

1.4.2 Específicos.

1. Colocar a punto un método para registro de múltiples vistas de imágenes de rango, enfocado en el problema de la obtención de la mejor correspondencia entre puntos, empleando un método robusto de búsqueda para imágenes que están parcialmente solapadas.
2. Diseñar un método para integrar imágenes de rango, de objetos de forma libre, que permita el llenado automático de huecos de modelos triangulares, empleando funciones de base radial.
3. Proponer un método para la obtención del modelo computacional que ajuste en forma suave objetos de forma libre a partir de una red parametrizada de parches NURBS, empleando técnicas de optimización.
4. Validar la aproximación metodológica propuesta, para la reconstrucción de objetos de forma libre en aplicaciones tales como: esculturas, rostros, entre otros. Empleando como parámetros de evaluación la exactitud y la complejidad computacional.

1.5 Contribuciones.

El aporte general de esta tesis es una metodología para la reconstrucción de superficies de objetos de forma libre, a partir de imágenes de rango, utilizando cuadrilaterización de mallas triangulares como paso intermedio para el ajuste de una red de parches NURBS optimizados.

Las contribuciones específicas son:

- Un método semiautomático para hallar la correspondencia de múltiples vistas de imágenes de rango a partir de un enfoque basado en algoritmos genéticos. Este método se orienta al problema de la obtención de la mejor correspondencia entre conjuntos de puntos, a través de un método robusto de búsqueda sobre imágenes que están parcialmente solapadas. Este conjunto de correspondencias, permitirá calcular una transformada que superponga con precisión las imágenes. El algoritmo genético emplea subdominios para reducir el espacio de búsqueda y mejorar la eficiencia.
- Un método automático para la detección, análisis y llenado de huecos en modelos triangulares. En el método, cada hueco es representado como un conjunto de aristas límites que definen su contorno, el cual es ajustado mediante una curva de tipo B-Spline, sobre la cual se realiza una estimación de la variación de la torsión, en cada uno

de los vértices de la curva de contorno. Cada hueco es finalmente clasificado mediante un umbral de la variación de la torsión, y es llenado mediante un interpolante de funciones de base radial definidas localmente alrededor del contorno. El interpolante es utilizado para generar nuevos puntos con igual densidad que la malla original del vecindario existente alrededor del hueco. El conjunto de puntos es triangulado junto con el conjunto de vértices del contorno, lo que permite una unión de la nueva región de la superficie con la malla original.

- Un método para ajustar superficies de objetos de forma libre empleando una red de parches NURBS que son generados a partir de la cuadrilaterización de una malla triangular mediante análisis espectral y teoría de Morse. En el método, la teoría de Morse es usada para analizar la estructura de los campos que se generan directamente sobre la forma intrínseca de la superficie. Posteriormente, la malla cuadrilateral es regularizada mediante curvas geodésicas y B-Spline para trazar los parches NURBS, los cuales son optimizados mediante estrategias evolutivas, y unidos de tal manera que se garantiza continuidad C^1 .

1.6 Organización.

Esta tesis está organizada de la siguiente manera: En el capítulo 2, se presenta la fundamentación del proceso de reconstrucción tridimensional. En el capítulo 3, se describe el método propuesto para el registro de imágenes de rango parcialmente solapadas empleando algoritmos genéticos. En el capítulo 4, se describe el método propuesto para el llenado de huecos en mallas triangulares empleando funciones de base radial. En el capítulo 5 se describe el método para el ajuste de superficies mediante parches NURBS optimizados. En el capítulo 6 se presentan los resultados obtenidos de la aplicación de la metodología propuesta para la reconstrucción tridimensional de objetos de forma libre. Por último, en el capítulo 7 se presentan las conclusiones.

Capítulo 2

RECONSTRUCCIÓN DE OBJETOS A PARTIR DE IMÁGENES DE RANGO: FUNDAMENTOS.

La reconstrucción tridimensional no es una tarea trivial, ella cuenta con una variedad de etapas cuyo objetivo principal es obtener un algoritmo que sea capaz de realizar la conexión del conjunto de puntos representativos del objeto en forma de elementos de superficie, ya sean triángulos, cuadrados o cualquier otra forma geométrica.

