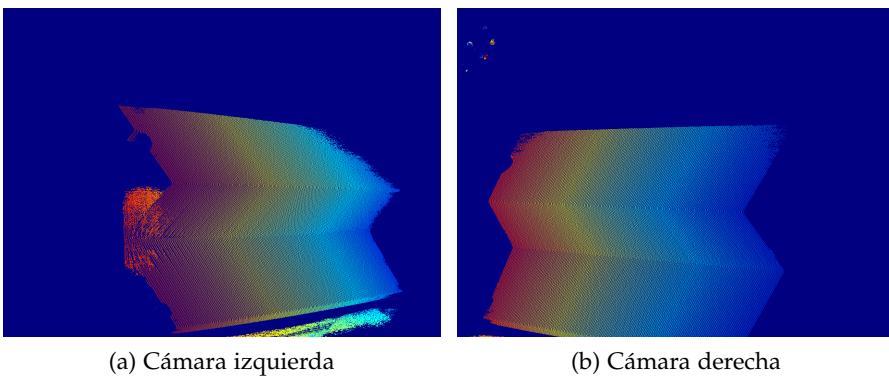


Figura 40: Bordes entre franjas decodificadas

4.4 RECONSTRUCCIÓN

Para la obtención de la nube de puntos en tres dimensiones se deben utilizar los conceptos vistos sobre la calibración de la cámara y el principio de triangulación. En principio se asume que tenemos una buena aproximación de los parámetros del modelo de cada cámara y de su ubicación espacial respecto a un único sistema de coordenadas. Respecto al principio de triangulación utilizaremos exclusivamente la intersección entre líneas.

En esta etapa ya no utilizaremos imágenes sino el resultado de su procesamiento, es decir que partimos desde los conjuntos de coordenadas 2D de los puntos correspondientes a cada límite entre franjas. Para entender mejor el procedimiento vamos a centrarnos en el procesamiento de un solo límite entre dos franjas, es decir trataremos un sólo conjunto de puntos por cada cámara. Conocemos que los puntos de cada conjunto corresponden a una misma línea sobre la superficie del objeto observado¹⁶, con lo cual mediante triangulación se puede obtener su ubicación en tres dimensiones. En este momento se puede establecer una clara relación con los conceptos involucrados en la visión estéreo, con la ventaja de que en este caso no hace falta buscar correspondencias entre las imágenes debido a que ya las conocemos por el uso de patrones de luz estructurada con codificación única.

Al inicio de esta sección dijimos que íbamos a utilizar intersección entre líneas. Cada punto en la imagen (2D) puede ser transformado a una recta en el espacio (3D) mediante los conceptos ya vistos en el [Capítulo 3](#). La intersección se debe realizar entre puntos de diferentes cámaras, pero todavía no establecimos la correspondencia entre los distintos puntos pertenecientes a las distintas cámaras, es decir, no conocemos qué punto del conjunto de la otra cámara debemos intersectar con cada punto de la primer cámara. En principio hubie-

¹⁶ Se podría decir que corresponden a un mismo plano de luz, ya que a fines prácticos podemos considerar que el proyector es perfecto. La probable deformación generada por la óptica del proyector no nos afecta.

ramos podido reducir la complejidad de este problema agregando patrones en la dirección opuesta (en nuestro caso significaría agregar patrones con franjas horizontales), lo que nos brinda información adicional para establecer las correspondencias entre los puntos de cada conjunto. Sin embargo esto significaría la duplicación de la cantidad de patrones necesarios (y el procesamiento del doble de imágenes), con su correspondiente incremento del tiempo de medición. Una solución práctica a este problema surge de la utilización de la geometría epipolar. Como se dijo previamente en la [Sección 3.1](#), el uso de un modelo de cámara central nos permite formular el mapeo como una proyección, lo que nos posibilita utilizar todos los conceptos de la geometría proyectiva [8], incluyendo la geometría epipolar. Este concepto es muy utilizado en visión estéreo para minimizar la búsqueda de correspondencias entre imágenes [10], y en nuestro caso nos ayuda de igual manera a reducir considerablemente el procesamiento.

La geometría epipolar nos permite realizar una proyección de las imágenes de un par de cámaras hacia un único plano de imagen virtual, el cual está a igual distancia de ambas cámaras y cuenta con el eje x alineado al vector translación que separa las dos cámaras. A partir de la proyección de los puntos de cada conjunto hacia este plano único podemos establecer fácilmente la relación entre ellos (solamente comparamos sus coordenadas y en el plano de imagen virtual) y también nos permite simplificar el cálculo de la triangulación. La intersección aproximada entre los pares de rectas (las proyecciones correspondientes a los puntos) se reduce a un cálculo muy rápido de la distancia z utilizando solamente la separación en la dirección x del plano de imagen virtual. Este concepto es muy utilizado en el ámbito de visión estéreo y se conoce como *disparidad* (ver [Figura 11](#)).

Una vez que calculamos todas las intersecciones obtenemos lo que se denomina comúnmente como *nube de puntos*, que no es más que un conjunto de puntos en el espacio. En este momento se puede utilizar la imagen obtenida con el patrón completamente blanco para asignarle un valor de intensidad (o color en el caso de imágenes color) a cada punto.

Esta nube de puntos se puede utilizar como punto de partida para una gran cantidad de aplicaciones (como las mencionadas en el [Capítulo 1](#)), y existe una gran cantidad de algoritmos y software dedicados a su procesamiento y visualización, incluyendo proyectos open-source como MeshLab¹⁷ o CloudCompare¹⁸, además de soluciones comerciales muy completas como SolidWorks¹⁹, Geomagic Design²⁰, SpaceClaim²¹, entre muchos otros. Algunos ejemplos de las tareas que se pueden realizar son la aplicación de filtros, la reducción de

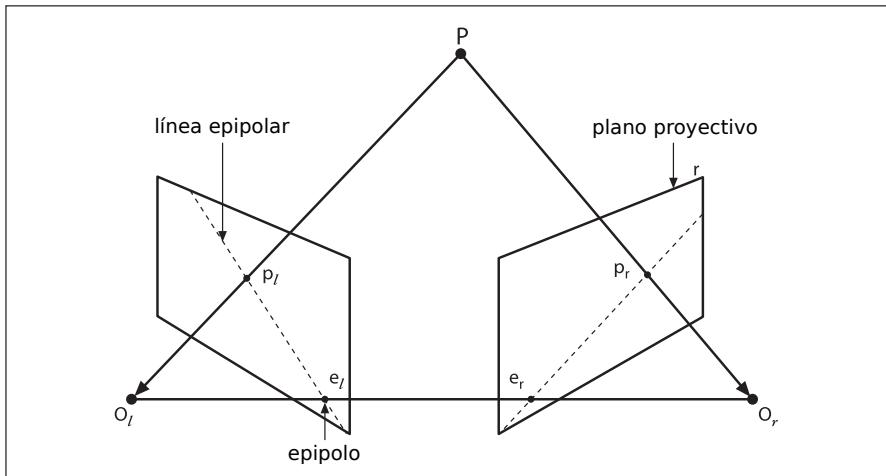
¹⁷ <http://meshlab.sourceforge.net/>

¹⁸ <http://www.danielgm.net/cc/>

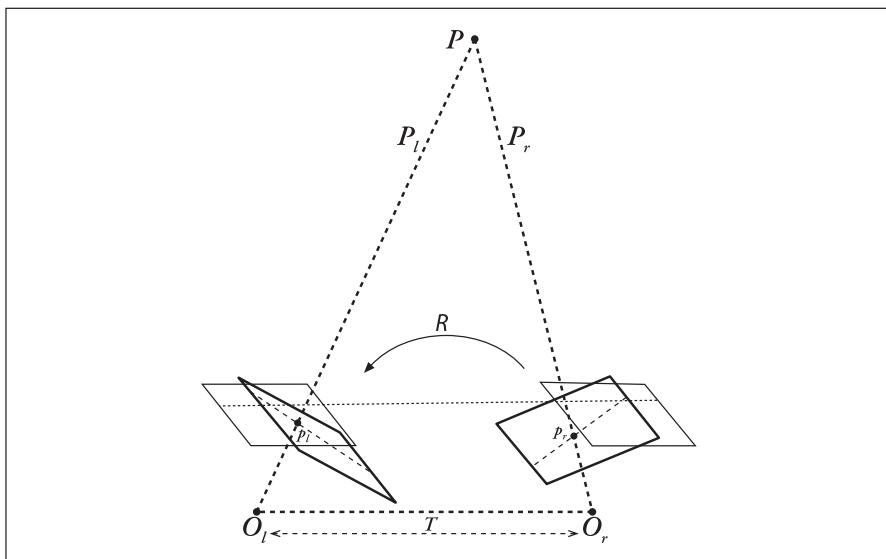
¹⁹ <http://www.solidworks.com/sw/products/3d-cad/scanto3d.htm>

²⁰ <http://geomagic.com/en/products-landing-pages/3d-design>

²¹ <http://www.spaceclaim.com>



(a) Principales conceptos de la geometría epipolar



(b) Proyección sobre un único plano virtual

Figura 41: Geometría epipolar

la cantidad de puntos a partir de un submuestreo, la generación de superficies, la segmentación de objetos y aproximación a formas primivas, la integración de nubes de puntos tomadas desde diferentes puntos de vista (o diversos sensores) para obtener un objeto completo, etc. En nuestro caso lo único que nos interesa es realizar mediciones sobre formas que ya conocemos por lo que no entraremos en detalle en estos aspectos.

En la [Figura 42](#) se observa la nube de puntos obtenida para las imágenes del ejemplo. El software de visualización desarrollado es básico pero es muy útil para observar rápidamente el resultado del proceso y aplicar algunos algoritmos básicos de post-procesamiento (por ejemplo para la eliminación de valores espúreos y para realizar un suavizado de la nube de puntos). Se desarrolló con la ayuda de

la biblioteca Point Cloud Library (PCL)²² y del Visualization Toolkit (VTK)²³.

