

Proyecto Fin de Máster
Máster en Ingeniería de Telecomunicación

Técnicas de reconstrucción en tres dimensiones y
realización de un programa informático para el
trabajo con los modelos

Autor: David Morales Asegurado

Tutoras: Begoña Acha Piñero - Carmen Serrano Gotarredona

Dep. Teoría de la Señal y Comunicaciones
Escuela Técnica Superior de Ingeniería
Universidad de Sevilla

Sevilla, 2016



Proyecto Fin de Máster
Máster en Ingeniería de Telecomunicación

Técnicas de reconstrucción en tres dimensiones y realización de un programa informático para el trabajo con los modelos

Autor:

David Morales Asegurado

Tutoras:

Begoña Acha Piñero - Carmen Serrano Gotarredona

Profesoras titulares

Dep. de Teoría de la Señal y Comunicaciones

Escuela Técnica Superior de Ingeniería

Universidad de Sevilla

Sevilla, 2016

Proyecto Fin de Máster: Técnicas de reconstrucción en tres dimensiones y realización de un programa informático para el trabajo con los modelos

Autor: David Morales Asegurado

Tutoras: Begoña Acha Piñero – Carmen Serrano Gotarredona

El tribunal nombrado para juzgar el Proyecto arriba indicado, compuesto por los siguientes miembros:

Presidente:

Vocales:

Secretario:

Acuerdan otorgarle la calificación de:

Sevilla, 2016

El Secretario del Tribunal

*Dedicado a todas aquellas
personas sin las que mi vida
académica no hubiera sido
posible – Familia, amigos y
docentes*

Resumen

En este proyecto se realiza la puesta en marcha de dos sistemas de reconstrucción de modelos en tres dimensiones: el primero basado en fotografías convencionales y el segundo realizado mediante un escáner de superficie. La motivación original del trabajo se encontraba en la aplicación de estas tecnologías para la reconstrucción de mamas para fines médicos, fin que fue descartado una vez analizadas las características y limitaciones de la tecnología disponible en la actualidad al alcance del mundo universitario.

El proyecto se divide en dos partes principales. La primera trata de una revisión del estado del arte de las técnicas disponibles para la reconstrucción de objetos en tres dimensiones mediante fotografías convencionales, sin referencias numéricas de la localización de la cámara. Se presentan los métodos analizados y se exponen el conocimiento y tecnologías aprendidas, eligiendo una de ellas para su análisis. Posteriormente se explica el método de reconstrucción 3D a partir de un escáner de superficie, paradigma completamente diferente al anterior pero de gran repercusión y uso en la actualidad.

En la segunda parte se ha realizado un programa que permite analizar modelos genéricos de reconstrucciones en tres dimensiones en un formato concreto. En este programa original creado para este trabajo se pone en práctica todo lo aprendido en las secciones anteriores. Con él se pueden realizar numerosas operaciones de tratamiento de modelos en 3D. Se termina el documento con las conclusiones aprendidas.

Resumen	9
Índice	11
Índice de Imágenes	13
1 Análisis de la Tecnología Elección de Técnica SFM	15
1.1 ANÁLISIS DEL ESTADO DEL ARTE	15
1.1.1 ALGORITMOS BASADOS EN CARACTERÍSTICAS MORFOLÓGICAS	16
1.1.2 ALGORITMOS BASADOS EN ESTRUCTURA A PARTIR DE MOVIMIENTO	16
1.2 ELECCIÓN DE SFM. EXPLICACIÓN	16
1.2.1 CÓMO SE LOGRA LA RECONSTRUCCIÓN	17
1.2.2 CONDICIONES DESEABLES DE LAS FOTOGRAFÍAS	20
1.2.3 VENTAJAS E INCONVENIENTES DE LOS MÉTODOS SFM	22
1.3 MÉTODO	22
1.3.1 EQUIPOS UTILIZADOS	22
1.3.2 DECISIÓN ENTRE FOTOGRAFÍAS Y FOTOGRAMAS DE UN VÍDEO	23
1.3.3 TIPOS DE OBJETOS, JUSTIFICACIÓN DEL LUGAR DE CAPTURA Y CONFIGURACIÓN DE LA CÁMARA	24
1.3.4 MOTIVO DE LA ELECCIÓN DE VISUALSFM. COMBINACIÓN CON ALGORITMOS MVS	25
1.4 EJEMPLOS	26
2 Reconstrucción mediante un Escáner de Superficie	31
2.1 ESPECIFICACIONES DEL ESCÁNER	33
2.2 MÉTODO	33
2.2.1 CALIBRACIÓN	34
2.2.2 ESCANEADO DE LOS OBJETOS	35
2.2.3 FICHERO SALIDA	36
2.3 CONJUNTOS MUESTRA	36
3 Programa de Análisis	39
3.1 CARGA DEL FICHERO	40
3.2 CONSTRUIR MALLA UNIFORME	40
3.3 VER SECCIÓN / VER SECCIÓN RELLENA	41
3.4 LIMPIEZA DEL MODELO	44
3.5 HOJA DE DATOS DEL MODELO	46
3.6 COMPARACIÓN	47
3.6.1 DIFERENTES REFERENCIAS	47
3.6.2 DIFERENTES RESOLUCIONES	47
3.6.3 RESULTADO	48
3.7 COMENTARIO FINAL	49
4 Resultados de la Comparativa con el Conjunto de Ejemplo	51
4.1 TOLERANCIA	53
5 Conclusiones	55
6 Glosario	57
Referencias	59
7 Aviso Legal	63

Índice de Imágenes

Ilustración 1: Ejemplo de deslumbramiento	20
Ilustración 2: Ejemplo de mala iluminación.	21
Ilustración 3: Foto desenfocada vs Foto enfocada en el objeto.	21
Ilustración 4: Reconstrucción de Puerta lateral de la iglesia de San Dionisio en Jerez de la Frontera (Cádiz).	27
Ilustración 5: Reconstrucción de un limón.	28
Ilustración 6: Varias reconstrucciones	29
Ilustración 7: Reconstrucciones de un molinillo y de una plancha antigua	30
Ilustración 8: Escenario de escaneo	32
Ilustración 9: Escáner con patrón de calibración	34
Ilustración 10: Escaneos de los objetos de prueba	35
Ilustración 11: Libro en perspectiva y planta	36
Ilustración 12: Plancha en perspectiva y alzado	37
Ilustración 13: Quinqué en perspectiva y alzado	37
Ilustración 14: Capura de reconstrucción ajustada a malla uniforme, con intervalo de paso entre posiciones de la malla de magnitud 3 y factor de aumento de resolución 1	40
Ilustración 15: Detalle de reconstrucción ajustada a la malla uniforme	41
Ilustración 16: Sección sin rellenar. Plano 'x', posición 0	42
Ilustración 17: Sección rellena en horizontal (rellenado en dirección correcta). Plano 'x', posición 0	43
Ilustración 18: Sección rellena en vertical (rellenado en dirección incorrecta). Plano 'x', posición 0	43
Ilustración 19: Antes y después de la limpieza de ruido (las perspectivas pueden variar ligeramente)	45
Ilustración 20: Superposición de los dos modelos generados (escáner 3D y SFM-MVS), tras ejecutar el algoritmo de ajuste de referencias	49

1 ANÁLISIS DE LA TECNOLOGÍA ELECCIÓN DE TÉCNICA SFM

Durante el transcurso de este apartado se presentan numerosas siglas y conceptos. Algunas de ellas están comentadas en el mismo texto para dar más coherencia al resultado expresivo final. Si se encuentra con alguna no explicada, consulte el apartado Glosario al final de la memoria, donde encontrará una explicación de los términos clave empleados.

1.1 ANÁLISIS DEL ESTADO DEL ARTE

Como paso previo a cualquier elección de una tecnología, y como buena práctica a la hora de empezar a trabajar en un tema, se hace necesario realizar un análisis del estado del arte de las tecnologías existentes. Para este apartado se ha consultado numerosa bibliografía de todo tipo, especialmente ha sido de gran ayuda el artículo de Paul Bourke [1].

En el tema que nos ocupa, el de la reconstrucción de modelos en tres dimensiones a partir de fotografías convencionales –sin referencias numéricas de la localización de la cámara– existen dos grandes grupos de técnicas: las que emplean una sólo imagen para ello, y las que utilizan varias de ellas.

En una primera aproximación puede parecer que existen multitud de técnicas para estos propósitos. Sin embargo, tras una posterior investigación acerca de los métodos disponibles, aparece la realidad: la gran mayoría de software de “primer nivel” disponible en la actualidad es de pago, en ocasiones su uso supone un desembolso elevado de dinero, inasumible para el desarrollo de este trabajo.