En general, el proceso de reconstrucción tridimensional, involucra las etapas conocidas como: adquisición, registro, integración y ajuste, siendo la última etapa, la que proporciona el modelo computacional del objeto representado. Generalmente una sola imagen de rango no es suficiente para representar un escena completa, así que múltiples imágenes deben ser adquiridas y registradas para formar un simple conjunto de puntos. La Figura 2.1 muestra el proceso de conversión de una escena del mundo real, representada en un conjunto de imágenes parciales a un modelo tridimensional completo [76].

La etapa de adquisición consiste en el sensado de la superficie del objeto en un número determinado de vistas o imágenes de rango. El registro de múltiples imágenes solapadas consiste en llevar todas las imágenes adquiridas a un sistema de coordenadas común, esto puede crear datos redundantes, los cuales pueden ser integrados en un conjunto de puntos menos denso. La etapa de integración tiene como objetivo disminuir los datos redundantes y generar datos en regiones con ausencia de información. Sobre la nube de puntos obtenida, se estima un modelo matemático en la etapa de ajuste de superficies. Los modelos de superficie pueden tener una variedad de representaciones, una de las más utilizadas es la malla triangular. Las

técnicas de creación de mallas pueden producir superficies demasiado densas y pobremente formadas, por lo que generalmente, se aplican técnicas de simplificación de mallas para mejorar el modelo.



Figura 2.1: Proceso de reconstrucción tridimensional.

Este capítulo está organizado de la siguiente manera: En la sección 2.1, se presenta la definición de imágenes de rango. En la sección 2.2, se presenta la definición de objetos de forma libre. En la sección 2.3, se presentan las aplicaciones. En la sección 2.4 se describe la etapa de adquisición. En la sección 2.5 se describe la etapa de registro. En la sección 2.6 se describe la etapa de integración, y por último, en la sección 2.7, se describe la etapa de ajuste de superficies.

2.1 Imágenes de Rango.

Las imágenes de intensidad son de uso limitado en términos de estimación de superficies. Los valores del píxel están relacionados con la geometría de la superficie sólo indirectamente. Las imágenes de rango son una clase especial de imágenes digitales, conocidas como imágenes de profundidad, mapas de profundidad, mapas xyz, perfiles de superficies e imágenes 2.5-D. Éstas pueden ser representadas en dos formas básicas: como una lista de coordenadas 3-D en un marco de referencia dado (nube de puntos), para lo cual no se requiere ningún orden específico, o como una matriz de valores de profundidad de puntos a lo largo de las direcciones de los ejes x, y de la imagen, lo cual hace explícito la organización espacial. En términos generales, las imágenes de rango codifican la posición de la superficie directamente, por lo tanto, la forma puede ser obtenida fácilmente. Cada píxel de una imagen de rango, expresa la distancia entre un marco de referencia conocido (generalmente asociado al sensor) y un

punto de vista visible en la escena. Por lo tanto, una imagen de rango reproduce la estructura 3-D de una escena (ver Figura 2.2).

Formalmente una imagen de rango se define como un conjunto de k muestras discretas y una función escalar $J : I^2 \rightarrow \mathbb{R}$ con $r_i = J(u_i)$ donde $u_i \in I^2$ es el índice de la malla que define la imagen, $r_i \in \mathbb{R}$ e $i = \{1, 2, \dots, k\}$. Esta imagen puede ser expresada como información de profundidad, definiendo un conjunto de k muestras puntuales de la función vectorial $H : I^2 \rightarrow \mathbb{R}^3$, $d_i = H(u_i)$, donde $d_i \in \mathbb{R}^3$ e $i = \{1, 2, \dots, k\}$, donde cada punto muestreado tiene coordenadas 3-D (x_i, y_i, z_i) , decir, $(x_i, y_i) = u_i$ y $z_i = r_i = J(u_i)$ [34].

2.2 Objetos de Forma Libre.

Las definiciones de superficies y objetos de forma libre son a menudo intuitivas en lugar de formales. Los adjetivos sinónimos incluyen esculpidas, flujo libre, suave por trozos para algunos grados de continuidad n . A menudo, “la forma libre” es una caracterización general de un objeto cuyas superficies no son de una clase fácilmente reconocida, como las superficies planas y/o cuadráticas. Se asume a menudo que un objeto de forma libre está compuesto de una o más superficies no planas y no cuadráticas (superficies de forma libre). Una caracterización aproximadamente equivalente fue proporcionada por Besl [10]: “*a free-form surface has a well defined surface normal that is continuous almost everywhere except at vertices, edges and cusps*”. Dorai y Jain [40], Besl [10], y Stein y Medioni [102] hacen referencia a esculturas, carrocerías, cascos de barco, aviones, rostros humanos, órganos, y mapas de terreno como ejemplos típicos de objetos de forma libre (ver Figura 2.3). Específicamente excluidos de esta clase de objetos, están las formas definidas estadísticamente como texturas y espumas, objetos infinitamente detallados que poseen autosimilaridad, que se describen mejor usando modelos fractales, y superficies no orientables como las cintas de Moebius y botellas de Klein [21].