²² <http://pointclouds.org/>

²³ <http://www.vtk.org/>

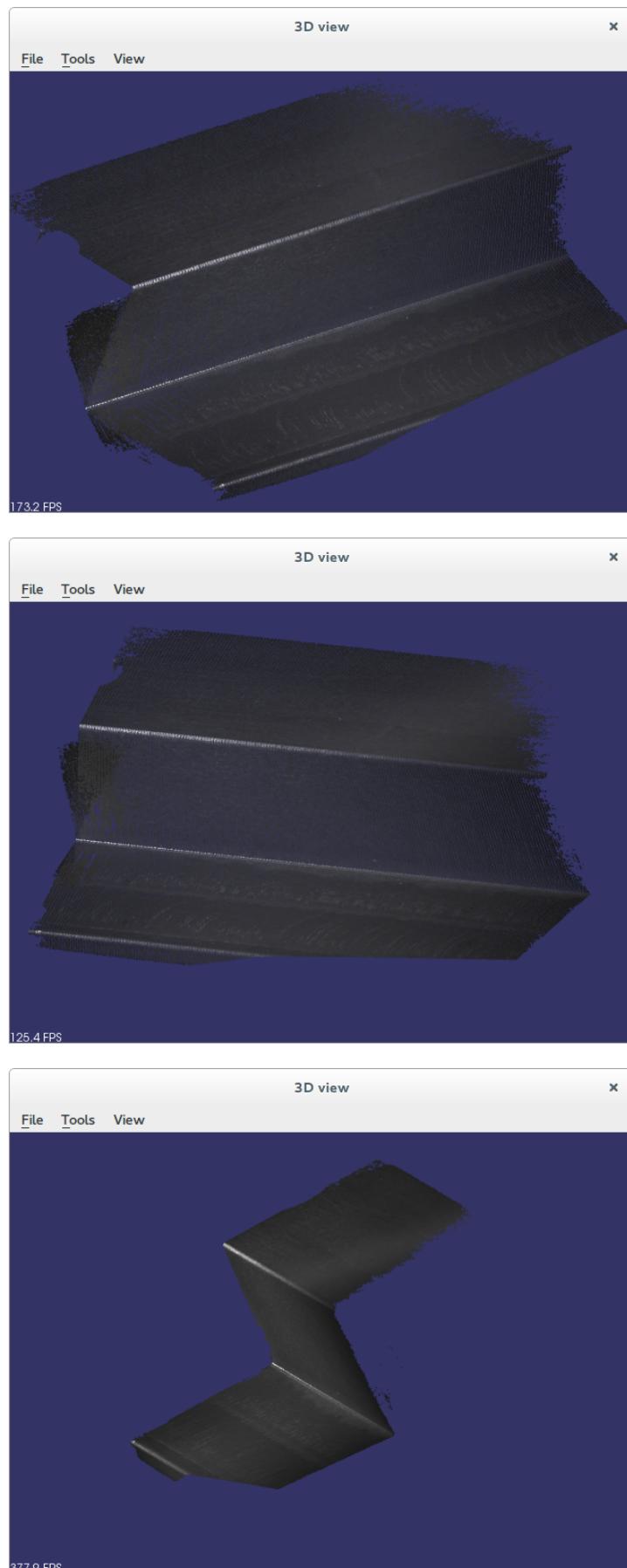


Figura 42: Visualización de la nube de puntos 3D

4.5 CALIBRACIÓN DEL DISPOSITIVO

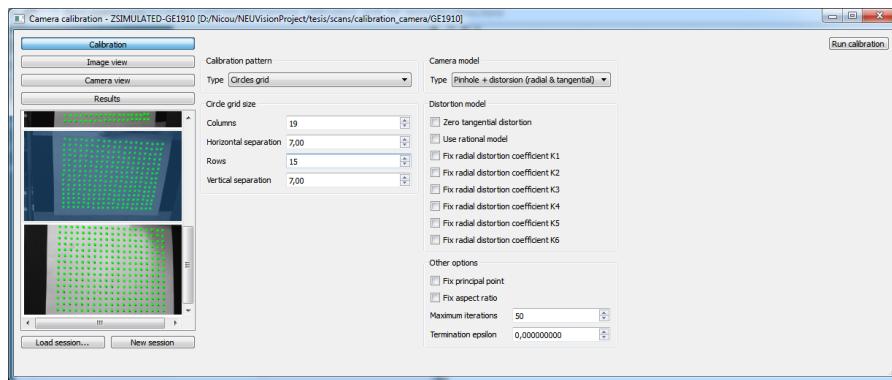
Junto con el software desarrollado se incluye una interfaz de usuario para facilitar la calibración de las cámaras y del dispositivo completo. En la [Figura 43](#) se observan las distintas secciones de la interfaz de calibración de los parámetros intrínsecos de las cámaras. Se puede ver que se incluye una sección que permite la selección del patrón de calibración a utilizar y el ajuste de los diversos parámetros del método de calibración. También cuenta con una vista previa de la cámara para tomar las fotos necesarias, permitiendo guardar las imágenes solamente cuando el patrón de calibración es detectado correctamente. El programa también incluye la funcionalidad necesaria para el uso de archivos de imágenes previamente capturadas, muy útil en caso que sea necesario repetir el procedimiento de calibración en otro momento.

Al finalizar el procedimiento de calibración se pueden guardar y actualizar los datos de la calibración actual de la cámara como así también observar los resultados y un mapa de colores de la magnitud de la distorsión en las distintas partes de la imagen.

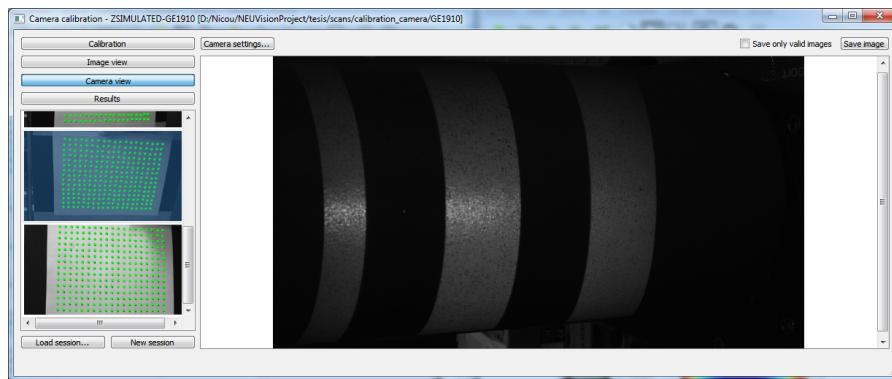
Para la calibración de los parámetros intrínsecos de la cámara se recomienda utilizar sólo 4 o 5 imágenes del patrón de calibración en diferentes ubicaciones, pero siempre se debe tratar de que los puntos de control abarquen el mayor área posible dentro de la imagen. De esta forma aproximaremos de mejor manera la distorsión, y se evita obtener una solución que sólo aproxima bien de manera local.

En nuestro caso, de acuerdo a la experiencia adquirida luego de varias pruebas y conociendo que el rango de medición es acotado, resulta suficiente con obtener las imágenes variando solamente la inclinación del patrón de calibración respecto a la cámara.

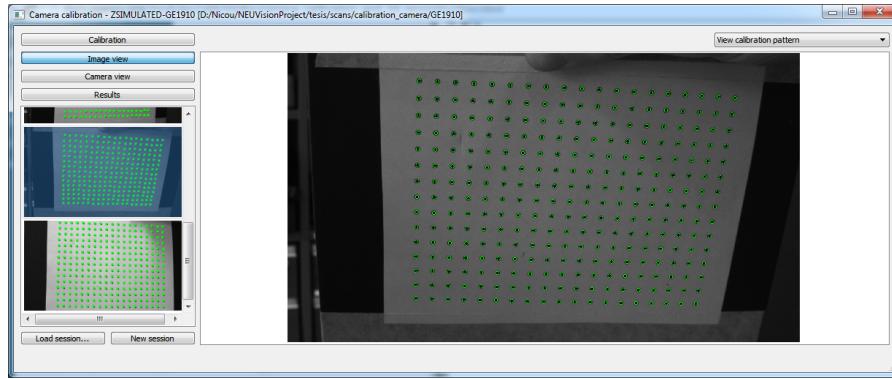
Para completar la calibración del dispositivo debemos obtener los parámetros que permiten relacionar las cámaras respecto a un único sistema de referencia. Para esto necesitaremos imágenes donde el patrón de calibración es observado simultáneamente por ambas cámaras. En este caso se recomienda el uso de una cantidad más generosa de imágenes (siempre en distintas posiciones) sin darle tanta importancia a la ubicación de los puntos de control dentro de la imagen, ya que estaremos utilizando los parámetros intrínsecos previamente calculados. Estas imágenes son utilizadas por un algoritmo de optimización para obtener la traslación y rotación entre las cámaras.



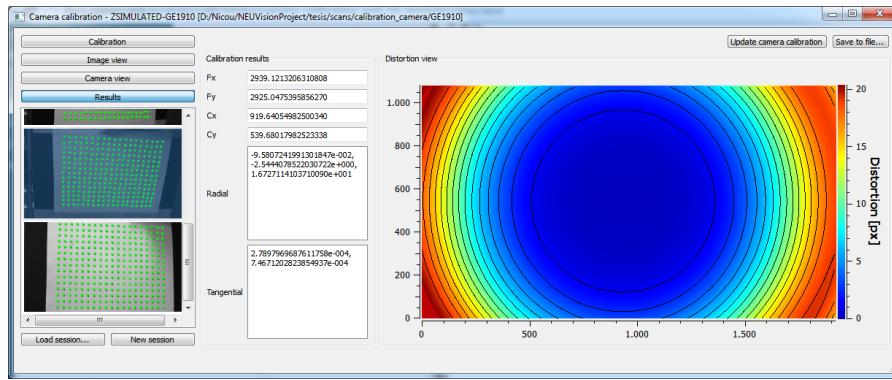
(a) Parámetros de la calibración



(b) Vista previa de la cámara



(c) Visor de imágenes adquiridas/guardadas



(d) Resultados de la calibración

Figura 43: Software de calibración de la cámara

4.6 LIMITACIONES

El dispositivo desarrollado presenta diversas limitaciones inherentes al método de medición utilizado.

La principal limitación, debido a tratarse de un método óptico, es la incapacidad de realizar mediciones directas sobre superficies brillosas, objetos que presenten alta transparencia o superficies como la piel donde la luz no se refleja sólo en la superficie sino que penetra el objeto, un efecto conocido como *subsurface scattering*²⁴. En estos casos (de ser posible) se puede cubrir la superficie del objeto con una capa fina de polvo/talco, lo que permite la obtención de resultados válidos sin importar el tipo de superficie original. En el [Apéndice C](#) se muestra un ejemplo de un inconveniente que surgió durante las pruebas.

Otra limitación presente en cualquier método basado en triangulación es la imposibilidad de realizar mediciones de agujeros profundos y secciones ocultas en superficies concavas. En este caso no hay mucho que se pueda hacer excepto reducir el ángulo de triangulación, con la perdida de precisión que eso conlleva.

Finalmente debemos mencionar también la limitación causada por la utilización de patrones de luz estructurada con codificación temporal, lo que imposibilita la realización de mediciones de escenas dinámicas u objetos que cambian su forma rápidamente.

²⁴ http://en.wikipedia.org/wiki/Subsurface_scattering

5

RESULTADOS

En este capítulo se presentan los resultados obtenidos para la medición de objetos de geometrías simples. En la primer sección se detallan los pasos realizados para calibrar el dispositivo y prepararlo para realizar las mediciones. En la segunda sección se presentan los objetos utilizados para realizar las pruebas y finalmente en la tercera sección se presentan los resultados obtenidos.

5.1 PREPARACIÓN

Para las pruebas el dispositivo fue armado con el objetivo de realizar mediciones a una distancia de entre 20 y 30 cm. La distancia entre las cámaras es de aproximadamente 17 cm. Ambas cámaras apuntan hacia un punto ubicado en el centro, a unos 20 cm del dispositivo. Las cámaras se utilizan en su máxima resolución: 1920x1080 en el caso de la cámara izquierda (AVT GE1910) y 2448x2050 en la cámara derecha (AVT GC2450).

Al completar el armado y configuración, y antes de comenzar a realizar mediciones, se debe calibrar el dispositivo. Primero se realiza la calibración de cada cámara de manera individual, obteniendo la matriz de la cámara y los parámetros de distorsión de la lente. Luego se debe realizar una calibración de los parámetros extrínsecos que relacionan una cámara respecto a la otra.

Con el fin de obtener los parámetros intrínsecos de cada cámara se tomaron entre 4 y 5 fotos del patrón de calibración. El patrón de calibración utilizado es una grilla de 19x15 círculos con una separación de 7mm entre ellos, los cuales fueron impresos sobre un papel común, el cual luego fue pegado sobre una superficie plana. La impresión se realizó desde AutoCAD utilizando una impresora láser. El patrón de calibración fue validado a partir de mediciones realizadas con instrumentos manuales, y se encontró que presenta una calidad acorde a la precisión deseada (diferencias menores a los 0,1mm entre extremos). Las imágenes obtenidas para realizar la calibración de la cámara izquierda pueden observarse en la [Figura 44](#) y en la [Figura 45](#) pueden observarse las correspondientes a la cámara derecha.

Para la obtención de la relación entre una cámara y la otra se deben obtener imágenes simultáneas desde cada cámara, en cada una de

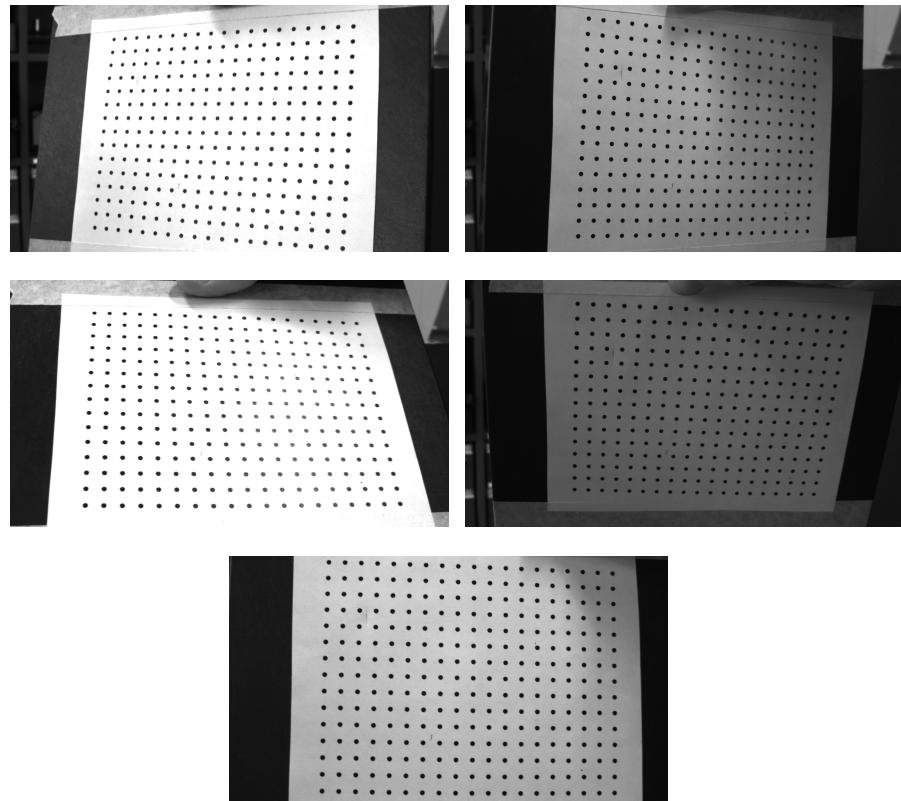


Figura 44: Imágenes de calibración para la cámara izquierda

las cuales se debe observar el patrón de calibración. En este caso el patrón de calibración es similar pero de un menor tamaño, con una grilla de círculos de 15×11 , para facilitar la vista simultánea desde ambas cámaras en diversas ubicaciones. Las imágenes utilizadas se pueden observar en la [Figura 46](#).