Por tanto, la búsqueda se centrará en los algoritmos gratuitos en el momento del desarrollo del trabajo. Aunque estos presenten varios nombres comerciales, la realidad que se encuentra al explorar más en profundidad es que la práctica mayoría se engloban en dos grupos:

1.1.1 ALGORITMOS BASADOS EN CARACTERÍSTICAS MORFOLÓGICAS

En este grupo se encuentran los algoritmos que generalmente emplean una sola imagen para la reconstrucción. Se basan en la búsqueda de elementos morfológicos en la fotografía que suministren información tridimensional. Ejemplos de estos indicadores son: pares de líneas oblicuas que indican perspectiva, gradientes de color, perspectiva en las texturas, etcétera.

Para el trabajo se consideró en concreto el programa Make3D [2], en su variedad para funcionar con una sola fotografía. En los resultados de las pruebas realizadas, aunque sí es cierto que conseguía generar una reconstrucción en tres dimensiones, éstos no fueron a nuestro juicio lo suficientemente satisfactorios para ser empleados en modelos relativamente complejos. Esto, unido a las posibles restricciones de licencia aplicables sobre las imágenes / modelos, descartaron su uso.

1.1.2 ALGORITMOS BASADOS EN ESTRUCTURA A PARTIR DE MOVIMIENTO

Se conoce como técnicas de Estructura a partir de movimiento (*Structure From Motion o SFM*) a aquellas que realizan una reconstrucción en tres dimensiones de un escenario a partir de un conjunto de fotografías.

Analizando sintácticamente el significado literal de la expresión: “Estructura a partir de movimiento” se puede entender su base de funcionamiento. Si un observador es capaz de ver un objeto desde todas las perspectivas de tal forma que pueda apreciar todas sus características constituyentes, podrá reconstruirlo de forma aproximada. Expresado de otra forma, si se dispone de información del objeto desde todas las perspectivas, y además se tiene cómo relacionar esa información entre sí para formar un todo, se podrá intentar reconstruir el objeto.

La mayoría de tecnologías encontradas se engloban en este apartado. En concreto, el software VisualSFM [3], que será empleado posteriormente en el trabajo entraría en este subtipo de técnicas, aunque posee más funcionalidad que la puramente SFM.

1.2 ELECCIÓN DE SFM. EXPLICACIÓN

Finalmente, la tecnología escogida para ser analizada es la de Structure From Motion (SFM). En la revisión del estado del arte realizada como fase constituyente del trabajo, esta se presenta como una de las principales en cuanto a la reconstrucción de objetos. Se presentarán sus peculiaridades, sus pros y sus contras, que serán analizados en las siguientes páginas. Antes, es importante explicar cómo funcionan, aunque bien es cierto que cada software particular de este tipo de técnicas puede diferir en la implementación.

1.2.1 CÓMO SE LOGRA LA RECONSTRUCCIÓN

Para poder abordar el tema que sigue, conviene tener muy claro el principio de funcionamiento de las técnicas de Structure From Motion para poder entender los requisitos aplicables. El sistema se puede ver como el problema de la caja negra en ingeniería. Se tienen entradas a un sistema, en este caso el conjunto de fotografías; y a la salida aparece un modelo en tres dimensiones del objeto que representan las imágenes. ¿Qué ha pasado en medio?



Lógicamente, cada programa es diferente, y cada cuál tendrá sus algoritmos propios. Sin embargo, se pueden extraer unos grandes subbloques en este proceso (referencias [4] a [20]) (ver avisos legales).

A diferencia de otros métodos, en los que se hace necesario conocer datos sobre la posición de captura de las cámaras, en los algoritmos de este tipo se parte de fotografías estándar, en las que no se tiene etiquetada de antemano la posición de la cámara y la distancia al objeto en el momento de la captura. En este caso, se dice que la cámara no está calibrada o no tiene ajustados los parámetros. Hará falta iniciar un proceso que calcule a partir de la imagen capturada esta información faltante. Este proceso se denomina calibración de cámara. Aparte de conocer la ubicación, es posible que sea conveniente disponer de datos que sirvan para corregir las imperfecciones en la captura (como por ejemplo, la distorsión de la lente).

En algunas ocasiones, se dispone de ayuda para este proceso de calibración. La mayoría de cámaras digitales exportan algunas características físicas sobre el momento de captura, y las guardan en el mismo archivo de la fotografía. Por ejemplo, algunos algoritmos de calibración de cámara emplean la información almacenada en EXIF [18][19][20]. Este formato permite almacenar los llamados metadatos, que no son más que información adicional, que no forma parte en sí de los datos del archivo, pero que lo complementan. Un ejemplo de metadato en imágenes es la distancia focal.

La distancia focal se define como la separación, expresada normalmente en milímetros, que hay entre el centro óptico del objetivo y el plano de formación de la imagen (que en una cámara digital es el sensor de imagen). A distancias focales cortas, el ángulo de apertura será mayor y se capturará más escena. Por el contrario, las distancias focales largas provocarán que una parte menor de la escena se capture, con lo que parecerá que el objeto está aparentemente más cerca de la cámara.

Más interesante aún es la distancia focal equivalente en 35mm. Cada objetivo tiene unas características, pero poder expresar la distancia focal en un marco de referencia común permite tener un indicativo de la distancia al objeto. Este marco de referencia común es el formato de película de 35mm, muy conocido en fotografía y cine.

$$d_{35mm} = K \cdot d_{objetivo}$$

Se puede decir que, para cada objetivo, se tiene un denominado factor multiplicador de la distancia focal. Este valor, representado como K en la fórmula anterior, permite transformar de la distancia focal propia del objetivo a la distancia focal equivalente en 35mm con sólo realizar una multiplicación o división. Éste es un tipo de dato que se puede aprovechar para el proceso de calibración de cámara.

Dejando a un lado la calibración, se debe extraer información de las fotografías propiamente dichas. Lógicamente, cada imagen dará una perspectiva del objeto diferente, luego habrá que buscar características reconocibles desde diferentes perspectivas. Éstos son los denominados puntos singulares. Existen diferentes técnicas para ello, en las que a cada punto singular se le asocia una serie de características que no dependen del lugar de la fotografía. La técnica más conocida es SIFT (Scale-Invariant Feature Transform) [21]. Este algoritmo toma su nombre debido a que las características obtenidas de cada punto singular serán un conjunto de valores que no dependen de la escala (Scale-Invariant) ni de la orientación, para así poder identificar los mismos puntos en diferentes fotografías tomadas desde perspectivas distintas. Por cada imagen, se obtendrá un listado de puntos singulares, llamados así por tener características reconocibles. Por cada punto singular se tienen las características que se le han extraído mediante SIFT.

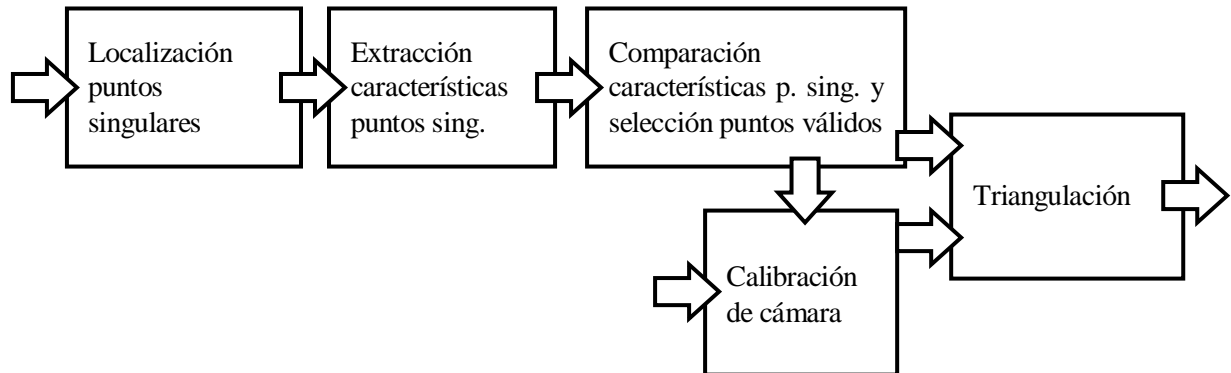
Una vez se tienen las características de los puntos singulares de cada una de las fotografías, habrá que compararlos con los del resto de ellas. Observando los valores de diferentes puntos singulares en diferentes fotografías, si se encuentran dos con características extraídas con el algoritmo SIFT que los haga candidatos a tratarse del mismo lugar físico, cada uno en una imagen diferente, se puede considerar (en principio) que estos dos puntos podrían representar una misma posición física en la escena de captura.

Con la información correspondiente a muchos puntos singulares se puede obtener también la localización angular de la cámara con respecto a la escena. Ya se tiene por tanto una estimación de la posición física de la cámara. Por triangulación, si un mismo punto es visto desde diferentes perspectivas espaciales, se puede aproximar su ubicación en un espacio tridimensional.