2.3 Aplicaciones.

Las aplicaciones de la digitalización y reconstrucción de formas 3-D son de gran alcance e incluyen manufactura, simulación virtual, exploración científica, medicina, negocios, entre otros [33].

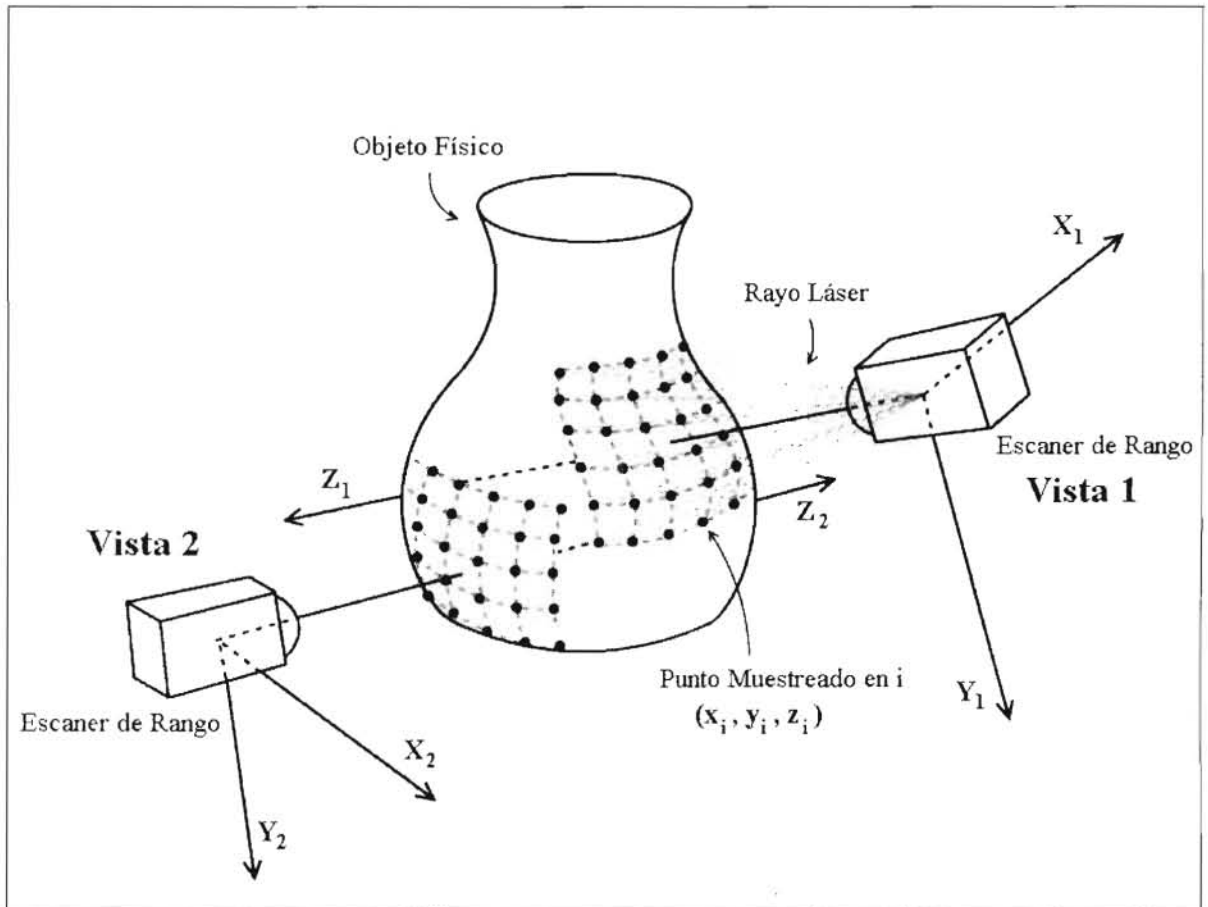


Figura 2.2: Proceso de adquisición de la imagen de rango para dos vistas diferentes de un mismo objeto [100].