La configuración (ubicación y orientación de las cámaras y el proyector, resolución de las cámaras, etc) y calibración del dispositivo utilizadas en las pruebas nos permite obtener una resolución (distancia entre puntos) aproximada de entre 0,2 mm (a una distancia de 20 cm) y 0,3 mm (a 30 cm). Esta corresponde a la peor resolución, la cual se observa en el eje x (dirección horizontal), limitada por la resolución del proyector (sólo 480 columnas).

En estas condiciones la nube de puntos obtenida tiene entre quinientos mil y un millón de puntos aproximadamente. El tiempo de decodificación de los patrones de luz estructurada es de alrededor de medio segundo y el cálculo de la nube de puntos se lleva a cabo en menos de un segundo. Estos tiempos corresponden a la ejecución del software en una laptop actual.

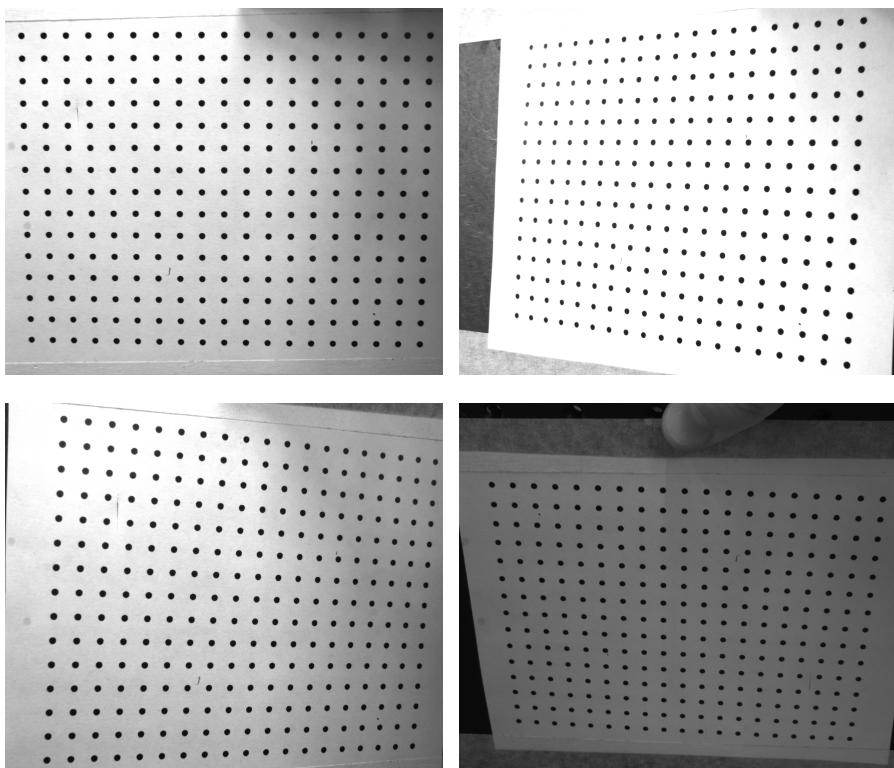


Figura 45: Imágenes de calibración para la cámara derecha

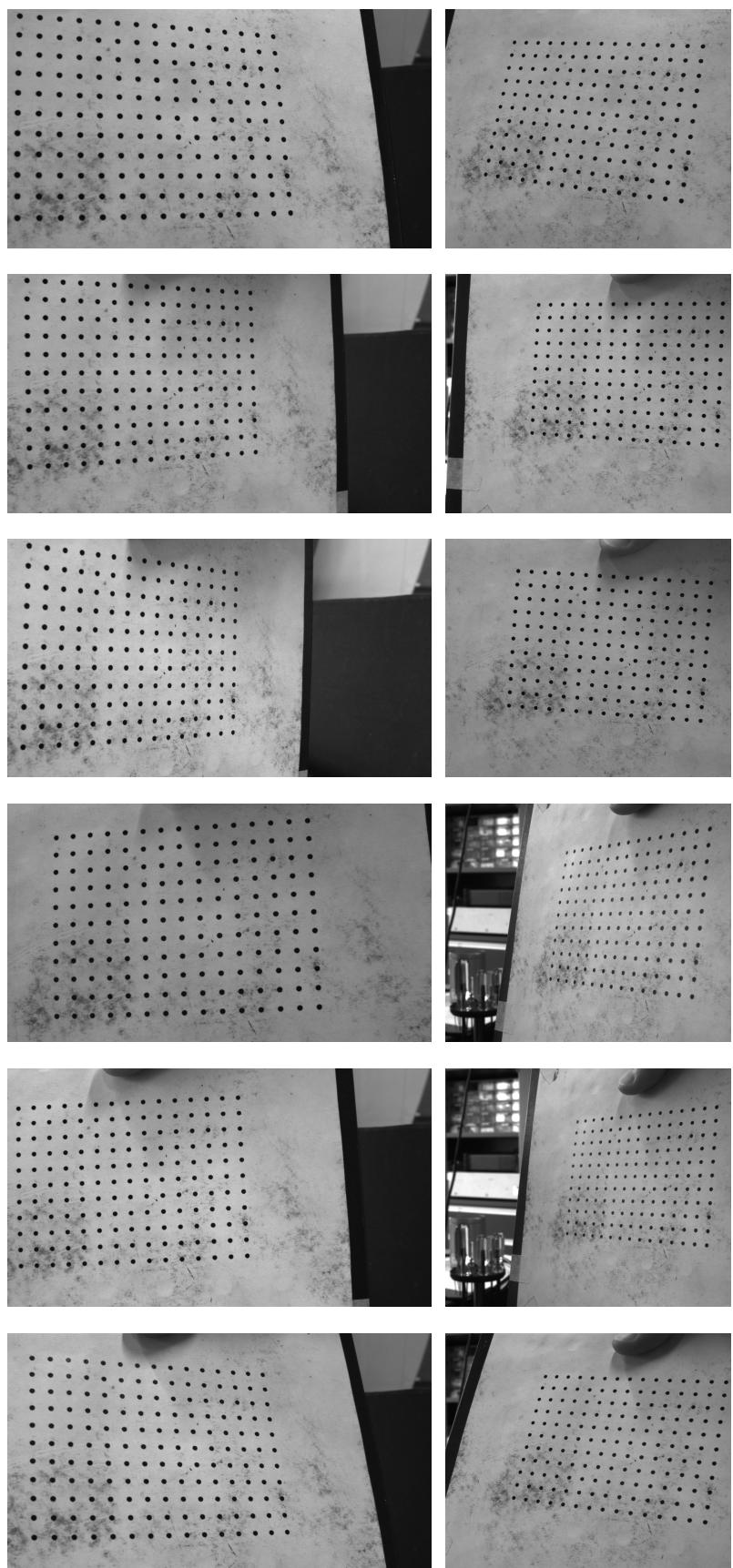


Figura 46: Imágenes para la calibración extrínseca entre cámaras. Izquierda: cámara 1, derecha: cámara 2

5.2 OBJETOS DE PRUEBA

El objetivo de las pruebas es verificar el correcto funcionamiento del dispositivo y del procedimiento de calibración, y nos permitirán obtener una estimación del error. Para facilitar esta tarea se utilizarán dos tipos de objetos simples: un objeto con una superficie plana y un objeto cilíndrico.

El objeto que utilizaremos como superficie plana es de muy alta precisión¹ y presenta una superficie un tanto brillosa de color gris oscuro, similar al mármol. La superficie oscura no es la óptima para un método de medición óptico pero la reflexión resulta suficiente para permitir su utilización. Este objeto nos permitirá evaluar fácilmente si estamos compensando correctamente la distorsión de los lentes.

El objeto cilíndrico es una cupla² pintada de color rojo que presenta una superficie relativamente opaca. Este objeto nos permitirá evaluar la calibración completa del dispositivo, los parámetros intrínsecos y extrínsecos funcionando en conjunto.

5.3 RESULTADOS

Ambos objetos fueron escaneados en distintas posiciones dentro del rango de medición para observar el comportamiento del error respecto a la distancia de medición. Los resultados presentados en esta sección fueron obtenidos sin ningún tipo de filtrado ni post-procesamiento aplicado a la nube de puntos.

5.3.1 Objeto plano

La superficie plana fue ubicada en diversas posiciones a diversas distancias respecto al dispositivo de medición. Si bien el desplazamiento no es aleatorio, tampoco fue controlado con precisión. En la Figura 47 se puede observar el objeto junto al dispositivo de medición.

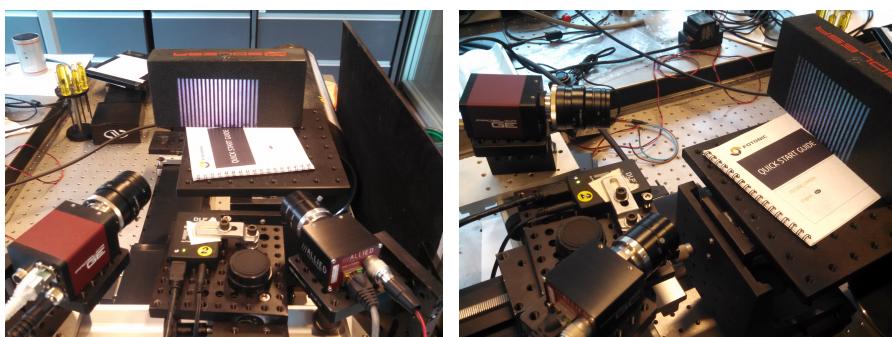


Figura 47: Medición del patrón de planicidad

¹ Patrón de planicidad marca Solesa <http://www.solesa.com.ar/>

² Una cupla es un tubo que tiene rosca por dentro y sirve para unir dos tubos

El resultado que nos interesa obtener es la estadística del error de cada punto obtenido respecto al plano que minimiza el error. El error observado presenta una distribución gaussiana por lo que puede ser caracterizado a partir del error medio y el desvío estándar. En la [Tabla 2](#) se pueden ver los resultados obtenidos.

	Distancia de medición [mm]					
	180	200	220	240	260	290
Error medio [um]	4,0	0,2	0,6	0,4	0,6	8,8
Desvío estandar [um]	77,5	55,3	51,2	57,4	65,2	87,4

Tabla 2: Resultados de la medición del objeto plano

Lo que podemos observar es que la calibración funciona razonablemente bien, con un desvío estándar del error menor a los 0,1mm. En la [Figura 48](#) se puede ver el comportamiento del desvío estándar del error respecto a la distancia.

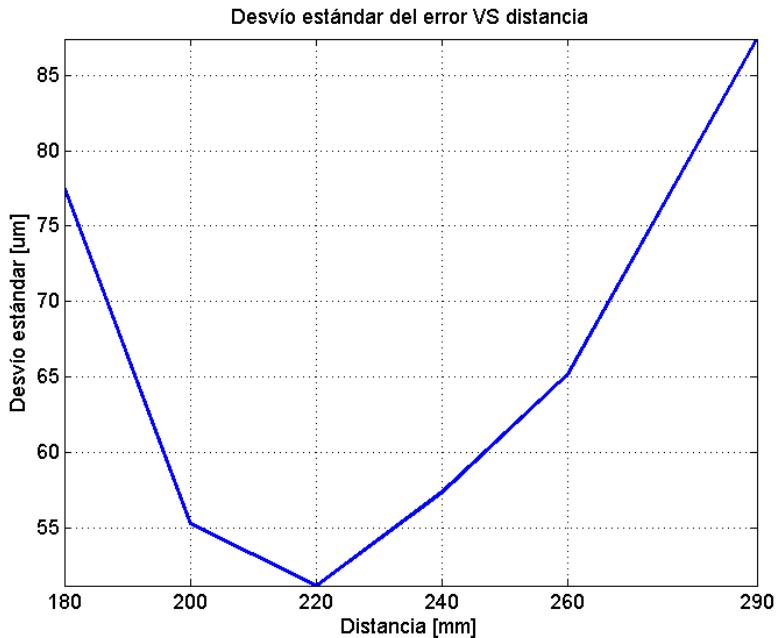


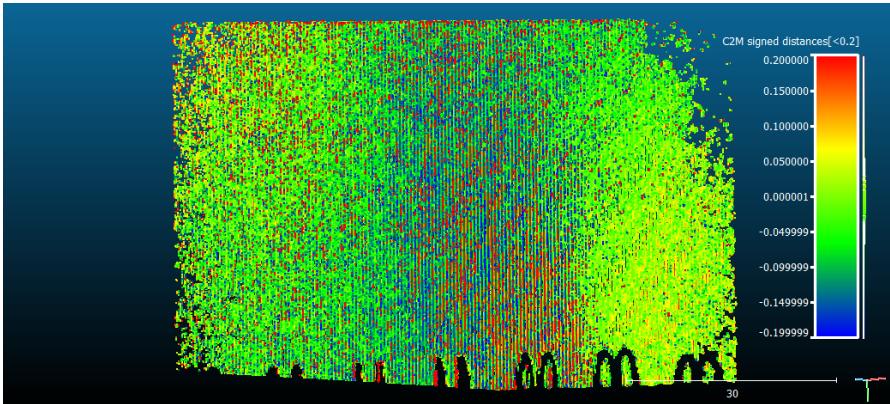
Figura 48: Objeto plano: desvío estándar del error en función de la distancia

El error debería presentar un comportamiento monotónicamente creciente respecto a la distancia, sin embargo se observa que el error es alto a la menor distancia, 180mm, y luego disminuye levemente. Las imágenes obtenidas para el plano a 180mm no estaban bien definidas, se observan fuera de foco, lo que puede explicar este comportamiento.