Este es un problema típicamente sobredimensionado. Se tiene más información de la necesaria para resolver el sistema matemático. Habrá que abordar este problema con algún método, por ejemplo, el de mínimos cuadrados. Como resultado, se tendrá la localización en el espacio tridimensional de ese punto.

Con todos los puntos posibles ubicados en el volumen, se construye una nube de puntos que representa al objeto fotografiado. Ésta nube es, por tanto, la reconstrucción SFM del objeto.

A continuación se presenta un diagrama de flujo resumen del proceso completo:



Al terminar el proceso de triangulación se tendrá un modelo tridimensional reconstruido del objeto fotografiado. Sin embargo, el proceso puede no acabar aquí; en algunos programas se incorpora una cadena posterior de tratamiento de este modelo, incluso puede que la salida de SFM se dé como entrada a otros métodos posteriores que realicen otra reconstrucción. Aquí se ha descrito el procedimiento básico de funcionamiento de SFM.

En este proceso pueden ocurrir varios problemas. Dos de los más frecuentes son:

- Modelos reconstruidos con pocos puntos: si las fotografías y/o los objetos no cumplen unos determinados requisitos, es posible que el resultado no tenga la densidad de puntos necesaria para que se vea como un volumen. Generalmente esto se verá como un conjunto de puntos en el espacio tridimensional que pueden o no corresponder al objeto fotografiado, pero con los que no se representan las características de éste.
- Producción de varios modelos. Si hay regiones del objeto de las que no se ha podido extraer la información y otras sí, es posible que se produzcan varios modelos separados; esto se debe a que no existen puntos en común suficientes entre las diferentes perspectivas. Dicho de otra forma, no hay suficiente información para unir las partes que sí se han podido reconstruir por separado.

Para evitar estos problemas, a continuación se expresan una serie de condiciones recomendables.

1.2.2 CONDICIONES DESEABLES DE LAS FOTOGRAFÍAS

En primer lugar, hay que seleccionar las fotografías adecuadas para darlas como entrada. Si alguna fotografía ha salido borrosa, desenfocada, movida o con algún otro problema es conveniente descartarla. Introducir fotografías inadecuadas puede provocar problemas en la reconstrucción, como por ejemplo, la creación de modelos alternativos o la no detección de puntos.

Por otro lado, a la hora de capturar las imágenes es importante que se tenga una buena iluminación y poner un valor de sensibilidad ISO a la cámara bajo para eliminar el ruido del sensor de imagen, especialmente si se usa un teléfono móvil para realizar las fotografías. La definición de la imagen es también importante, aunque en la experiencia, valores muy elevados de resolución resultan contraproducentes por el aumento de tiempo para procesarlas.

En el trabajo, con el dispositivo de captura utilizado (ver apartado 1.3.1 para más información), valores de sensibilidad ISO de 100 y resolución de 8 Megapíxeles han dado buenos resultados.

Las condiciones específicas que son necesarias para una buena reconstrucción, basadas en la bibliografía y en la experiencia tras la realización de este trabajo, son:

- Número suficientemente elevado de fotografías. Dependiendo del modelo, del ángulo a cubrir y de la complejidad del objeto capturado serán necesarias más o menos fotografías. En las pruebas realizadas, y para modelos reales (no sintéticos), son necesarias del orden de varias decenas de fotografías. Al aumentar el número de fotografías, normalmente aumentará el espacio necesario para almacenarlas y procesarlas informáticamente, por lo que una relación de compromiso en este aspecto es importante.
- Iluminación de calidad. Uno de los aspectos más importantes a la hora de capturar una escena (ver Ilustración 2). Fotografías con poco contraste o en las que no se puedan apreciar con nitidez los detalles del objeto pueden llegar a ser incluso contraproducentes. La iluminación es muy importante también; en las pruebas realizadas los mejores resultados se han obtenido con luz natural en exteriores.
- Deslumbramientos. Es un subtipo de “mala iluminación”, pero se considera de suficiente entidad como para comentarlo independientemente. Al capturar imágenes en distintos ángulos es posible que en algunos de ellos se tengan deslumbramientos por estar a contraluz. Es una situación que se debe evitar, no sólo hará que haya partes del objeto que no se vean, sino que puede provocar posiciones erróneas a la hora de interpretar las imágenes para reconstruir el modelo en tres dimensiones.

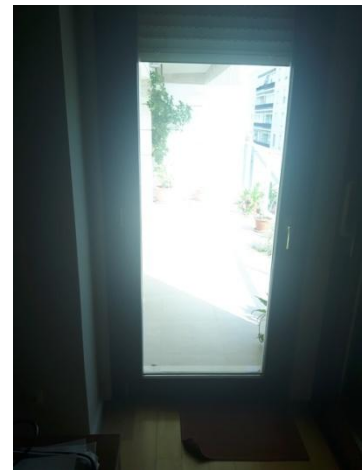


Ilustración 1: Ejemplo de deslumbramiento



Ilustración 2: Ejemplo de mala iluminación.

Detalle del ruido e imprecisión (retocada informáticamente para aumentar el brillo y el contraste)

- Los objetos tienen que tener detalles claros y distinguibles. Si un objeto es demasiado simple, liso y sin texturas, es probable que no se encuentren suficientes características con las que enlazar las fotografías entre sí, no hay “referencias” comunes para que los algoritmos entiendan que se trata de un único objeto. Ejemplos de este tipo pueden ser una pared monocromática o un objeto con un color homogéneo. Por supuesto, los objetos que sean transparentes confunden al método por mostrar lugares que deberían estar ocultos con ese ángulo de fotografía.
- Objetos brillantes. Los brillos y reflejos del entorno de un objeto pueden provocar que un mismo lugar se vea diferente desde lugares distintos, por lo que no se podría comparar bien las fotografías.

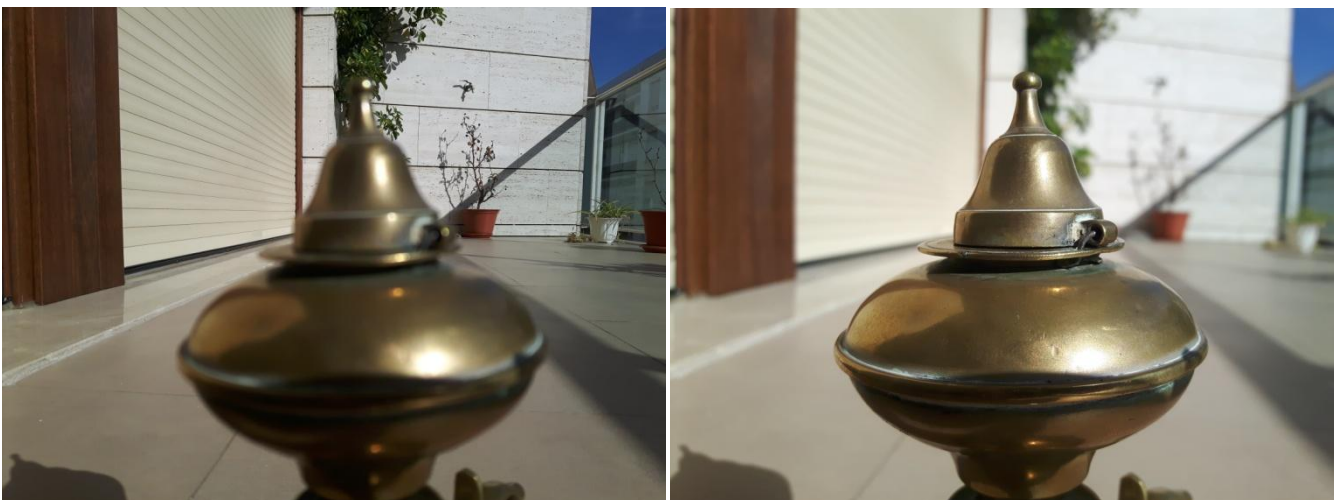


Ilustración 3: Foto desenfocada vs Foto enfocada en el objeto.

En la fotografía con el enfoque puesto en el fondo no se conseguirá obtener demasiada información del objeto de interés a capturar (en este caso el quinqué)

1.2.3 VENTAJAS E INCONVENIENTES DE LOS MÉTODOS SFM

Se han analizado en sí los requisitos que tienen que cumplir las fotografías. Yendo ya al análisis de los métodos SFM disponibles, se pueden identificar varias ventajas e inconvenientes:

Las ventajas de esta técnica son:

- ✓ Ampliamente extendida. Numerosos tutoriales de funcionamiento.
- ✓ Permite captura de la información de color.
- ✓ Con un número suficientemente elevado de fotografías puede tener bastante calidad, especialmente si se combina con otros métodos posteriores.

Por el contrario, las desventajas son:

- × Las fotografías y los objetos tienen que cumplir unos requisitos (ver listado anterior).
- × Es necesario un número elevado de fotografías (normalmente >25).
- × Tiempo de reconstrucción elevado. Requisitos informáticos potentes.