Figura 2.3: Objetos de forma libre.

2.3.1 Ingeniería inversa.

Muchas partes manufacturables actualmente son diseñadas con software de Diseño Asistido por Computador (CAD). Sin embargo, en algunas instancias, una parte mecánica existe y pertenece a un sistema activo, pero no se tiene ningún modelo computarizado para reconstruir las partes. Con frecuencia, este es el caso de máquinas actualmente en servicio que fueron diseñadas antes del advenimiento de los computadores y sistemas CAD, así como partes que fueron encajadas a mano dentro de maquinarias existentes. En los casos donde no existen ni partes de repuesto ni moldes para cambiarla, entonces sería posible remover una parte de un sistema activo y digitalizarla precisamente para su manufactura.

2.3.2 Inspección.

Después que un constructor ha creado un modelo computarizado de una pieza, ya sea por medio de la digitalización de la forma de un modelo físico, o a través de un diseño CAD interactivo, él posee una variedad de opciones para fabricar esta parte, como un prototipo funcional y como punto de partida para su producción por lotes. Finalmente, las dimensiones de la parte fabricada deben caer dentro de algunas tolerancias de funcionamiento. En este caso, la digitalización de la figura puede ayudar a determinar dónde y en qué medida difieren el modelo computarizado y la forma de la parte real. Estas diferencias pueden servir como guía para modificar el proceso de manufactura hasta que la parte sea aceptable.

2.3.3 Piezas museables.

Las piezas museables representan objetos únicos en su clase, que atraen el interés de científicos y público en general en todo el mundo. Tradicionalmente, para poder apreciar estos objetos, ha sido necesario visitar museos potencialmente distantes u obtener imágenes no interactivas o secuencias de video. Digitalizando estas piezas, los curadores de los museos pueden ponerlas a disposición para una visualización interactiva. Para los científicos, estos modelos computarizados ofrecen la oportunidad de estudiar y medir artefactos remotamente, usando poderosas herramientas de computación.

2.3.4 Medicina.

Las aplicaciones de la digitalización de formas 3-D en medicina, tienen un amplio alcance. Las prótesis pueden ser diseñadas a la medida, cuando las dimensiones de cada paciente son conocidas con alta precisión. Los cirujanos plásticos pueden usar la forma del rostro de un individuo para modelar el proceso de cicatrización del tejido y apreciar el resultado final de la cirugía. En el procedimiento de un tratamiento radioactivo, un modelo de la figura del paciente puede ayudar a guiar al médico a dirigir la radiación con precisión.

2.4 Adquisición.

Los métodos para digitalizar las formas de objetos tridimensionales han evolucionado rápidamente en las últimas décadas. La velocidad y precisión de las tecnologías de digitalización 3-D deben mucho a los avances en las áreas de ingeniería eléctrica y física, en lo relacionado con la disponibilidad de nuevos componentes, nuevas tecnologías y nuevos dispositivos, incluyendo el desarrollo de láseres, CCD's y circuitos de alta velocidad de muestreo. Los láseres son las fuentes más comunes para los sistemas digitalizadores. Son estables, monocromáticos, disponibles en un amplio rango de longitudes de onda, son potentes y operan con una alta eficiencia óptica [92, 103].

La digitalización de los objetos tridimensionales se efectúa mediante puntos o líneas. En el pasado esta tarea se realizaba a mano por un operario calificado, el cual realizaba mediciones sobre el modelo físico con un pie de rey. Un paso importante en la automatización de este proceso fue la aparición de las primeras máquinas por medición de coordenadas o CMM, con las que se podía obtener de una manera semiautomática una nube de puntos del objeto a digitalizar. Sin embargo, esta tarea era lenta y compleja, aunque supuso un gran avance.

Durante los 70's y principios de los 80's. la disponibilidad de luces de posición y otros dispositivos electro-ópticos, y la introducción de los computadores hizo posible el desarrollo de sensores de rango automáticos de bajo costo. Al principio, los dispositivos de posicionamiento eran casi todos análogos por naturaleza. Por ejemplo, el fotodiodo de efecto lateral (LEP), y las cámaras tipo Vidicon se usaban para convertir información óptica a una señal eléctrica para extraer información de rango. La precisión limitada hizo que estas tecnologías fueran difíciles de calibrar [89].