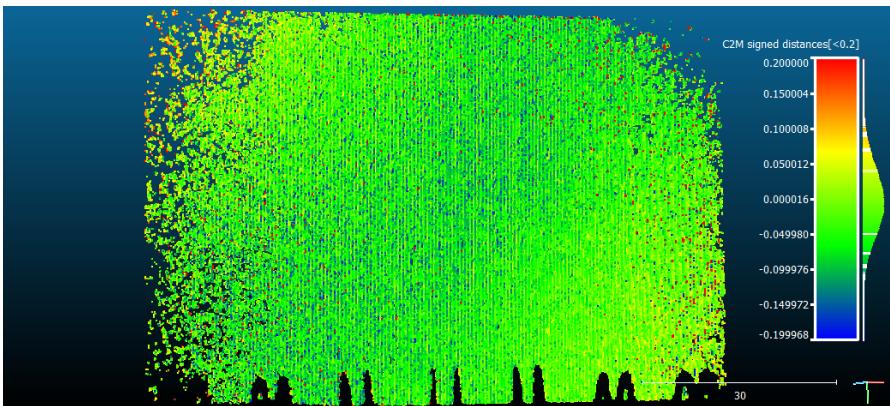
La distancia de medición que presenta los mejores resultados es la zona para la cual el dispositivo fue calibrado y los lentes fueron enfocados. Luego el comportamiento del error parece adecuarse a una

función cuadrática de la distancia, lo que se ajusta al error aleatorio teórico esperado para un sistema basado en triangulación [5].

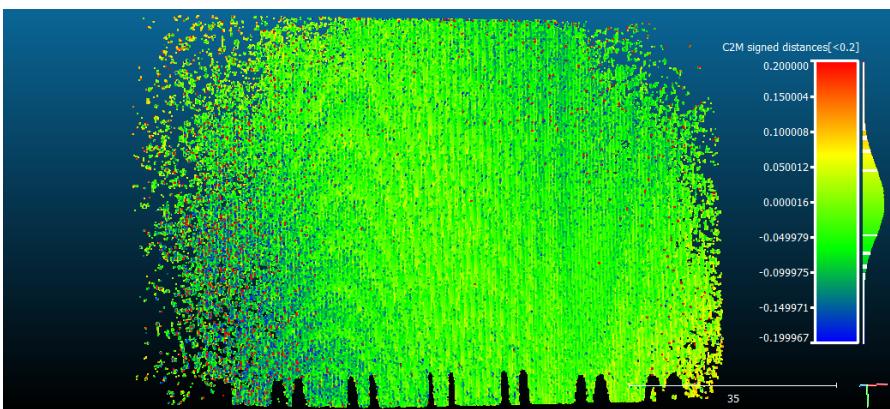
En las figuras 49 y 50 se muestran cada una de las nubes de puntos obtenidas. El color de cada punto corresponde al error respecto al plano fiteado, y a la derecha se puede observar la escala de colores junto con un histograma. Estas imágenes fueron obtenidas con el software CloudCompare.



(a) Medicion del patrón de planicidad a $\approx 18\text{cm}$

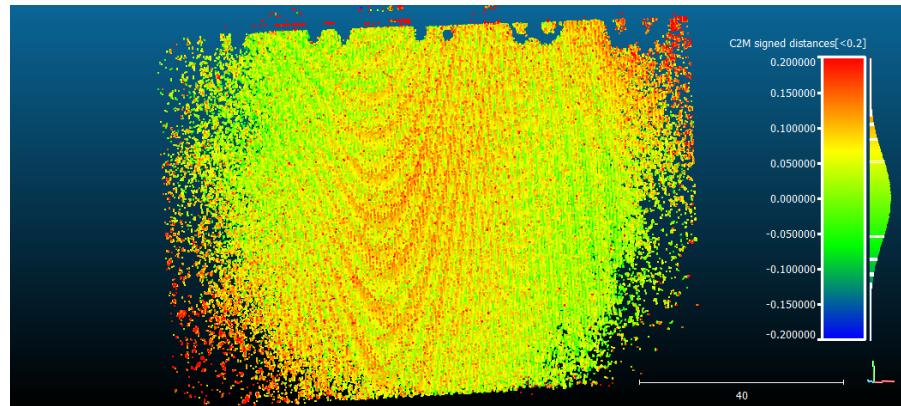
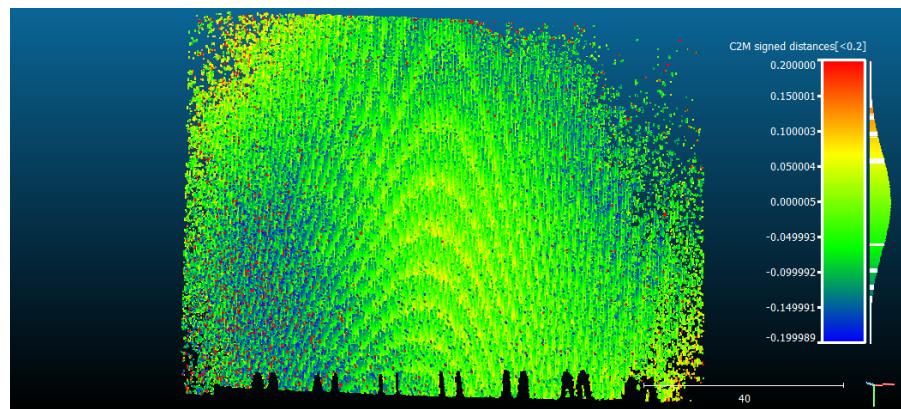
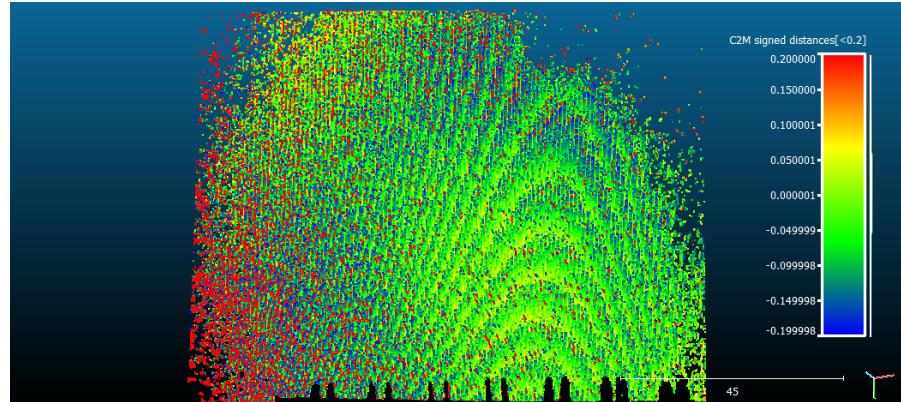
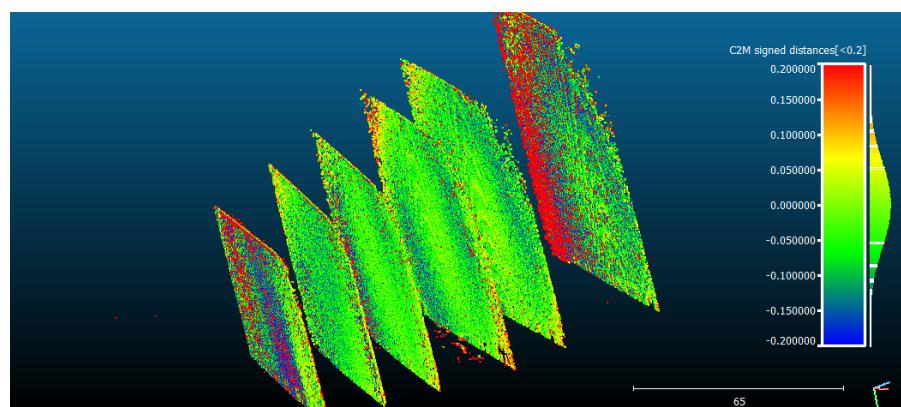


(b) Medicion del patrón de planicidad a $\approx 20\text{cm}$



(c) Medicion del patrón de planicidad a $\approx 22\text{cm}$

Figura 49: Diversas vistas de las mediciones del objeto plano

(a) Medicion del patrón de planicidad a $\approx 24\text{cm}$ (b) Medicion del patrón de planicidad a $\approx 26\text{cm}$ (c) Medicion del patrón de planicidad a $\approx 29\text{cm}$ 

(d) Todas las mediciones del patrón de planicidad

Figura 50: Diversas vistas de las mediciones del objeto plano (continuación)

5.3.2 Objeto cilíndrico

La cupla fue medida en seis distintas posiciones, y el desplazamiento fue realizado sobre una plataforma de desplazamiento lineal motorizada, abarcando un rango de 100mm (una medición cada 20mm). El desplazamiento se realiza sobre el eje z, en dirección alejándose del dispositivo de medición. En la [Figura 51](#) se puede observar la cupla al momento de la medición.

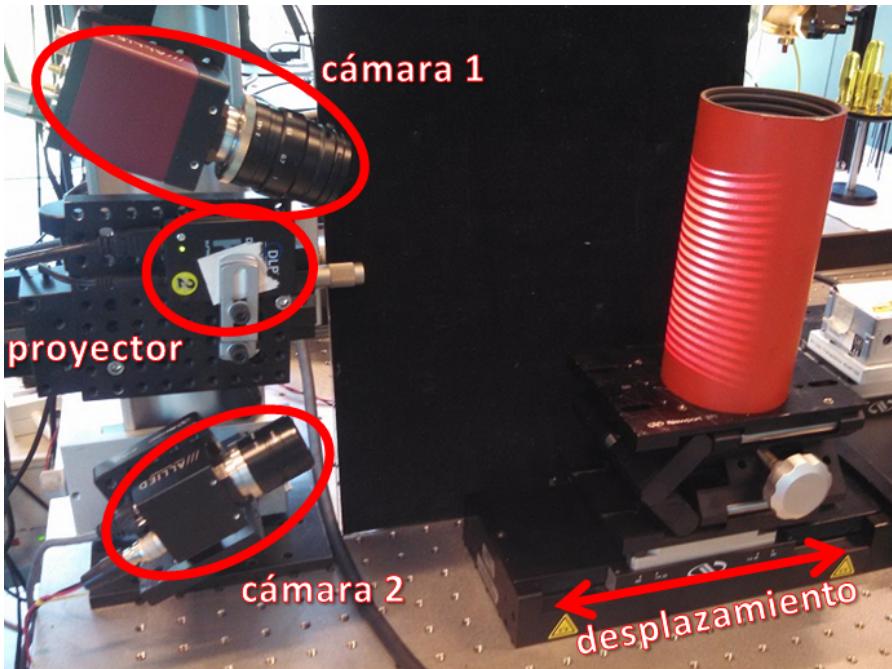


Figura 51: Medición de una cupla

Para cada medición se realiza una aproximación del cilindro que minimiza el error, y en este caso no sólo observaremos la variación de los puntos respecto al cilindro ajustado sino también el radio, la ubicación y la orientación de éste.

En la [Tabla 3](#) se muestra el radio estimado y su error respecto al real. El desvío estándar corresponde a la diferencia entre la nube de puntos y la superficie del cilindro aproximado. El diámetro del cilindro real es aproximadamente 82,55mm, con lo cual su radio es aproximadamente 41,27mm. Se puede ver que el cilindro estimado representa un error relativamente bajo (alrededor de los 0,5mm) pese a que solamente estamos observando una sección de la curvatura del objeto. En la [Figura 52](#) se muestra un gráfico donde se puede observar la variación del error respecto a la distancia entre el dispositivo de medición y el objeto. Por otro lado el desvío estándar observado es menor a 0,1mm, lo cual representa menos del 0,1 % del rango de medición (el cual es mayor a los más de 100mm que abarcamos con el desplazamiento del objeto). En la [Figura 53](#) se puede observar la distribución del error y en la [Figura 54](#) se puede observar la varia-

ción del desvío estándar respecto a la distancia a la que se realiza la medición.