1.3 MÉTODO

1.3.1 EQUIPOS UTILIZADOS

Para la captura de las fotografías se analizaron dos alternativas. O bien emplear una antigua cámara digital de fotografía de 3,2 Mpx o usar un teléfono móvil con una resolución de 16 Megapíxeles, aunque capturando principalmente a 8 Mpx por los motivos expuestos en el apartado anterior.

El teléfono móvil fue elegido como cámara porque proporcionaba una buena calidad de imagen, suficiente resolución y buenos requisitos de movilidad y transporte, aunque con problemas como la ausencia de zoom óptico.

Los equipos en los que se puso en marcha el programa fueron dos diferentes. Uno de ellos fue un equipo portátil con procesador Intel i3 (primera generación) de bajo consumo y gráfica ATI-AMD. El segundo equipo fue un ordenador de sobremesa con procesador AMD Athlon II X2 y gráficos NVIDIA. En ambos el comportamiento fue muy parecido. Al ser ordenadores de una cierta antigüedad (alrededor de 5 años), el rendimiento experimentado fue bastante pobre.

Al intentar reconstruir modelos a partir de muchas fotografías (>25) el proceso se hace muy largo, llegando en algunas ocasiones a las 2 horas de duración. En las pruebas, el consumo de CPU fue de un 100%, excepto en aquellos momentos en los que se agotaba la memoria física, en los cuales había otros cuellos de botella diferentes a la CPU (por ejemplo, el disco duro al ser usado como memoria principal al agotarse la RAM).

1.3.2 DECISIÓN ENTRE FOTOGRAFÍAS Y FOTOGRAMAS DE UN VÍDEO

Llegados a este punto se debe tomar una decisión: o bien hacer un vídeo que capture el objeto entero o, por el contrario, hacer un número elevado de fotografías desde todas las direcciones. Visto como una primera aproximación técnica la idea es parecida, puesto que un vídeo es básicamente un conjunto de imágenes tomadas en movimiento. Sin embargo, hay pequeñas diferencias entre ambos, en especial relativas a la captura y compresión de la información.

Tras realizar varias pruebas con vídeos se presenta el problema que lleva a la decisión de usar fotografías en este trabajo: las imágenes en movimiento obtenidas a partir de los medios que están al alcance no proporcionan una calidad comparable a la de las fotografías. Esto se hace especialmente patente si la escena tiene poca luz y se intenta hacer una captura en vídeo relativamente rápida: para poder obtener la escena la cámara tendrá que realizar un tiempo de exposición más prolongado; esto unido al movimiento conlleva un efecto indeseado que se hizo patente en las pruebas con vídeo: las imágenes movidas o borrosas.

Si el vídeo se observa en movimiento, estos problemas quedan camuflados, incluso aportan a la escena una sensación de naturalidad –realizar en un vídeo tiempos de exposición demasiado rápidos implican que la imagen se ve con aspecto entrecortado-. Sin embargo, cuando se intentan extraer fotogramas de la secuencia los resultados no suelen ser satisfactorios.

Si a todo lo anterior se le añade el uso de un algoritmo de compresión de vídeo, los problemas se agravan aún más. Los algoritmos de compresión con pérdidas emplean técnicas para reducir el tamaño del fichero de vídeo a costa de perder información poco apreciable para el sistema de visión humano cuando el vídeo es visto como sucesiones rápidas de imágenes. Sin embargo, esta pequeña pérdida –o grande, según como se comprima- puede suponer un efecto muy negativo para el fin que se persigue, que es el de extraer imágenes fijas (fotogramas del vídeo).

Por tanto, queda justificado el uso de fotografías estáticas para este trabajo. En las fotografías también están presentes algunos de los problemas comentados anteriormente, en concreto cabe destacar el problema de la compresión. Si se le aplica una compresión con pérdidas a las imágenes para reducir el espacio que ocupan en disco, los resultados que se obtendrán serán generalmente peores. Además, la información que se pierde en la compresión ya no se podrá recuperar. De forma similar a la compresión en vídeo, el algoritmo comprimirá para que la vista humana no aprecie demasiado la pérdida, pero el algoritmo de reconstrucción puede necesitar esa información perdida.

Aparte de lo expresado anteriormente, las fotografías proporcionarán la información necesaria si cumplen unas determinadas condiciones específicas para SFM, ya discutidas con anterioridad. Posteriormente se analizarán aspectos técnicos avanzados más allá de esta dicotomía vídeo-fotografías.

1.3.3 TIPOS DE OBJETOS, JUSTIFICACIÓN DEL LUGAR DE CAPTURA Y CONFIGURACIÓN DE LA CÁMARA

Llegados a este punto se hace necesario distinguir dos tipos de objetos a capturar:

El primer grupo de objetos se pueden clasificar como inmóviles. En este conjunto se encuentran las fachadas de edificios, aceras, etcétera. Se encuentran en la calle y no se puede hacer nada por trasladarlos a un lugar adecuado para la captura. Lo único controlable en esta situación son las condiciones de iluminación, por ello estos objetos son fotografiados con buena luz solar. Como último detalle que comentar, se ha tenido especial cuidado en que no aparezcan personas en las fotografías, con un doble sentido: el respeto a la imagen e intimidad, y el de crear un escenario lo más estático posible (si las personas se mueven se puede confundir el algoritmo).

En el segundo grupo se encuentran los objetos transportables. Éstos suelen ser pequeñas piezas o aparatos que se pueden mover con facilidad de su posición original, que puede que no sea la adecuada para la captura, a una que cumpla los requisitos. Lo ideal sería disponer de una sala especial, preparada a modo de estudio de fotografía, en la que la iluminación esté controlada y en la que la cámara tenga libertad de movimiento para capturar el objeto desde las posiciones deseadas. En el proyecto se intentó usar las instalaciones y facilidades de fotógrafos profesionales, pero por su alto coste se descartó. Por tanto, sólo quedaba la posibilidad de buscar un lugar adecuado para la captura en el que hubiera buenas condiciones de iluminación y se pudiera fotografiar el objeto desde el mayor rango de ángulos posibles. Se escogió una azotea de un edificio, a horas del día que no produjeran sombras excesivas ni desvirtuantes de las características del objeto. Se optó por colocar los objetos en una superficie elevada con respecto al suelo para disponer de un ángulo de fotografía más cómodo, tomando todas las precauciones necesarias.

El dispositivo de captura se configuró, poniendo la resolución a 8 Megapíxeles (ver apartados anteriores para justificación) y se dejaron el resto de ajustes en modo automático, sólo teniendo que configurar manualmente –y en contadas ocasiones- el enfoque al objeto de interés.

El tiempo medio para realizar el proceso completo varía según el objeto, pero tomando una media se puede decir que, una vez localizado y preparado el escenario, para realizar las fotografías basta con unos 15 minutos por objeto. El tiempo destinado a la reconstrucción informática del objeto en tres dimensiones se ha discutido en apartados anteriores.

1.3.4 MOTIVO DE LA ELECCIÓN DE VISUALSFM. COMBINACIÓN CON ALGORITMOS MVS

En el universo Structure From Motion existen muchos programas y métodos que persiguen el mismo fin, el de reconstruir un modelo tridimensional a partir de fotografías; desde programas gratuitos hasta algunos de pago de alto coste.

Entre todas las técnicas consideradas como candidatas en el análisis del estado del arte, este programa se presentó como uno de los mejores posicionados en la gama de los programas gratuitos en el momento de desarrollo del trabajo. Esto, unido a una relativa alta disponibilidad de la documentación, fue la clave de la decisión para elegir el programa.

En el momento de realización del trabajo, el programa está disponible en la página de su creador [3], y no necesita instalación para su funcionamiento, aunque sí puede requerir la descarga de otros componentes desde páginas externas para realizar algunas funciones.

1.3.4.1 INTEGRACIÓN CON HERRAMIENTAS MVS

VisualSFM, además de ser un programa basado en SFM, integra en la misma interfaz soporte para otras herramientas de reconstrucción avanzadas que toman como entrada el modelo SFM. Estas técnicas son del tipo *MultiView Stereo* (MVS), y las integrables en VisualSFM son llamadas como PMVS2 [22] y CMVS [23].

Los algoritmos MVS permiten explotar las características geométricas de la visión estereoscópica para conseguir las reconstrucciones. Por eso, se hace necesario en este caso darles como entrada la reconstrucción previa resultado de SFM, que incorpora la posición de las cámaras y algunos puntos reconstruidos. Mediante el empleo de estas técnicas, se consigue una reconstrucción con más puntos y generalmente más detallada que la proporcionada por SFM.

Según la bibliografía consultada [23], los métodos MVS (en este caso PMVS2) no consiguen escalar bien a medida que la complejidad de la tarea aumenta. Ese es el motivo de la existencia de CMVS; este algoritmo permite separar en subprocesos la reconstrucción MVS y poder paralelizarlos, pudiendo abordar el problema de la reconstrucción de modelos complejos.