Con la aparición de sensores de luz, como el detector sensible a la posición (PSD) o el dispositivo de carga acoplada (CCD), se pensó que se podrían digitalizar volúmenes captando la luz reflejada por los objetos. De esta manera, se idearon tres técnicas de digitalización

conocidas como: punto a punto, por línea y por superficie. El problema fue que el procesamiento de las imágenes era complicado y la visión artificial aún se encontraba en un estado incipiente. Por aquel entonces, el láser ya se encontraba en desarrollo y a algunos científicos se les ocurrió aprovechar sus propiedades de coherencia espacial y temporal para emplearlo como instrumento de digitalización. El láser se empleaba como fuente de iluminación la cual es capaz de marcar una zona brillante de fuerte contraste y sencilla de procesar. La detección espacial de este punto por dos sensores planos era una tarea relativamente sencilla. Así, realizando un barrido de toda la superficie espacial del objeto se podrían recuperar sus coordenadas tridimensionales. En el caso de digitalización por superficie, se proyecta un mallado o rejilla láser habitualmente generado por un patrón de difracción o de forma dinámica [92].

Los sistemas comerciales de sensores de rango, muestran que la tecnología basada en 3-D ha pasado exitosamente la etapa inicial de curiosidad investigativa en los laboratorios de los 80's y las numerosas demostraciones de aplicaciones en los 90's. Poderosos algoritmos y software de procesamiento de datos han complementado gradualmente las técnicas sensoriales basadas en 3-D, probando que las aplicaciones dimensionales se desempeñan mejor usando imágenes 3-D, que usando imágenes 2-D. Sin embargo, los sistemas de visión 3-D están aún en sus etapas iniciales, cuando son comparados a sus contrapartes, en términos de calidad de imagen, renderizado y facilidad de uso. Generalmente, los métodos estándar para extraer datos de rango, son precisos sólo para objetos planos uniformes. Usando estos métodos, superficies curvas, superficies discontinuas y superficies de reflexión variante, causan distorsiones sistemáticas de los datos de rango.

Durante las últimas décadas, varios métodos para la adquisición de imágenes de rango fueron citados por Jarvis [61]. Besl [9] presentó una revisión de diferentes técnicas de medida de rango y sensores comerciales. También propuso una función que compara el desempeño de los sensores. Otros trabajos fueron presentados por Tiziani [105] y más recientemente por Chen [24] y Blais [12]. Estos trabajos describen este amplio tema desde diferentes perspectivas e inclusive, con una amplia lista de referencias que sólo tratan "*tip of the iceberg*".

Métodos de digitalización 3-D.

Las técnicas de visión 3-D, son clasificadas en métodos pasivos y activos. En una clasificación general, los métodos pasivos se limitan al estudio de la información visual de una o más vistas, tomadas en distintos instantes de tiempo y observando la escena, sin actuar sobre ella. Por el contrario en los métodos activos, existe retroalimentación de la información visual sobre:

- El movimiento, localización y orientación del sensor.

- Los lentes y el sistema de adquisición de imágenes (zoom, focos, iris, iluminación).
- La métrica y el estado interno del sistema (calibración visual); se suele trabajar generalmente modificando la iluminación.

La visión activa trata de imitar un sistema biológico, ya que se mejora el resultado de la percepción visual empleando mecanismos que reaccionan y controlan los sensores visuales y factores externos como la iluminación.

La clasificación en métodos activos y pasivos que se sigue en esta tesis, es la sugerida por Curless [35] y se complementa con la de Chi-Fang [27], Marshal [73] y González [49]. Es más sencilla, ya que no se tiene en cuenta si se modifican o no los parámetros internos de la cámara o la posición, sino simplemente, si se modifica la iluminación. De esta manera, se referirán como técnicas de visión activa aquellas en las que se actúa con algún tipo de energía externa sobre el objeto a medir, con el fin de capturar de forma más sencilla sus dimensiones y las técnicas de visión pasiva serán aquellas en las que no se emplea ninguna fuente de energía externa.