	Distancia de medición [mm]					
	210	230	250	270	290	310
Radio [mm]	41,82	41,82	41,81	41,79	41,77	41,74
Error en radio [mm]	0,55	0,54	0,53	0,51	0,49	0,47
Desvío estandar [um]	46,4	49,6	55,7	65,7	80	92,2

Tabla 3: Resultados de la medición del objeto cilíndrico

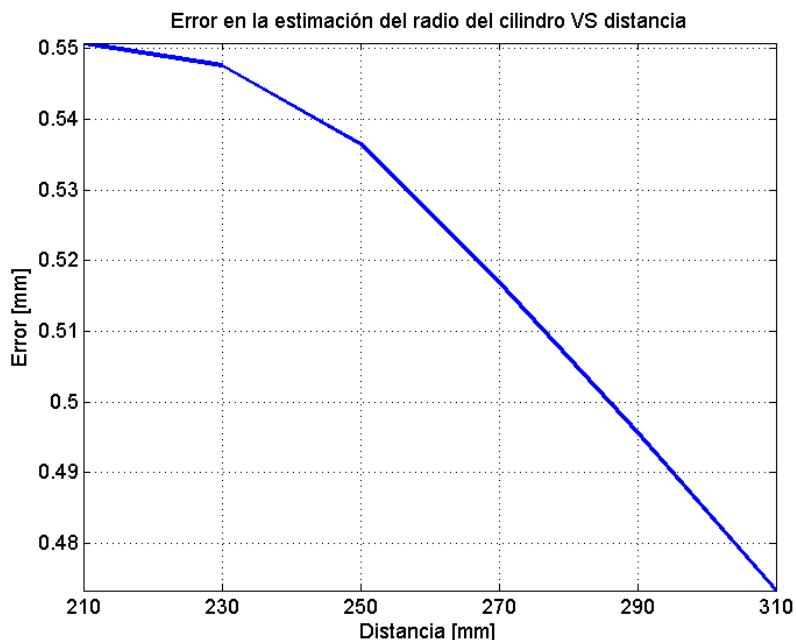


Figura 52: Objeto cilíndrico: error en la estimación del radio en función de la distancia

Otro resultado interesante de observar es la ubicación espacial del objeto, tanto la distancia observada entre el objeto y el dispositivo de medición como su orientación. Como el movimiento del objeto fue realizado sobre una plataforma de desplazamiento lineal motorizada de buena precisión, podemos asumir que todo el error observado corresponde a errores del dispositivo de medición. Para analizar este caso utilizaremos la aproximación del cilindro obtenido a partir de la nube de puntos. En la [Tabla 4](#) se pueden observar los resultados obtenidos.

La orientación puede ser analizada a partir de la dirección del eje del cilindro ajustado. En la [Figura 55](#) se puede observar la diferencia entre los ejes de los cilindros detectados en cada medición. La diferencia corresponde al ángulo entre los ejes. Se considera que el cilindro

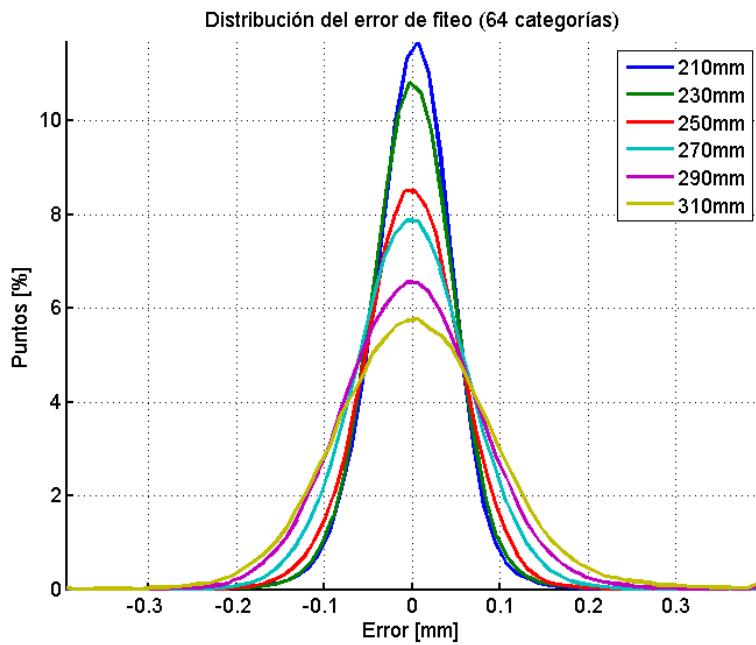


Figura 53: Objeto cilíndrico: distribución del error

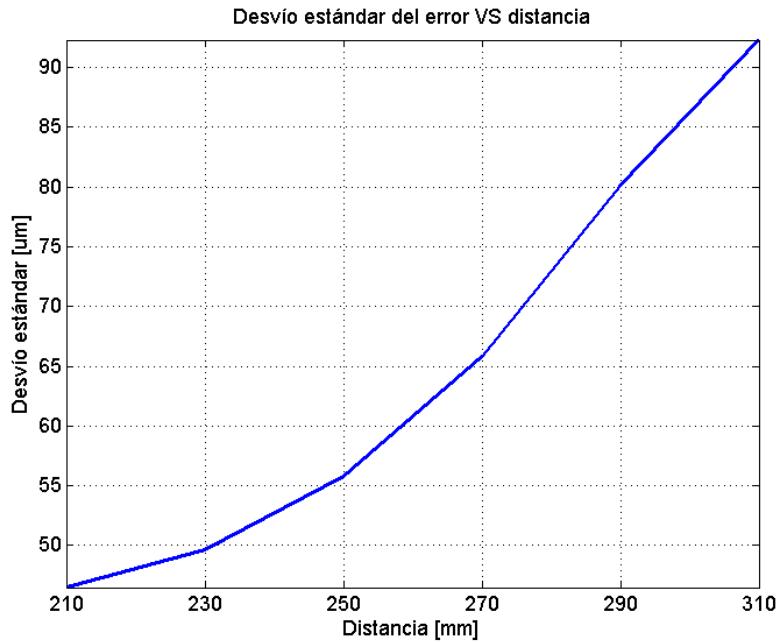


Figura 54: Objeto cilíndrico: desvío estándar del error en función de la distancia

correspondiente a la menor distancia de medición es el que presenta el menor error, por lo cual es usado como referencia. Se puede observar que la diferencia de cada eje respecto al primero es despreciable, siendo la máxima alrededor de $0,06^\circ$.

La distancia entre los cilindros detectados en cada medición se debe calcular en la misma dirección en que se realizó el desplazamiento, para ello utilizaremos una dirección perpendicular a la dirección media entre todos los ejes, debido a que el objeto se desplazó de esta manera. Los puntos entre los cuales se calcula la distancia corresponden a la intersección del eje de cada cilindro contra un plano centrado en las mediciones observadas y cuya normal corresponde a la media entre los ejes. En la [Figura 56](#) se puede observar la distancia entre los ejes de cada cilindro, nuevamente tomando como referencia el cilindro correspondiente a la menor distancia de medición. Se puede notar que el error se va acumulando y el error máximo es el observado entre los extremos, es decir entre la medición a $\approx 210\text{mm}$ y la medición a $\approx 310\text{mm}$, el cual es menor a los 0.2mm .

		Distancia de medición [mm]					
		210	230	250	270	290	310
Orientación	x	0,9232	-0,9233	-0,9234	0,9235	0,9235	-0,9236
	y	0,0362	-0,0362	-0,0361	0,0362	0,0361	-0,0364
	z	-0,3825	0,3823	0,3821	-0,3818	-0,3816	0,3815
Ubicación	x	13,449	88,428	78,929	13,324	-5,107	47,075
	y	-10,954	-9,924	-12,211	-16,699	-19,335	-19,213
	z	272,831	263,345	288,836	337,504	366,628	366,578

Tabla 4: Orientación y ubicación espacial (punto sobre el eje) del cilindro

En la [Figura 57](#) se pueden observar distintas vistas de las nubes de puntos obtenidas para cada ubicación junto con la aproximación del cilindro completo. Nuevamente el color de cada punto corresponde al error respecto al cilindro fiteado, y en la parte inferior de cada imagen se puede observar la escala de colores. Estas imágenes fueron obtenidas con el software desarrollado en este trabajo.

Se puede notar que en cada medición se obtiene información sobre aproximadamente una cuarta parte de la circunferencia del objeto cilíndrico. Esta es toda la información que utilizamos para realizar una aproximación del cilindro completo, lo que nos permite obtener una estimación de su radio y su ubicación en el espacio.

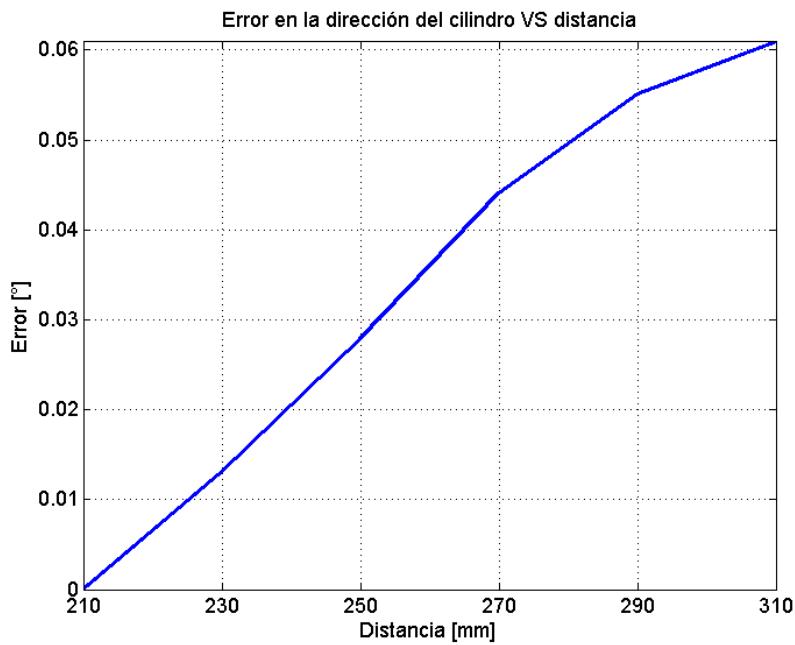


Figura 55: Objeto cilíndrico: error en la orientación del cilindro en función de la distancia

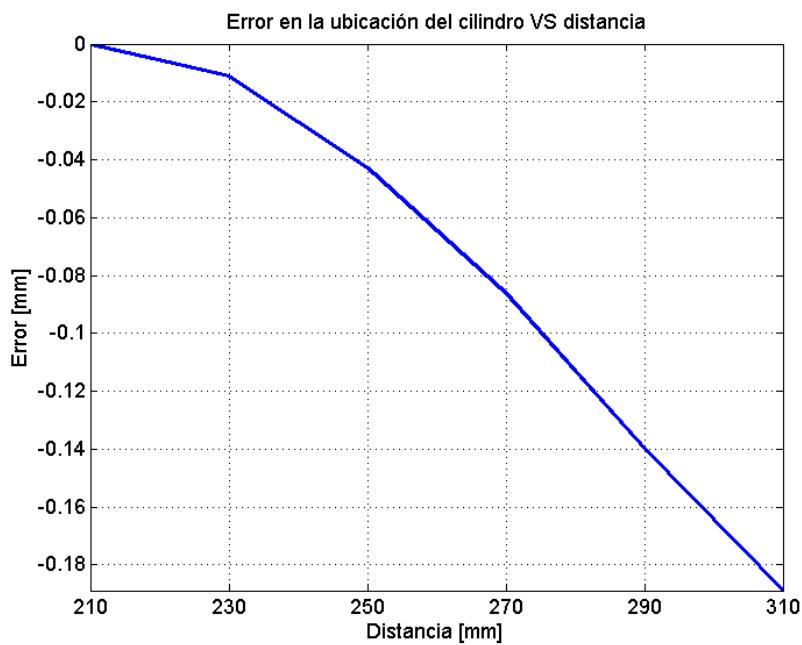


Figura 56: Objeto cilíndrico: error en la ubicación espacial del cilindro en función de la distancia