Los ejemplos que se muestran de reconstrucciones en el desarrollo de este apartado son el resultado de la reconstrucción de SFM y CMVS/PMVS2.

1.3.4.2 FICHERO SALIDA

Como salida del programa se producen numerosos ficheros intermedios y los resultados finales de la reconstrucción. Entre todos los ficheros de salida, más adelante en este documento se hablará de los de extensión PLY, que es el formato que leerá el programa creado en este trabajo. Este formato se conoce como Standford Polygon, para más información vea apartado número 4 de esta memoria.

1.4 EJEMPLOS

Se han hecho más de veinte pruebas diferentes en diferentes escenarios, con distinta iluminación, tamaño y forma. Algunos de estos ejemplos se reflejan en las páginas siguientes, ilustrando diferentes características de los resultados. Los ejemplos son resultado de las reconstrucciones consecutivas mediante SFM y MVS usando los programas VisualSFM, PMVS2 y CMVS, con diferentes configuraciones en cada objeto.

EJEMPLO 1: RECONSTRUCCIÓN DE GRAN CALIDAD

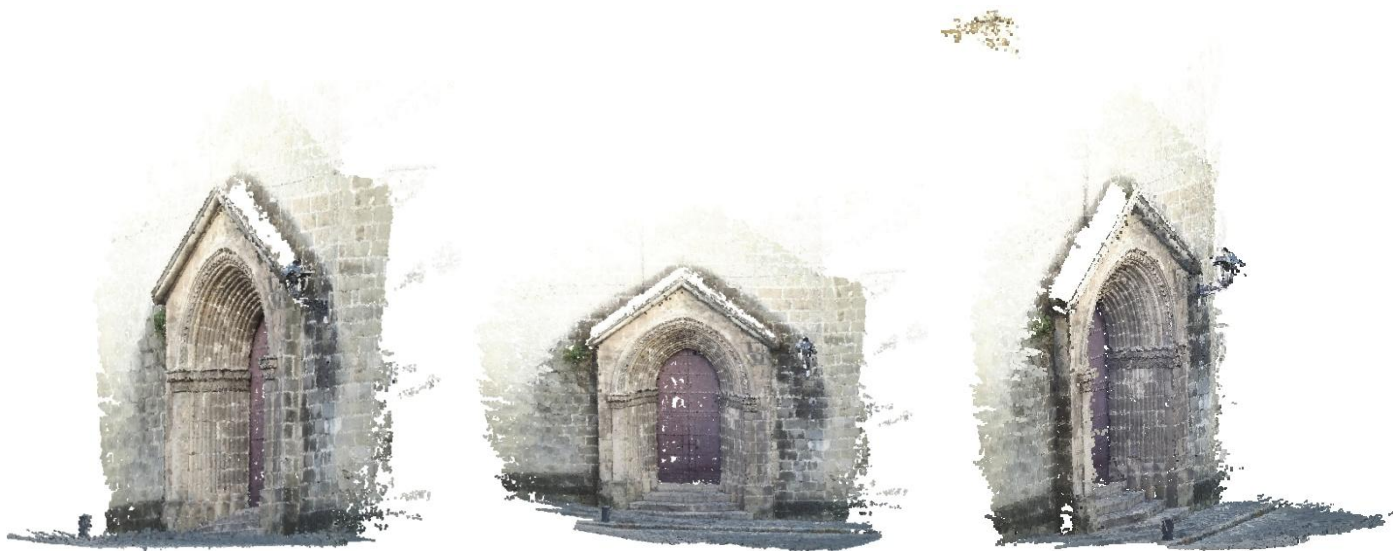
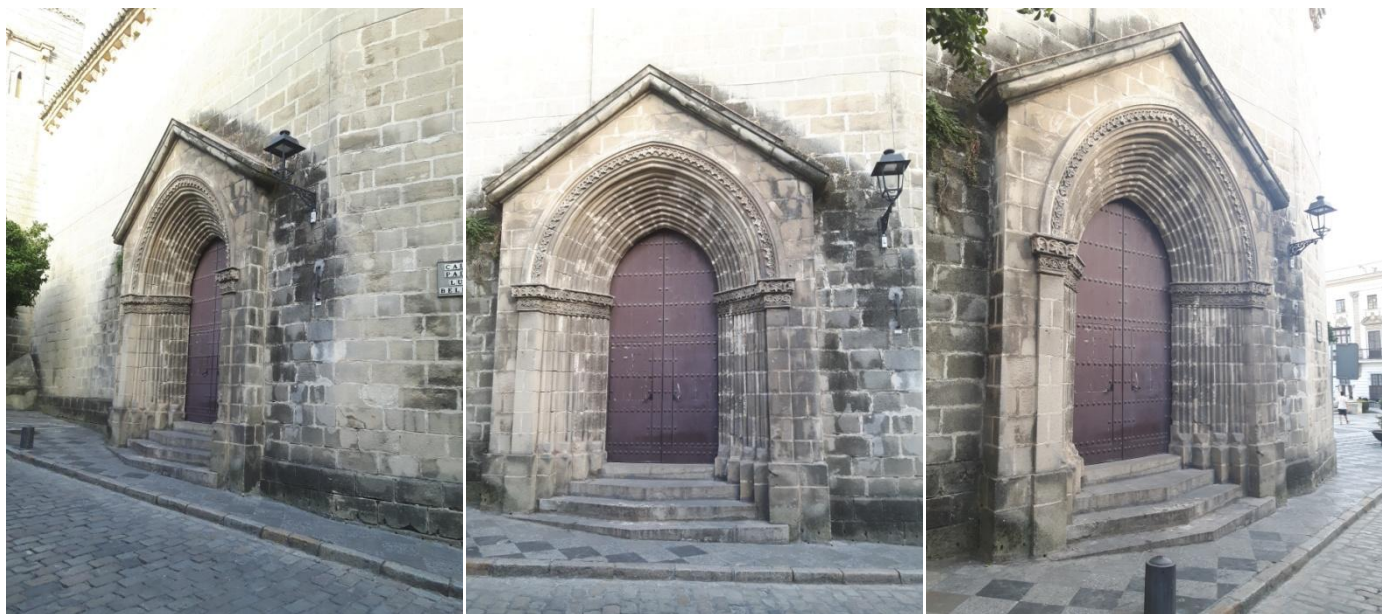


Ilustración 4: Reconstrucción de Puerta lateral de la iglesia de San Dionisio en Jerez de la Frontera (Cádiz).

En la parte superior, 3 de las 40 imágenes usadas para la reconstrucción. Abajo, 3 capturas diferentes de la reconstrucción en tres dimensiones.