Los sensores de rango activo aprovechan una variedad de principios físicos (la Figura 2.4 resume los métodos de adquisición de rango activo). Entre las técnicas de detección más comunes están la triangulación, radar/sonar, interferometría moiré y enfoque/desenfoque activo. La triangulación usa un proyector de luz y una cámara de intensidad, que es colocada a cierta distancia del proyector. El proyector emite un patrón de luz. Los patrones más comunes son el plano y el haz de luz sencillo (ver Figura 2.5). El radar/sonar usa una onda electromagnética o acústica corta y detecta el retorno (eco) desde las superficies. La distancia es obtenida como una función del tiempo tomado entre el golpe de la onda con la superficie y su retorno. Los sensores Moiré proyectan dos rejillas con patrones espaciados regularmente sobre la superficie y miden las diferencias de fase del patrón de interferencia observado. Otros sensores de diferencia de fase miden el cambio de fase del haz de retorno observado, en ellos la distancia es una función de diferencia de fase. Los sensores de enfoque/desenfoque activo usan dos o más imágenes de la misma escena, que son adquiridas bajo situaciones de enfoque variantes, una vez el mejor enfoque de imagen es determinado, un modelo enlaza valores de enfoque y distancia, produciendo la profundidad.

2.5 Registro.

El registro es una etapa del proceso de reconstrucción 3-D, que consiste en emparentar dos o más imágenes, que pueden ser tomadas en diferentes momentos, desde diferentes sensores

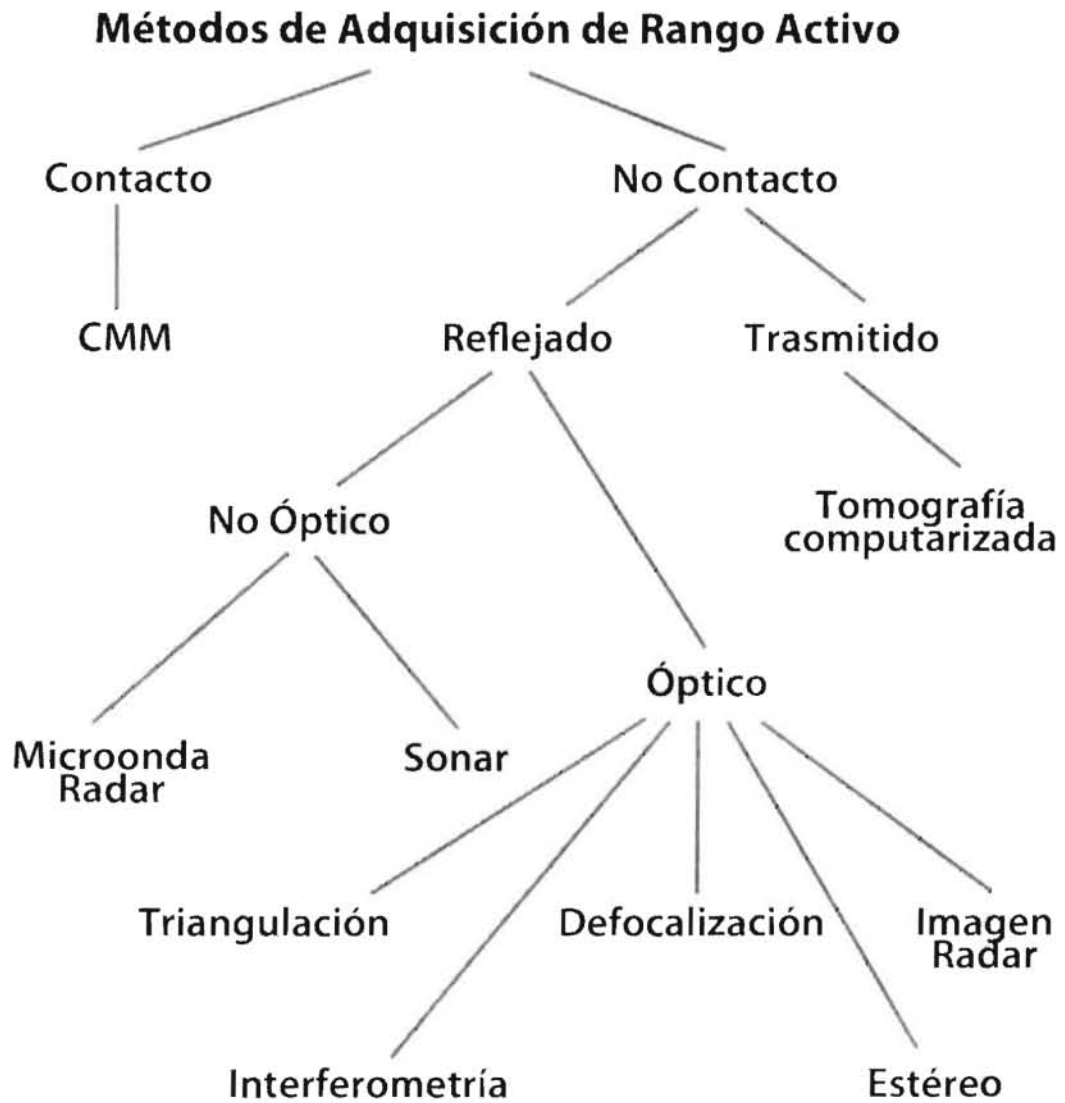


Figura 2.4: Taxonomía de los métodos de adquisición de rango activo [33].