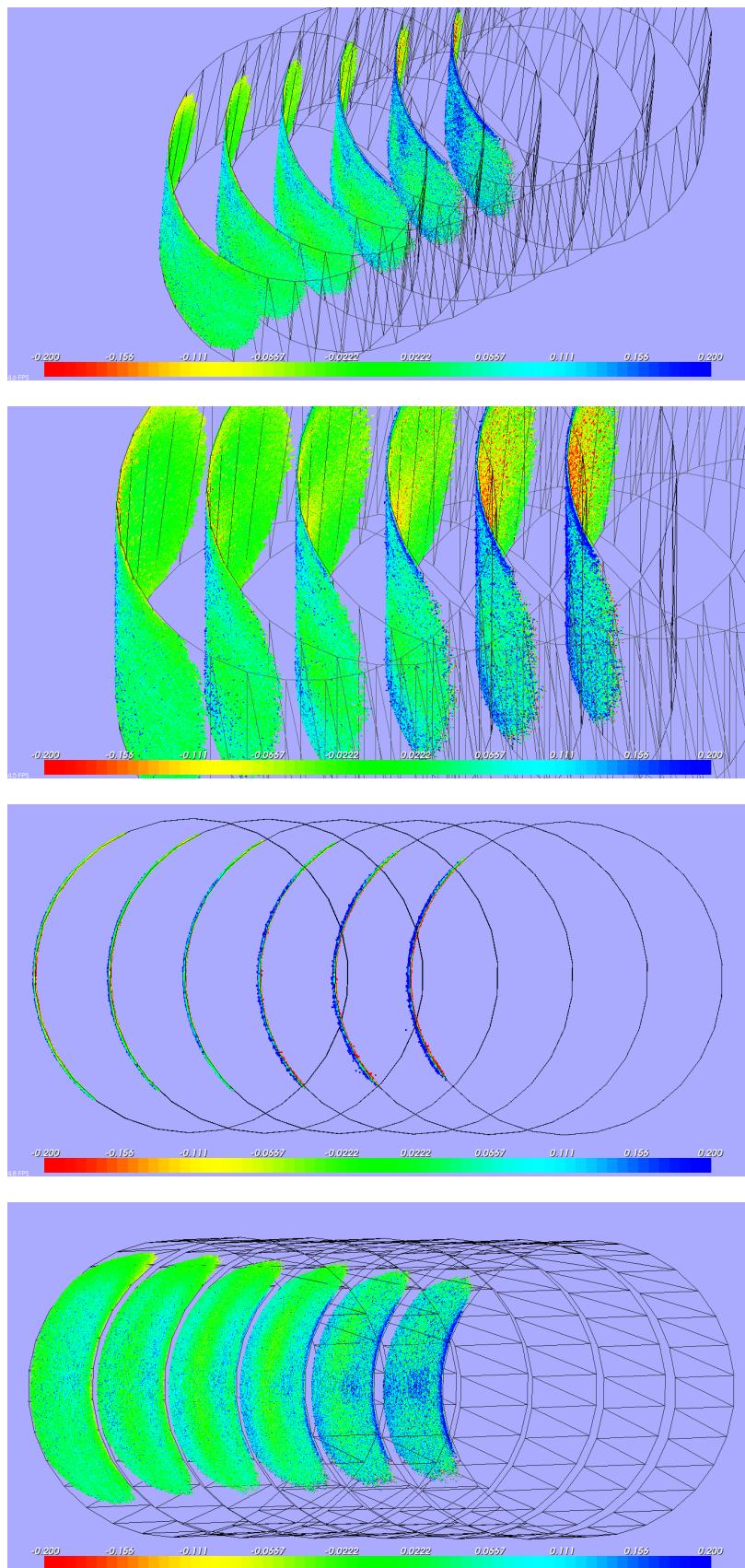


Figura 57: Diversas vistas de las mediciones del objeto cilíndrico

6

CONCLUSIONES Y TRABAJOS FUTUROS

El principal objetivo de este proyecto era el desarrollo de un dispositivo para detectar la presencia de un objeto cilíndrico en el espacio y determinar su ubicación y dimensiones. En este trabajo se realizó una introducción a los métodos ópticos de medición, se explicaron los principios matemáticos involucrados y los detalles del prototipo de luz estructurada desarrollado, incluyendo tanto el hardware como el software que lo acompaña. Luego se presentaron las diversas pruebas realizadas y los resultados obtenidos fueron expuestos. En este capítulo se presentan las conclusiones y se sugieren posibles mejoras a futuro.

6.1 CONCLUSIONES

El armado y la configuración del dispositivo son sencillos. El proceso de calibración propuesto es simple y rápido, y el software desarrollado permite que sea realizado por un usuario inexperto a partir de una serie de instrucciones sencillas. La arquitectura del software desarrollado junto con los diversos plugins implementados brinda gran flexibilidad respecto a la elección de las cámaras a utilizar, y permite la rápida incorporación de nuevos modelos matemáticos de cámaras junto a sus correspondientes métodos de calibración. El diseño modular permite también una rápida implementación e incorporación de nuevos métodos de luz estructurada.

Los resultados obtenidos muestran que el dispositivo puede realizar mediciones con un error acorde a las expectativas. Sin ningún tipo de post procesamiento y con una calibración rápida, los resultados para el objeto plano tienen un error con un desvío estándar menor a los 0,1mm para todo el rango de medición. Un desvío estándar tan bajo en el caso del objeto plano nos indica que la calibración modela correctamente la distorsión de la lente, y que el error se debe principalmente a la naturaleza discreta del sensor de las cámaras y las imágenes que se obtienen a partir de éste.

Los resultados de las pruebas de dimensionamiento del objeto cilíndrico muestran que se produce un leve sobredimensionamiento del radio del cilindro, del orden de los 0,5 mm. Este sobredimensionamiento parece ser constante para todo el rango en el que se desplaza

el objeto. Los algoritmos de ajuste de círculos a partir de arcos son conocidos por su tendencia a sobre-estimar su radio, como lo explica Berman [13]. Para confirmar esta situación se debería analizar si el error se reduce cuando utilizamos observaciones de una mayor sección del cilindro, para lo cual deberíamos unir mediciones desde distintos lugares alrededor del eje del objeto. Esta situación no es compatible con el fin práctico del dispositivo, por lo cual no fue considerado y no se realizaron las pruebas.

La detección de la ubicación del cilindro en el espacio presenta un error menor a los 0,2 mm en un rango de medición de más de 100 mm, un valor que incluye también el error debido al sobredimensionamiento del cilindro y el posible error en el eje de desplazamiento. La dirección del eje del cilindro presenta una variación despreciable, del orden de los 0,06 °. El error aleatorio observado entre la superficie del cilindro ajustado y la nube de puntos presenta nuevamente un desvío estándar menor a los 0,1 mm, similar al caso del objeto plano.

La velocidad de medición no era primordial pero incluyendo sincronización por hardware entre las cámaras y el proyector se pueden obtener las imágenes necesarias (y desocupar la pieza) en menos de un segundo si se utilizan cámaras capaces de obtener al menos 20 imágenes por segundo. El tiempo de procesamiento para la obtención de una nube de puntos de aproximadamente un millón de puntos es inferior a los dos segundos. El software utiliza lookup tables¹ para la calibración (para quitar la distorsión y realizar la transformación desde el sistema de coordenadas de la imagen al sistema de coordenadas de la cámara) y para la proyección al plano único virtual (el de la geometría epipolar), pero los algoritmos de decodificación y reconstrucción aún no fueron paralelizados, por lo que consideramos que el tiempo de procesamiento puede ser reducido aún más.

Los resultados obtenidos son prometedores y alentadores, sabiendo que existen muchas posibilidades para mejorar.

6.2 TRABAJOS FUTUROS

Durante el desarrollo de este proyecto se obtuvo una inmensa cantidad de conocimiento y se generó mucha experiencia en relación a los métodos ópticos de medición. La naturaleza acotada del proyecto hizo que muchas ideas queden sin probar, desde simples mejoras de usabilidad/practicidad hasta evidentes posibilidades de mejora en cuanto a la reducción del error. A continuación se mencionan algunas ideas que merecen un análisis mas detallado.

Desde el punto de vista del usuario final se hizo evidente que el proceso de calibración puede ser mucho más práctico. El uso de pa-

¹ Una lookup table es una estructura de datos, normalmente un vector, que se usa para substituir una rutina de computo con una simple indexación de los vectores. Más información en http://en.wikipedia.org/wiki/Lookup_table

trones de calibración con una estructura rectangular que debe ser observada por completo en cada imagen facilita las detección y validación automática del patrón. Sin embargo al momento de obtener las imágenes de calibración se observa que no es cómodo para una sola persona tener que sostener/ubicar el patrón y al mismo tiempo observar la previsualización de la imagen para conocer si es una imagen de calibración válida, y si es óptima o no. Una posible solución es el uso de patrones de calibración de mayores dimensiones y que sea posible la calibración a partir de imágenes del patrón incompleto. Otro beneficio que se obtendría es la calibración de ambas cámaras a partir de las mismas diversas ubicaciones del patrón. Si además se utiliza un patrón con alguna marca especial que nos permita conocer las coordenadas absolutas de cada punto (recordemos que se plantea la observación parcial del patrón), se pueden utilizar estas mismas imágenes para realizar la calibración extrínseca entre las cámaras. Si bien el método de calibración actual no lleva demasiado tiempo, esta propuesta reduciría aún más el tiempo requerido.

Para mejorar la resolución tanto lateral como en profundidad se puede implementar un método híbrido usando patrones binarios junto a patrones del tipo *phase-shift*. Como se comentó previamente, los patrones binarios solamente brindan información útil en los bordes entre franjas, mientras que los patrones del tipo *phase-shift* aprovechan los efectos de desenfoque de la lente del proyector para permitir el uso de información extra, incluida implícitamente entre una columna de pixels del proyector y la siguiente. Una clara ventaja es la posibilidad de aumentar la resolución, ya que se puede incluir todos los puntos visibles y no solamente los bordes entre franjas. Otra ventaja es el posible incremento en la resolución en profundidad a partir de la obtención de información más detallada respecto a los franjas, la cual podríamos asemejar a la obtención de una resolución subpixel (y no solamente cambios *discretos*, haciendo referencia a la detección del cambio entre una franja y la siguiente en el caso de patrones binarios).

A

FUENTES EN RUIDO EN CÁMARAS DIGITALES

En este apéndice se comentan brevemente las principales fuentes de ruido involucradas durante el proceso de adquisición de imágenes en una cámara digital¹.

A.1 RUIDO FOTÓNICO

La luz está compuesta por bloques discretos de energía llamados fotones. Cuanto más intensa es la luz, mayor es la cantidad de fotones por segundo que iluminan la escena. La corriente de fotones tendrá un flujo (fotones por segundo) promedio que incide sobre cierta área del sensor, y se producirán fluctuaciones alrededor de ese valor. Las fluctuaciones en la cantidad de fotones se observan como ruido en la imagen.

Estas fluctuaciones siguen una distribución de probabilidades conocida como distribución Poisson. Una característica de la distribución Poisson es que el desvío estándar es igual a la raíz cuadrada del valor promedio. Si por ejemplo la cantidad de fotones detectados en promedio es 10000, la fluctuación típica será de ± 100 fotones (el 1%). En cambio si el promedio es 100, la variación será de ± 10 (el 10%). Por lo tanto, al crecer el número de fotones, también crecerá el ruido fotónico, pero más despacio. Cuanto mayor sea la iluminación, menos aparente es el ruido fotónico; a menor iluminación, más aparente.

A.2 RUIDO DE LECTURA

Los fotones captados por los fotodetectores estimulan la emisión de electrones, uno por cada fotón. Luego de la exposición, los foto-electrones acumulados son convertidos a un voltaje proporcional a la cantidad. Luego este voltaje es amplificado y convertido a una señal digital mediante un conversor analógico-digital. En un mundo ideal, el valor digital debería ser directamente proporcional a la cantidad de fotones. Sin embargo en el mundo real estos valores no reflejan exactamente la cantidad de fotones. Cada componente del circuito electrónico sufre

¹ Una explicación mas detallada se encuentra en <http://theory.uchicago.edu/~ejm/pix/20d/tests/noise/>

fluctuaciones en el voltaje que contribuyen a una desviación del valor ideal. Estas fluctuaciones constituyen el ruido de lectura del sensor.

A.3 RUIDO TÉRMICO

Los fotodetectores son semiconductores y tienen asociado un ruido intrínseco, inherente a su constitución física, que solo depende de la temperatura absoluta. La agitación térmica en un fotodetector puede liberar algunos electrones, y estos electrones son indistinguibles de los electrones liberados por la absorción de fotones, por lo que pueden causar una diferencia en la cuenta de fotones. El sensor de la cámara se calienta tanto por la temperatura ambiente como por su propio funcionamiento.

A.4 RESPUESTA NO UNIFORME DEL PIXEL

No todos los pixels del sensor tienen exactamente la misma eficiencia en la captura de fotones. Siempre existe una variación en la cantidad de fotones detectados aún cuando no hubiera ruido de lectura, fotónico, etc, debido a la respuesta no uniforme de los pixels.

A.5 ERROR DE CUANTIZACIÓN

El conversor analógico-digital convierte el valor del voltaje del pixel en valores discretos. Debido a esto el valor real y el valor digital tienen una pequeña diferencia. Este error es conocido como error de cuantización y algunas veces también se le llama ruido de cuantización. La magnitud de este ruido es mucho menor que los anteriores.