EJEMPLO 2: RECONSTRUCCIÓN CON FALSOS PUNTOS

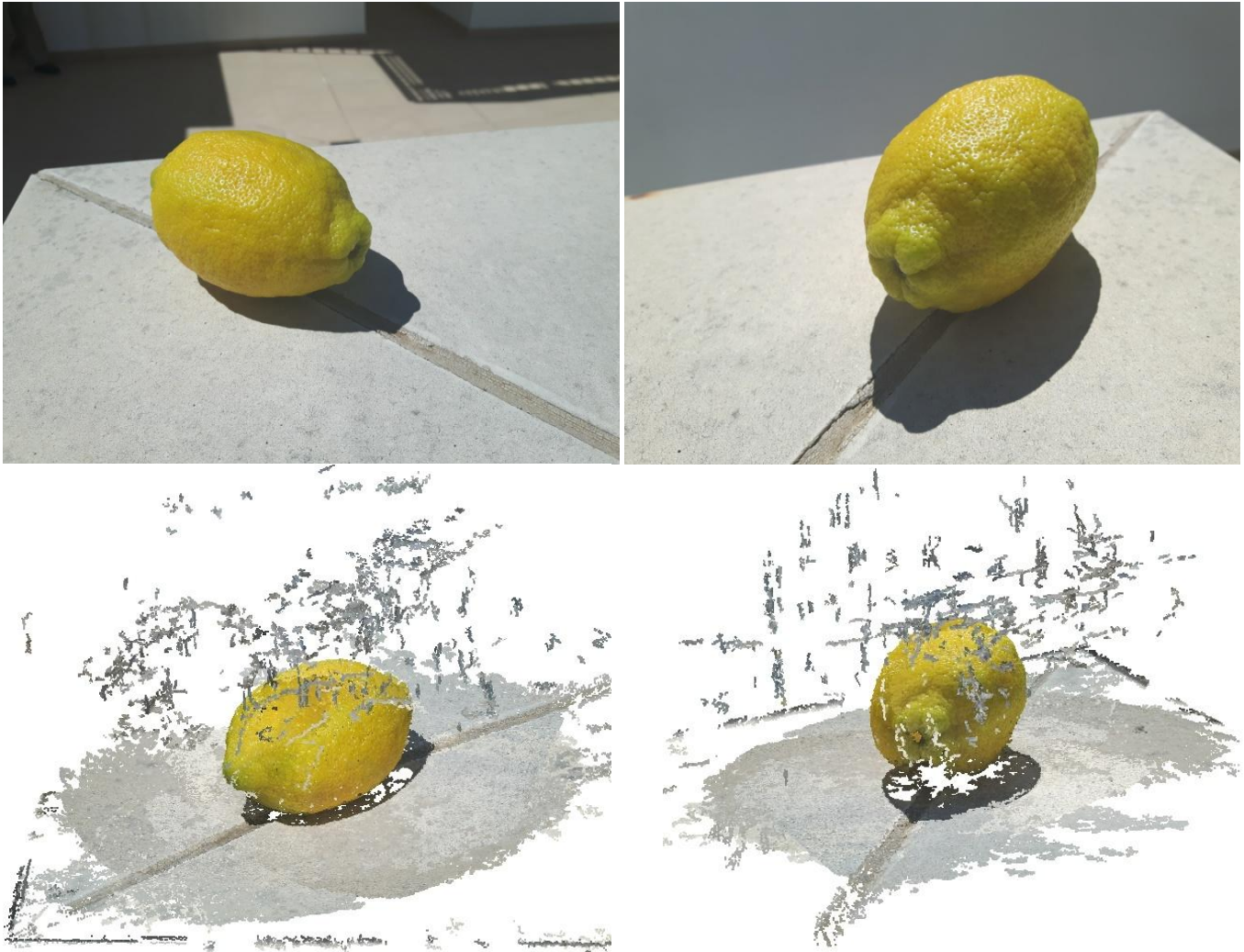


Ilustración 5: Reconstrucción de un limón.

Arriba, dos de las fotografías empleadas para la reconstrucción. Abajo, dos capturas de la reconstrucción. Nótese los puntos “falsos” alrededor de limón. Afortunadamente, estos puntos son fácilmente eliminables.

OTROS EJEMPLOS

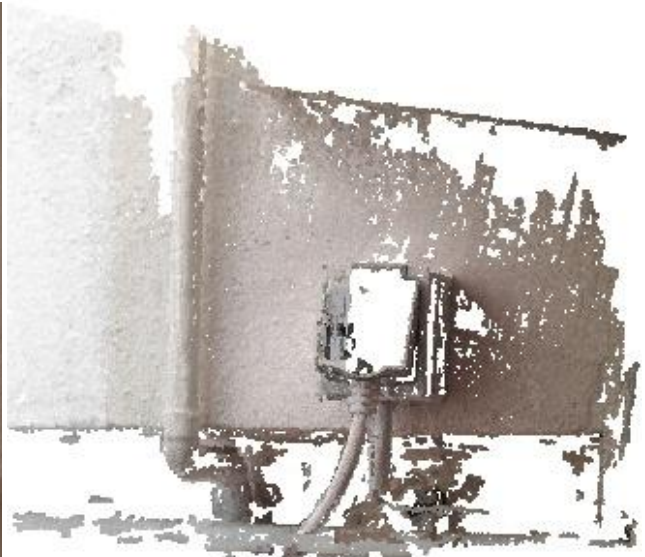


Ilustración 6: Varias reconstrucciones

De arriba abajo: Catedral de Jerez, enchufe, reproducción de la Victoria de Samotracia



Ilustración 7: Reconstrucciones de un molinillo y de una plancha antigua

2 RECONSTRUCCIÓN MEDIANTE UN ESCÁNER DE SUPERFICIE

Hay que tener la mayor información posible para poder elegir correctamente la mejor solución ante un caso concreto.

- Reflexión personal -

En el apartado anterior se ha visto un conjunto de técnicas válidas para la reconstrucción en tres dimensiones a partir de fotografías 2D. Cambiando de idea completamente nos encontramos con el otro gran paradigma existente en el momento, el de los escáneres de superficie.

Un escáner en tres dimensiones adquiere la información de volumen acerca del objeto a partir de una o diferentes fuentes físicas y la almacena en uno o varios ficheros informáticos.

La idea es completamente diferente a la de las técnicas de Structure From Motion. Mientras en aquellas se pretendía usar dispositivos genéricos y no especializados para realizar la reconstrucción (cámaras digitales y ordenadores “estándar”), en este caso se tiene la existencia de un hardware especializado. Con una idea parecida al dispositivo del que toma nombre (el escáner clásico de documentos o fotografías), estos dispositivos buscan ampliar el concepto original y extenderlo a un volumen.

Con una tecnología que era muy novedosa hasta hace relativamente escaso tiempo y todavía poco extendida en el ámbito doméstico, se empiezan a ver propuestas relativamente económicas de escáneres en tres dimensiones, en el pasado con precios prohibitivos y reservadas a sectores comerciales de alto poder adquisitivo. Casi a la par que el desarrollo de las impresoras en 3D, se está viendo en los últimos tiempos la aparición de dispositivos aplicables en el ámbito doméstico. Lo que hace unos años parecía ciencia ficción hoy es una realidad, creándose un mercado cada día más importante alrededor de este mundo. Paseando por las calles ya se puede observar algunas tiendas especializadas en el mundo de las tres

dimensiones (captura, diseño e impresión), incluso las tiendas informáticas clásicas ya poseen muchas en sus catálogos impresoras y escáneres en tres dimensiones a la venta con precios asequibles para el gran público.

Existen multitud de escáneres, desde lo más sofisticados hasta los más sencillos, algunos con capacidades de escanear pequeños objetos y otros de capturar personas enteras.

En este caso, el escáner empleado es el BQ Ciclop [24], ampliamente extendido y conocido. Dada la necesidad de disponer físicamente del dispositivo durante el tiempo requerido para los escaneos y los pasos previos a ellos, para este apartado se ha recurrido a la colaboración con una empresa externa a la universidad [25] que ha sido la que ha proporcionado el escáner, el lugar y la ayuda al escaneo que ha sido requerida. Los costes de los servicios de esta empresa han sido pagados con fondos propios del autor del trabajo, sin que haya repercutido a la Universidad en ninguna forma.

En concreto, el escenario ha sido el que aparece en la Ilustración 8, situado en el mismo establecimiento que ha proporcionado el servicio de escaneo y ayuda al proceso:



Ilustración 8: Escenario de escaneo

2.1 ESPECIFICACIONES DEL ESCÁNER

Es importante conocer las características del escáner para poder extraer juicios acerca de las reconstrucciones. En la página web del fabricante [24] se encuentran las especificaciones del mismo. A continuación se nombran las que se consideran más importantes:

- Máximo tamaño del objeto a escanear: 250mm de diámetro y 205mm de altura.
- Peso máximo del objeto a escanear: 3 kg
- Máximo número de pasos discretos para una vuelta entera: 1600
- **Precisión:** 0.5 mm

2.2 MÉTODO

Tras un primer contacto con las empresas de la zona especializadas en esta labor, se decidió escoger la indicada [25] (Fábrica Digital). El proceso de escaneado se realizó en dos sesiones diferentes de dos días consecutivos. En la primera sesión se dispuso de la ayuda de un trabajador del establecimiento, que se encargó de la configuración y primer escaneo. Posteriormente se analizó el fichero resultado del escaneo, viendo que cumplía los requisitos para el trabajo y que se podía visualizar con las herramientas de que disponía.

En la segunda sesión, se escanearon los dos restantes objetos hasta completar el conjunto de tres, realizando todo el proceso de una forma más ágil, tras haber aprendido en la sesión anterior cómo eran los pasos necesarios y las consideraciones que había que tomar en cuenta para ello.

El equipamiento informático empleado para controlar al escáner en tres dimensiones fue el existente en el establecimiento. Para realizar la captura, el escáner de superficie empleado utiliza la proyección de dos láseres sobre el objeto. Estos láseres provocan una línea visible sobre el objeto delimitando su silueta en el ángulo correspondiente. La cámara incorporada en el escáner se encarga de recoger esa silueta y procesarla. Como el objeto va girando sobre la superficie de escaneo, se captura toda la información de superficie del objeto.



2.2.1 CALIBRACIÓN

Antes de proceder al escaneo de volúmenes tridimensionales, hubo que calibrar el dispositivo con el patrón y siguiendo las instrucciones del tutorial del programa. En la Ilustración 9 se aprecia el proceso de calibración. Este proceso es importante realizarlo para un buen funcionamiento posterior.

Para una correcta medición de las distancias en el proceso de calibrado se usó un dispositivo de medida, el pie de rey digital. Este dispositivo es similar al clásico pie de rey pero las medidas, en vez de ser leídas por inspección sobre una escala impresa o grabada en el instrumento, son proporcionadas a través de una pantalla.