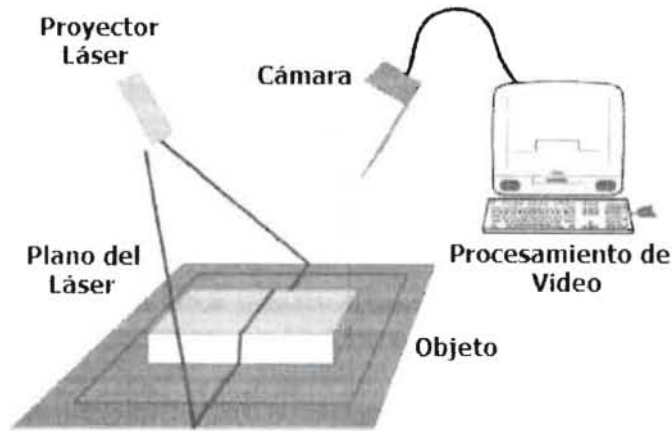


Figura 2.5: Triangulación óptica [12].

o desde diferentes puntos de vista. Dado que una simple vista de rango puede contener únicamente los puntos en la escena que son visibles al sensor, es necesario unir múltiples vistas de rango en un único conjunto de puntos, para obtener el modelo completo de la escena. El registro consiste en obtener una transformación (translación y rotación) que permita llevar un par de vistas de rango al mismo sistema de coordenadas.

Tomando la definición presentada por Blais y Levine, el registro de imágenes de rango consiste en:

“Dadas N vistas de un objeto en una escena, cada una describiendo la estructura 3-D del objeto como si fuera visto desde un punto de vista particular, deseamos encontrar N transformaciones rígidas de movimiento T_1, T_2, \dots, T_N , que especifiquen las verdaderas posiciones del sensor de rango con respecto a un único marco de referencia (elegido arbitrariamente y, usualmente, el marco de referencia de una de las vistas)” [13].

La desalineación que se produce inevitablemente al tomar dos o más imágenes de rango desde puntos de vista diferentes y sin ningún control de las posiciones relativas entre el objeto y el sensor, corresponde al tema central del registro. La información tridimensional capturada en cada imagen está referenciada al sistema de coordenadas de la cámara, el cual por regla general no corresponde al sistema coordinado del mundo. Una vez que la etapa de adquisición

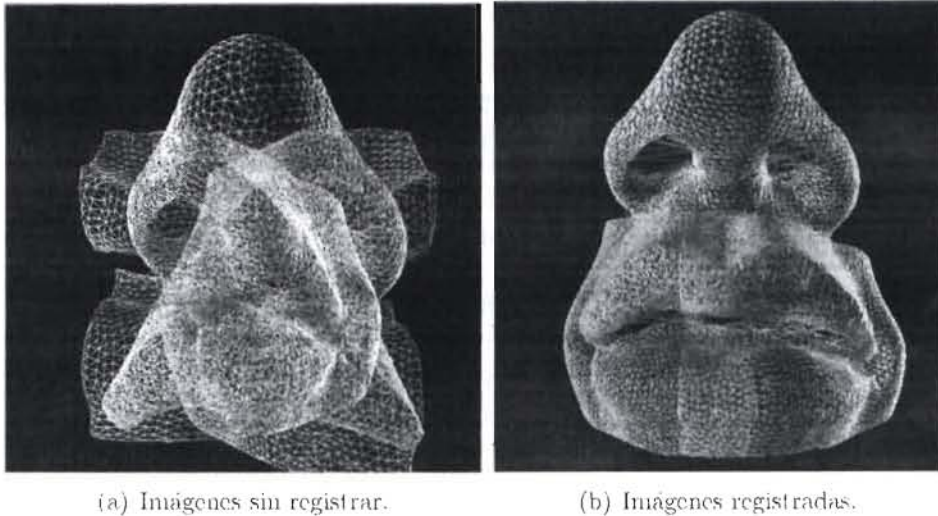


Figura 2.6: Registro de imágenes de rango.

culmina con el sensado de la superficie del objeto y se realiza la visualización de las imágenes en un sistema de coordenadas único, se obtienen diferentes vistas que se superponen unas a otras, sin coherencia alguna. El propósito del registro es alinear estas vistas de tal manera que la forma del objeto sea recuperado con la mejor precisión posible (ver Figura 2.6).