B

EJEMPLO COMPLETO DE UNA MEDICIÓN

En la [Figura 58](#) se puede observar la nube de puntos obtenida como resultado de una medición. El color corresponde a la intensidad de cada punto, utilizando un mapa de color para hacer más clara su visualización.

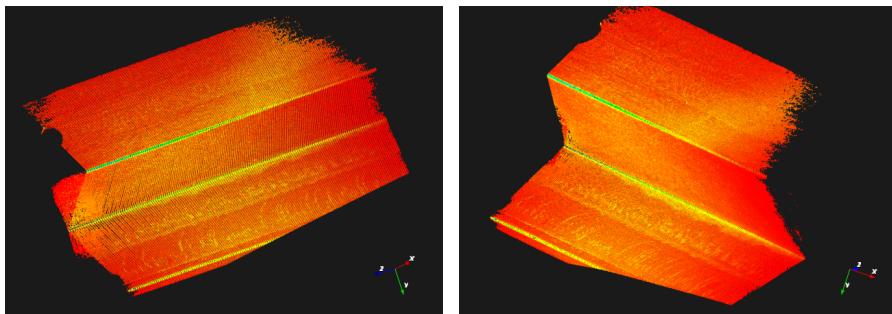


Figura 58: Ejemplo de una medición: nube de puntos

En la [Figura 59](#) se muestra la misma medición pero luego del post-procesamiento y con la superficie reconstruída utilizando triángulos. En este caso no mantenemos el color (la intensidad) de cada punto, ya que utilizamos un modelo simple de iluminación para que se puedan observar de mejor manera las variaciones en la superficie.

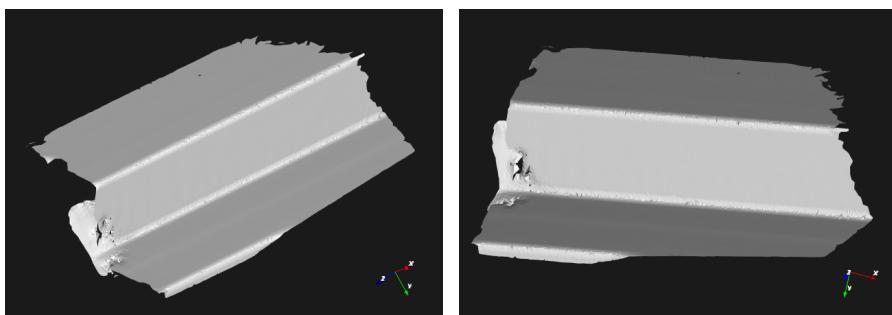


Figura 59: Ejemplo de una medición: superficie reconstruída

A continuación se muestra el conjunto completo de imágenes utilizados para obtener la medición.

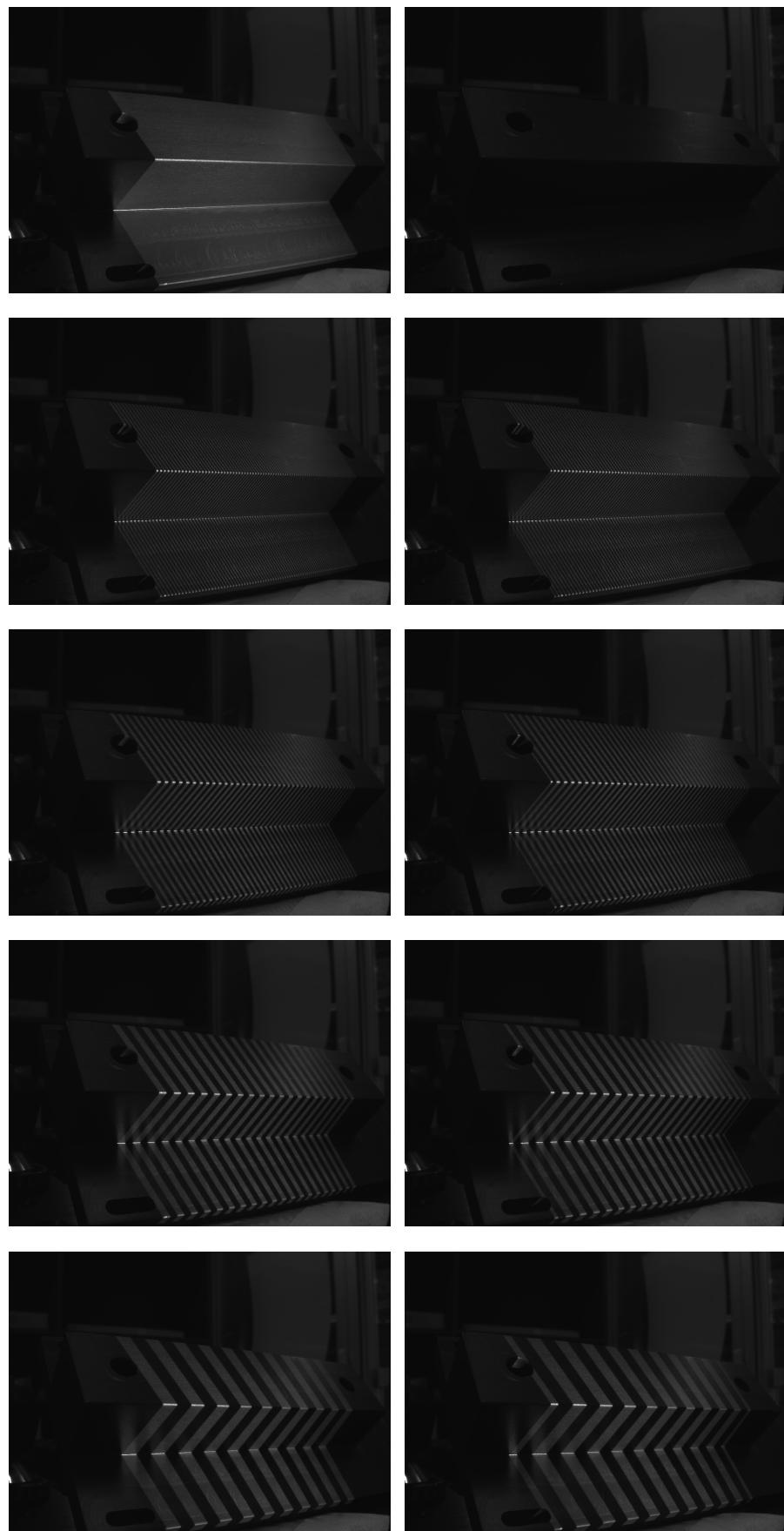


Figura 60: Cámara izquierda: 5 primeros patrones proyectados (izq.) y su inverso (der.)

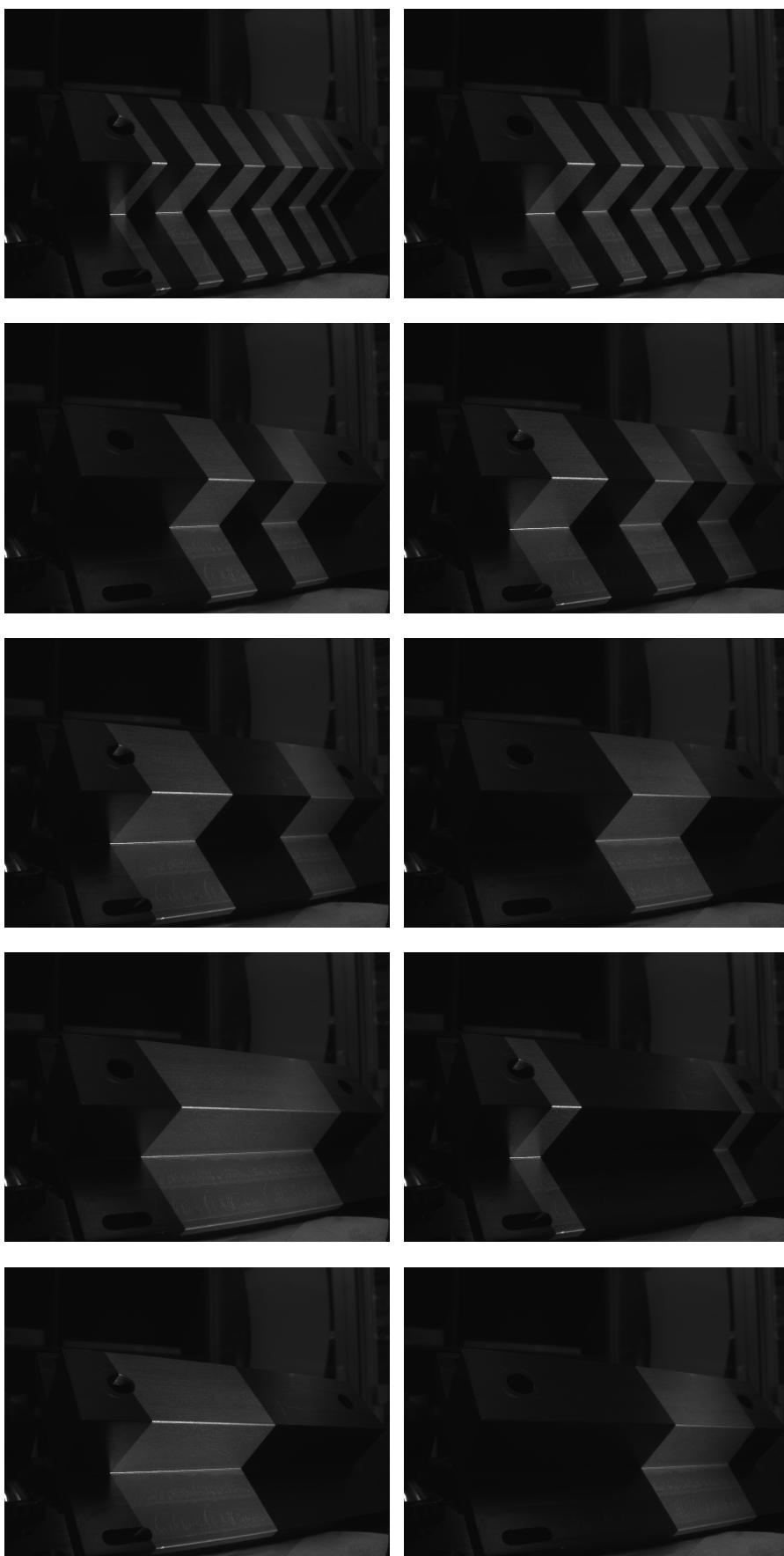


Figura 61: Cámara izquierda: 5 últimos patrones proyectados (izq.) y su inverso (der.)

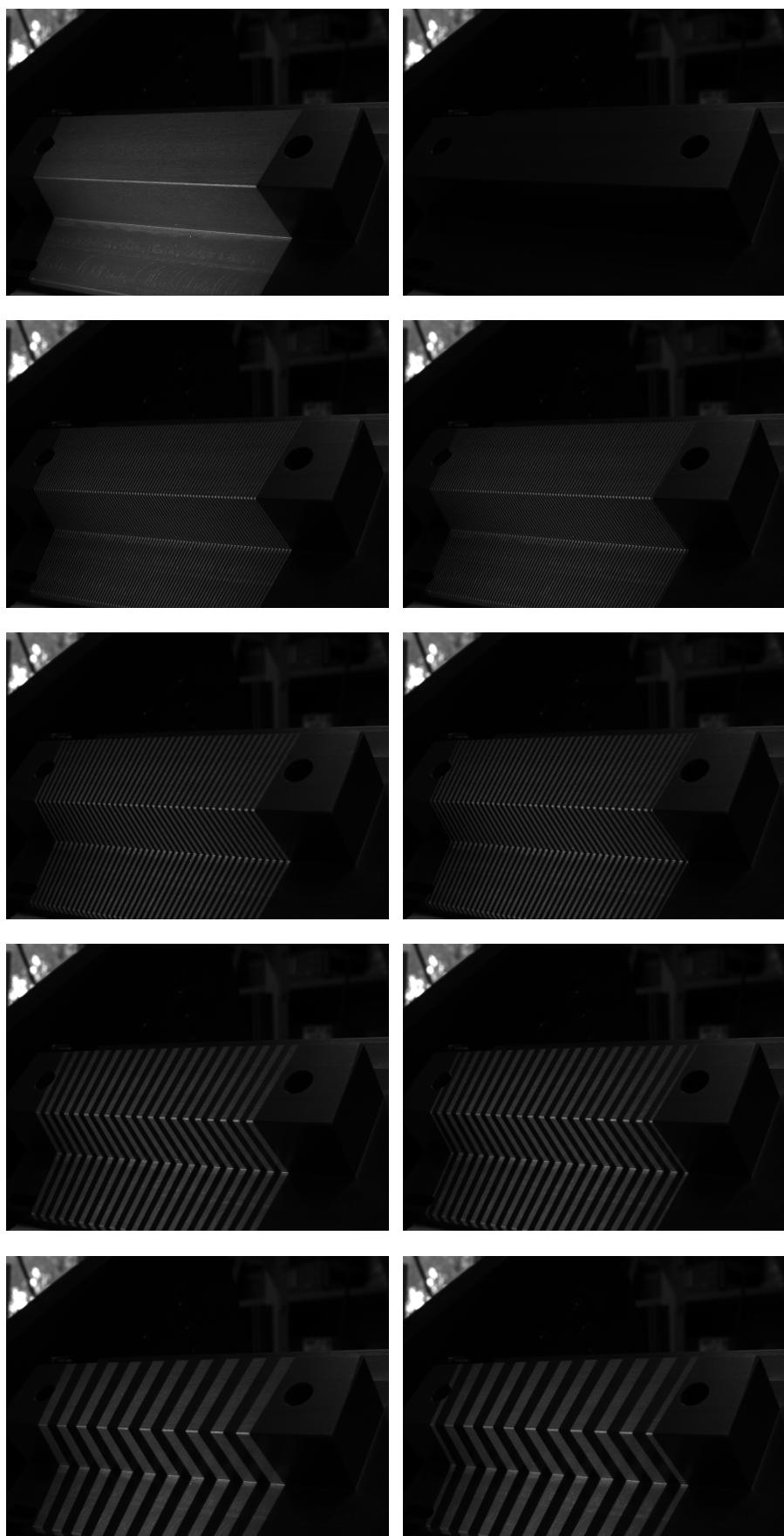


Figura 62: Cámara derecha: 5 primeros patrones proyectados (izq.) y su inverso (der.)