El procedimiento de calibrado se realizó de forma guiada siguiendo las instrucciones proporcionadas por pantalla durante la ejecución del mismo, y sólo requirió de la intervención del usuario en pocas ocasiones.



Ilustración 9: Escáner con patrón de calibración

2.2.2 ESCANEADO DE LOS OBJETOS

Posteriormente se pasó a escanear los tres modelos de ejemplo en este proyecto: el quinqué, la plancha y el libro. Todos ellos se encontraban dentro de los rangos de tamaño requeridos por el dispositivo en sus especificaciones.

Nótese en las fotografías las líneas rojas correspondientes al escaneo. En este proyecto se optó, por economía de recursos informáticos y por simplicidad de la captura, a sólo escanear información de volumen, sin la información de color.



Ilustración 10: Escaneos de los objetos de prueba

2.2.2.1 TIEMPO

El tiempo empleado para el proceso completo depende mucho de la pericia y práctica de la persona encargada. Una gran parte del proceso consiste en el montaje y calibración, por lo que una persona experimentada puede conseguir mejores tiempos al realizar estas tareas. El proceso exclusivamente de escaneo tardó unos 15-20 minutos por objeto.

2.2.3 FICHERO SALIDA

El fichero de salida es, al igual que en el apartado anterior, un archivo con extensión PLY, de tamaño variable dependiendo del objeto escaneado. Este tipo de archivo se comenta detalladamente el apartado 4 (Programa de análisis), más adelante en esta memoria.

El tamaño de estos ficheros es variable dependiendo del objeto escaneado. En este proyecto se manejaron tamaños de decenas de megabytes, conteniendo cada modelo del orden de varias centenas de miles de puntos. Este gran conjunto de datos supuso que trabajar con ellos en el equipamiento informático disponible conllevara un tiempo de computación elevado para las operaciones realizadas, pero asumible si se realiza con eficiencia.

2.3 CONJUNTOS MUESTRA

Se han empleado los tres objetos de referencia en este apartado: el libro, la plancha y el quinqué. A continuación muestro resultados gráficos de los escaneos usando MATLAB con el fichero “ply_read.m” de Pascal Getreuer [26].

A continuación se muestran capturas de la representación de cada uno de los tres objetos. En los tres se han incluido vistas en perspectiva; sin embargo, dada las características geométricas de los objetos, en el libro se ha añadido la vista en planta y en la plancha y el quinqué la vista en alzado. Con ello se pretende una mejor visualización.

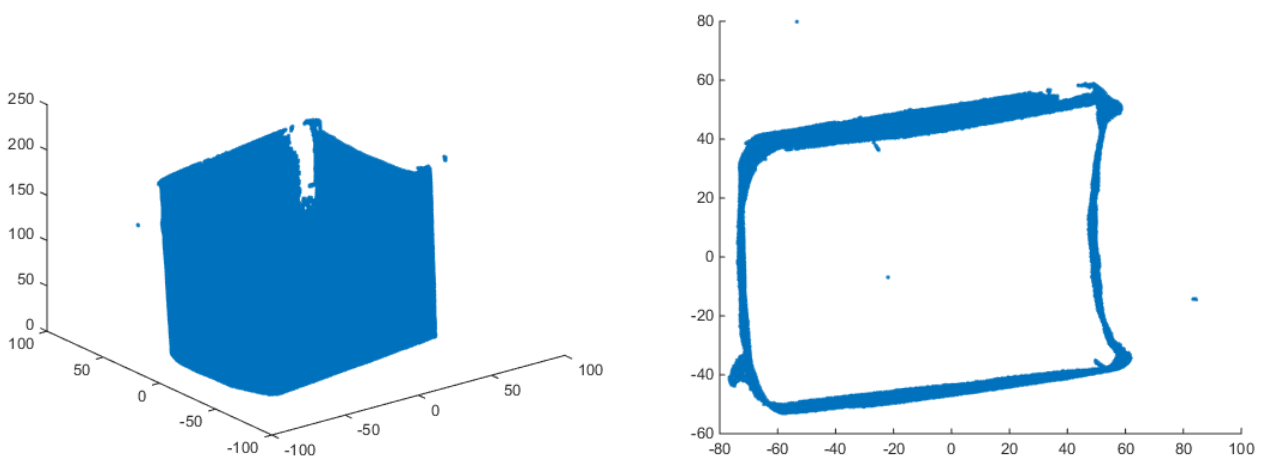


Ilustración 11: Libro en perspectiva y planta

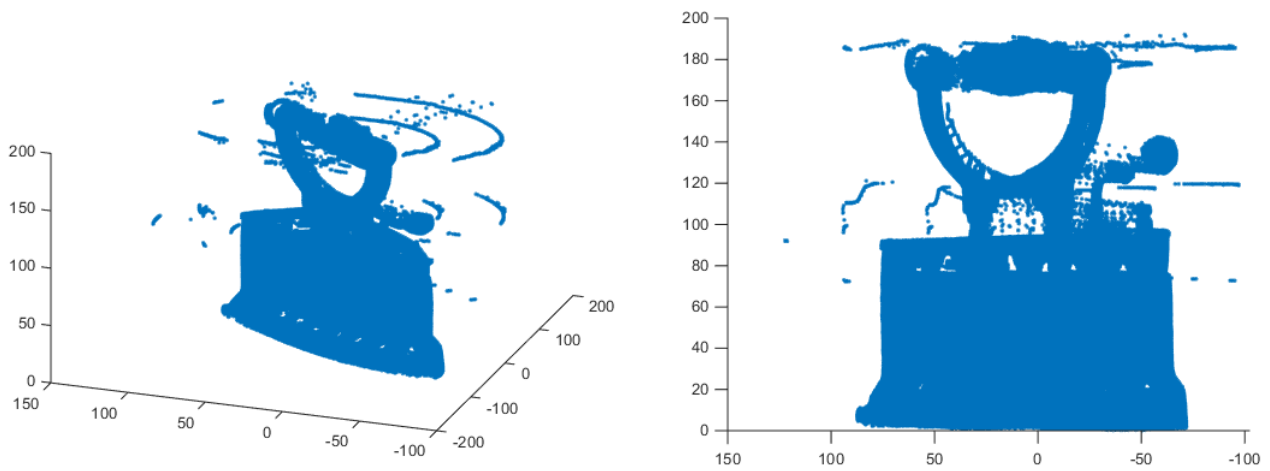


Ilustración 12: Plancha en perspectiva y alzado

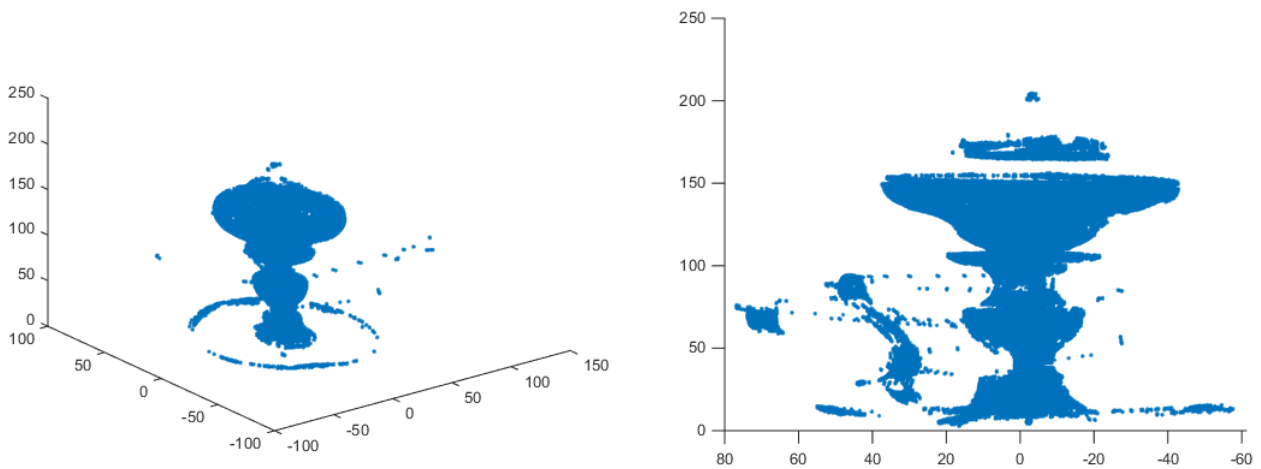


Ilustración 13: Quinqué en perspectiva y alzado

Como comentario final a este apartado, hay que decir que el modelo empleado en los ejemplos del programa de análisis es este último, el del quinqué resultado del escaneo 3D, debido a que el objeto presenta algunas características interesantes (complejidad, cierta simetría, etcétera).

3 PROGRAMA DE ANÁLISIS

Para culminar el aprendizaje y la incursión en el campo de conocimiento de la reconstrucción en tres dimensiones, se considera necesario aprender acerca del tratamiento de los modelos resultantes de las reconstrucciones o escaneos. El formato más extendido para el almacenaje de estos modelos es el formato PLY (Stanford Triangle Format) [27].