En el esfuerzo por ajustar las imágenes de manera coherente se hacen evidentes dos situaciones: en la primera, no se tiene certeza de que puntos en el sistema coordenado de una imagen, se corresponde con sus equivalentes en el sistema coordenado de la otra. Este hecho, es conocido como el problema de la correspondencia y constituye la etapa de mayor consumo de tiempo de ejecución del algoritmo. En la segunda, se requiere una transformación sobre la información tridimensional, de una de las imágenes en su sistema coordenado, al sistema de la imagen que se eligió como referencia, con el propósito de ajustar ambas imágenes utilizando la información común entre ellas. Dada la naturaleza inexacta de los datos y la incertidumbre en el conocimiento de las superficies comunes, el procedimiento de cálculo de esta transformada es iterativo, guiado por las estrategias y métricas elegidas por el investigador.

Históricamente, la correspondencia de objetos de forma libre usando datos 3-D, fue realizada en un principio por Faugeras y su grupo en INRIA [46], donde demostraron una correspondencia efectiva con el timón de un Renault a principios de la década de los 80's. Este trabajo popularizó el uso de cuaterniones para registro de puntos correspondientes mediante mínimos cuadrados. El uso alternativo del algoritmo de descomposición en valores singulares

(SVD), no fue ampliamente conocido en ese tiempo. La limitación primaria de este trabajo fue que partía de la existencia probable de regiones planas razonablemente grandes al interior de una forma libre. Numerosos trabajos posteriores como los de Schwartz [97], Haralick [53], Horn [59], Brou [19] y Taubin [104], constituyen aproximaciones interesantes a la solución del problema de registro, sin embargo, estos trabajos poseen limitaciones tales como la correspondencia en modelos con altas deformaciones, el alto costo de emparejamiento local para porcentajes de oclusión pequeños, entre otras. En las últimas décadas, se han propuesto varias aproximaciones al registro. Estas pueden ser clasificadas básicamente en dos grupos: el registro grueso y el registro fino.

En el registro grueso, la meta es encontrar un conjunto aproximado de transformaciones de registro, sin previo conocimiento de las posiciones espaciales relativas a las vistas. Dentro de los trabajos realizados en el registro grueso se encuentra el de Sappa *et al.* [93], que presentaron un método el cual emplea una técnica de segmentación basada en bordes para guiar el proceso de registro. Chua y Jarvis [29], usaron curvaturas principales y marcos Darboux para calcular características invariantes. También, Feldmar y Ayache [47] propusieron un método para estimar desplazamientos rígidos usando curvaturas principales de superficies.

Stein y Medioni [102] propusieron la estructura Splash, la cual es un mapeo que describe las distribuciones de superficies normales a lo largo de un círculo geodésico. La imagen spin presentada por Johnson y Hebert [62], la cual es un descriptor de forma a nivel de datos, ha sido usada en registro. Lucchese *et al.* [72], explotan la regularidad geométrica obtenida por la transformada de Fourier como un método basado en un dominio de frecuencias para el registro de imágenes de rango. El método DARCES basado en RANSAC de Chen *et al.* [23], es un método robusto basado en la búsqueda exhaustiva que puede chequear todas las alineaciones de datos posibles entre dos conjuntos de datos, para registrar dos vistas parcialmente solapadas.

En contraste a los métodos de registro grueso, los métodos de registro fino se basan en la suposición de que se obtuvo inicialmente una buena transformación inicial. Entonces, las alineaciones precisas se pueden obtener con criterios confiables para medir la calidad de las transformaciones refinadas. Adicionalmente, varias aproximaciones en la literatura proponen el empleo de una combinación de ambas técnicas: un registro grueso, seguido de un registro fino para lograr resultados de registros precisos y automáticos.

Los mejores métodos que se conocen para registro fino de imágenes son las variantes del algoritmo de iteración de punto más cercano (ICP). ICP es un procedimiento iterativo que minimiza cada vez el error cuadrado medio, calculado como la suma de las distancias entre puntos en una vista y los puntos más cercanos en la otra vista. Este método fue propuesto por Besl y McKay [11] y constituye el método clásico para abordar el problema del registro.