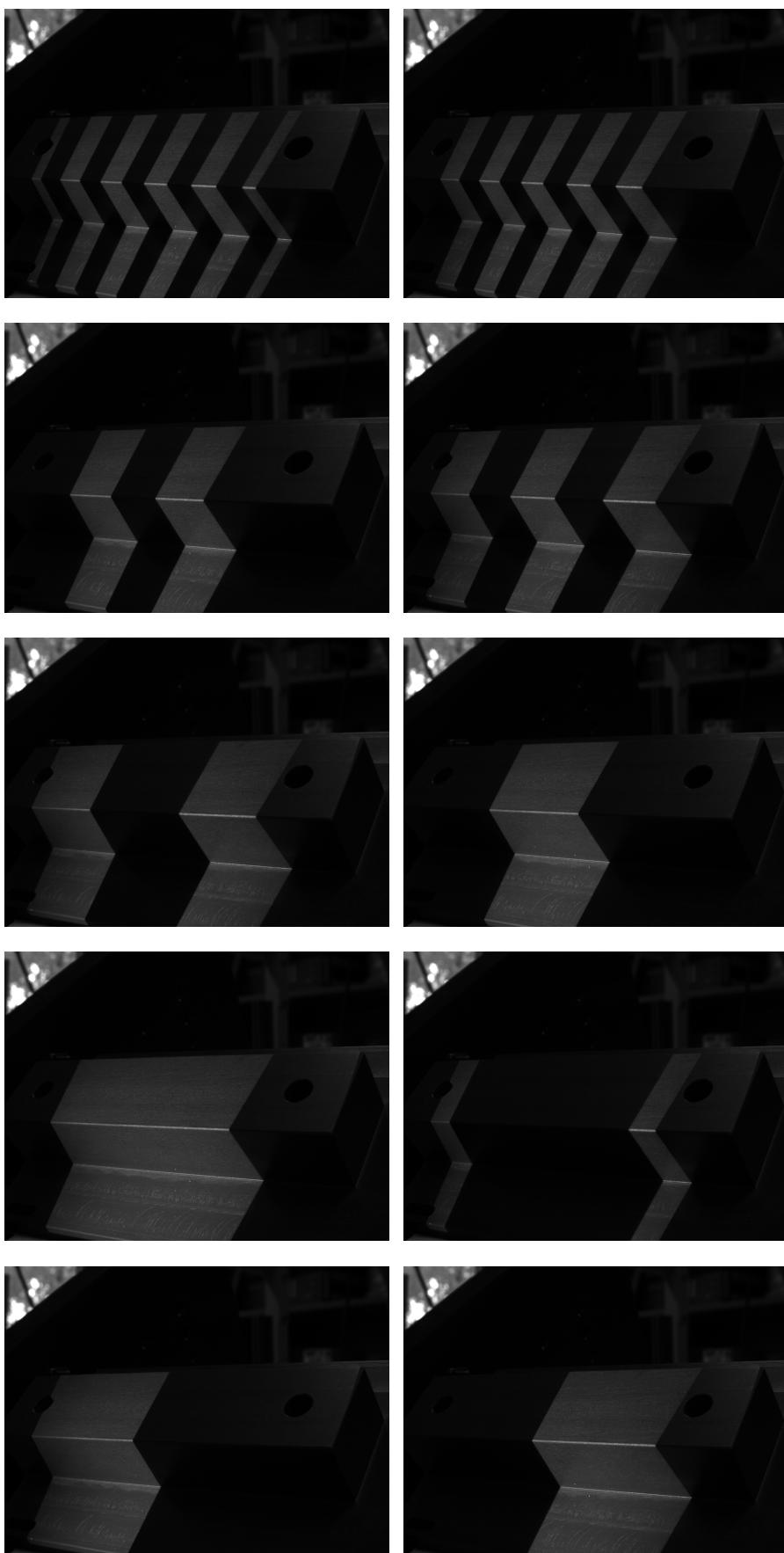


Figura 63: Cámara derecha: 5 últimos patrones proyectados (izq.) y su inverso (der.)

C

SUPERFICIES BRILLOSAS

En la [Figura 64](#) se puede observar un ejemplo de las limitaciones del dispositivo. Este ejemplo corresponde a un sector de la medición del [Apéndice B](#). La superficie real del objeto no tenía ningun defecto ni discontinuidad en la sección observada. Sin embargo las reflexiones causadas por la superficie brillosa hacen que la decodificación de los patrones sea incorrecta.

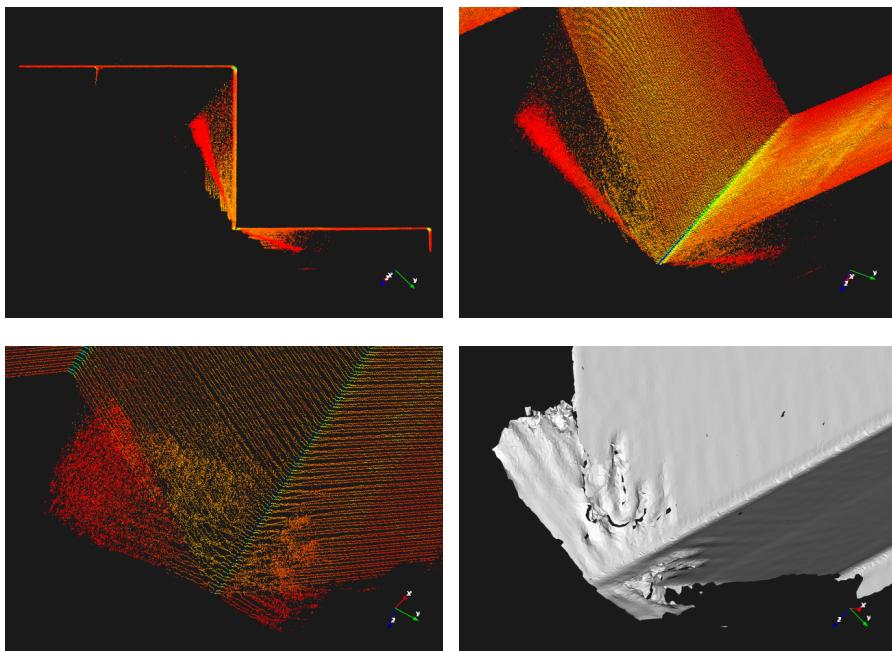


Figura 64: Ejemplo de un problema causado por superficie brillosa

En la [Figura 65](#) se puede observar en detalle la sección de las imágenes donde se produce el problema. Queda claro que el reflejo de la superficie inferior sobre la superior es observado como un patrón fantasma/incorrecto, y lo mismo ocurre en la otra dirección (reflejo de la superficie superior sobre la inferior). Esto se debe a la combinación de la superficie brillosa junto con la dirección del eje de observación (ubicación de la cámara), la inclinación de las superficies de la pieza y la dirección de incidencia de la luz (ubicación del proyector).

Este inconveniente no es observado en las imágenes correspondientes a la otra cámara, debido a que se forma otro ángulo entre la superficie, el proyector y la cámara, lo cual reduce considerablemente la magnitud del reflejo.

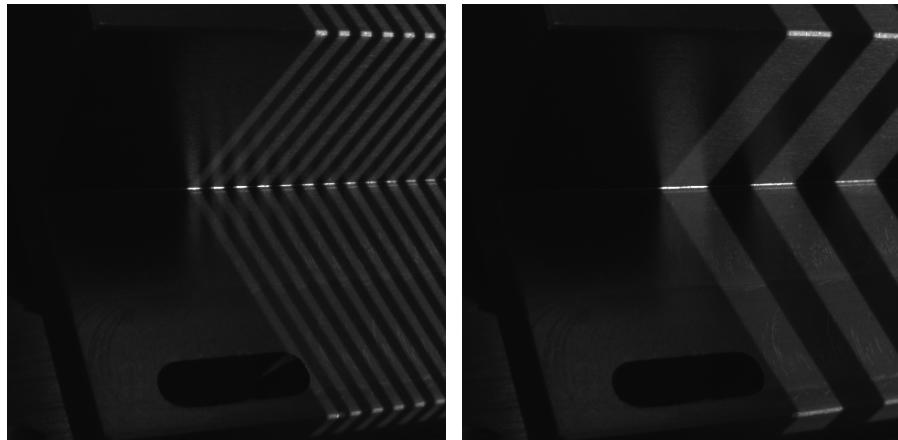


Figura 65: Causas de la decodificación incorrecta

En este caso la solución es simple, simplemente se busca una ubicación en la cual este inconveniente no afecte a ninguna de las cámaras. En el caso de que la superficie sea muy brillosa, muy oscura o transparente, se puede utilizar una capa muy fina de talco/polvo. Se pueden comprar aerosoles especialmente diseñados que cubren la superficie con una capa muy fina de polvo, logrando obtener una apariencia mate y de un color claro. Estos aerosoles contienen una suspensión de partículas blancas en un solvente de secado rápido. Un ejemplo¹ se puede ver en la Figura 66.



Figura 66: Spray para la medición de superficies brillosas

¹ https://ssl.david-vision-systems.de/shop/product_info.php/info/p105_3D-Coating-Spray.html

BIBLIOGRAFÍA

- [1] F. Chen, G. M. Brown, and M. Song, "Overview of three-dimensional shape measurement using optical methods," *Optical Engineering*, vol. 39, no. 1, pp. 10–22, 2000.
- [2] G. Sansoni, M. Trebeschi, and F. Docchio, "State-of-the-art and applications of 3d imaging sensors in industry, cultural heritage, medicine, and criminal investigation," *Sensors*, vol. 9, no. 1, pp. 568–601, 2009.
- [3] S. S. Gorthi and P. Rastogi, "Fringe projection techniques: whither we are?," *Optics and Lasers in Engineering*, vol. 48, no. IMAC-REVIEW-2009-001, pp. 133–140, 2010.
- [4] G. Häusler and S. Ettl, "Limitations of optical 3d sensors," in *Optical Measurement of Surface Topography*, pp. 23–48, Springer, 2011.
- [5] K. Khoshelham and S. O. Elberink, "Accuracy and resolution of kinect depth data for indoor mapping applications," *Sensors*, vol. 12, no. 2, pp. 1437–1454, 2012.
- [6] D. Lanman and G. Taubin, "Build your own 3d scanner: 3d photography for beginners," in *ACM SIGGRAPH 2009 Courses*, p. 8, ACM, 2009.
- [7] P. Sturm, S. Ramalingam, J.-P. Tardif, S. Gasparini, and J. Barreto, "Camera models and fundamental concepts used in geometric computer vision," *Foundations and Trends® in Computer Graphics and Vision*, vol. 6, no. 1–2, pp. 1–183, 2011.
- [8] T. Hanning, *High Precision Camera Calibration*. Springer, 2011.
- [9] J. E. Greivenkamp, *Field guide to geometrical optics*, vol. 1. SPIE Press Bellingham, Washington, 2004.
- [10] G. Bradski and A. Kaehler, *Learning OpenCV: Computer vision with the OpenCV library*. O'Reilly Media, Inc., 2008.
- [11] Z. Zhang, "A flexible new technique for camera calibration," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 22, no. 11, pp. 1330–1334, 2000.
- [12] M. Tröbina, "Error model of a coded-light range sensor," *Technique Report, Communication Technology Laboratory, ETH Zentrum, Zurich*, 1995.

- [13] M. Berman, "Large sample bias in least squares estimators of a circular arc center and its radius," *Computer Vision, Graphics, and Image Processing*, vol. 45, no. 1, pp. 126–128, 1989.