Para trabajar con este formato, se ha elegido el lenguaje de programación MATLAB, debido a las avanzadas capacidades que éste ofrece y su extensión en el mundo académico. Para poder importar los ficheros PLY a Matlab se ha empleado el fichero “ply_read.m” de Pascal Getreuer [26].

Los ficheros PLY tienen dos modalidades para almacenar los datos: de forma binaria y de forma ASCII. Para poder trabajar con la función de lectura anteriormente referenciada, se hace necesaria una conversión de los datos. Para esta conversión se empleó el programa MeshLab [28], que permite importar estos ficheros y posteriormente exportarlos en la codificación deseada (binaria o ASCII).

Los modelos con los que se ha trabajado representan el objeto como una nube de puntos, no como un conjunto de polígonos.

Una vez cargado el modelo en tres dimensiones en el entorno ya se puede trabajar con él como si de una matriz se tratara. El programa desarrollado permite varias opciones de funcionamiento. Al iniciar el programa se muestra un menú con las diferentes opciones de las que dispone, que se detallan a continuación.

3.1 CARGA DEL FICHERO

Esta opción pide el nombre del fichero por línea de comandos. Posteriormente hace una llamada al fichero “ply_read.m” [26] para abrirlo.

Hay que tener en cuenta que los datos en el fichero PLY deben estar en formato ASCII para que funcione.

3.2 CONSTRUIR MALLA UNIFORME

Este paso es muy importante para que funcionen otras opciones. Dado que la nube de puntos que representa al objeto en la reconstrucción puede estar en cualquier posición del espacio, una forma de acelerar todos los cálculos es forzar a que estos puntos se encuentren en una malla tridimensional. De esta forma se pasa de un conjunto de localizaciones posibles muy grande a uno más reducido. Una vez hecho esto, se tiene un conjunto de posiciones en el espacio que pueden ser puntos del objeto, pero se tiene la certeza que fuera de esas localizaciones no habrá puntos, por lo que para cualquier operación sobre el modelo no hay que buscar en “casi infinitas” posiciones. Dicho de otra forma, se realiza una cuantificación del espacio tridimensional.

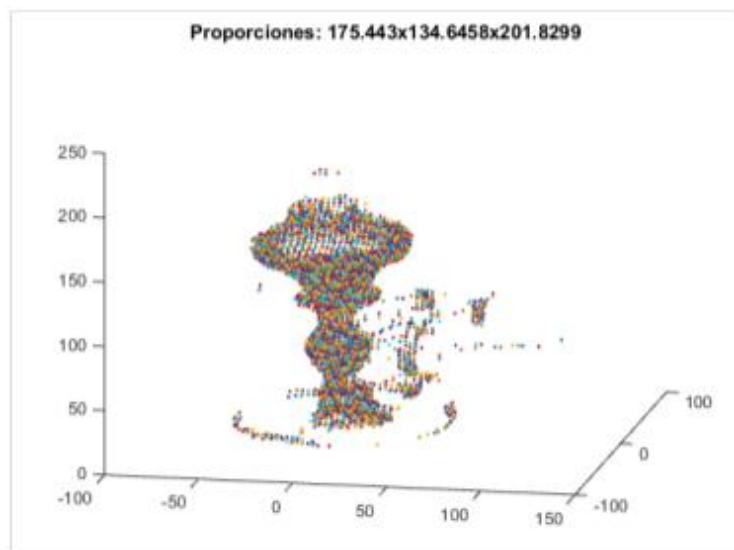


Ilustración 14:

Capura de reconstrucción ajustada a malla uniforme, con intervalo de paso entre posiciones de la malla de magnitud 3 y factor de aumento de resolución 1

En un principio esta tarea se realizó mediante la búsqueda una por una de los puntos en el espacio tridimensional. Sin embargo, este proceso se presentó muy lento y pesado, por lo que se optó por una solución mucho más sencilla pero muy eficaz: el redondeo. Dado un tamaño de paso, se redondean los puntos a esas posiciones. Como desventaja está el problema de que, al redondear, especialmente si el tamaño de paso o de “discretización” es muy elevado respecto a la densidad relativa de puntos en esa localización, puede ocurrir que haya varios puntos con la misma localización o localizaciones “repetidas”. Sin embargo, el ahorro de tiempo al trabajar con conjuntos grandes de nube de puntos justifica la elección de este segundo método de creación de malla.

Al finalizar la tarea, se tendrá una matriz tridimensional con tantas posiciones como lugares en la malla. Si la entrada correspondiente a la posición “x,y,z” es un 1, es un punto del modelo; si es 0 no lo es. Las índices de la matriz son valores absolutos, por lo que para conservar el tamaño relativo del objeto se tiene que guardar también la posición inicial y el paso de la malla tridimensional.

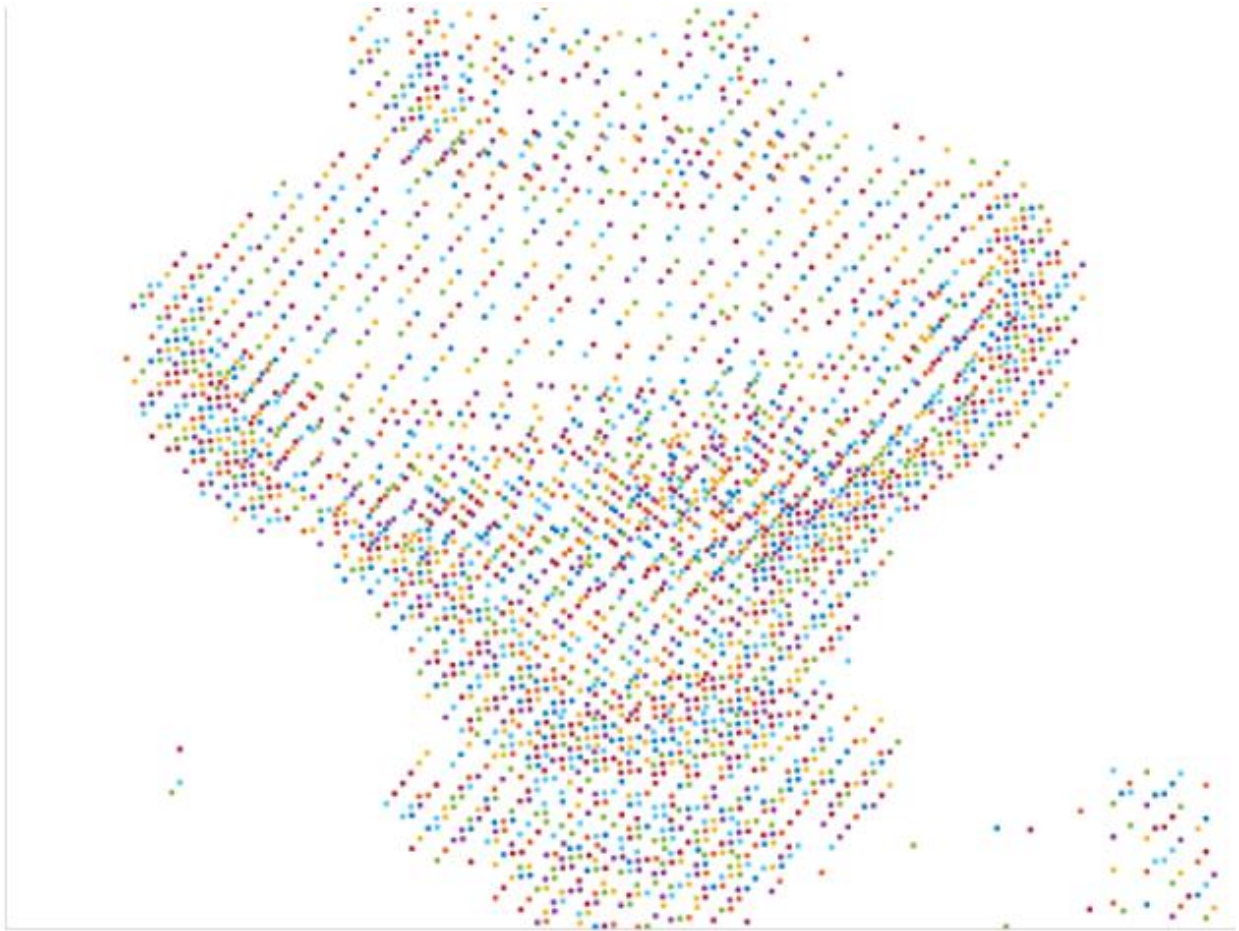


Ilustración 15: Detalle de reconstrucción ajustada a la malla uniforme

3.3 VER SECCIÓN / VER SECCIÓN RELLENA

Una de las opciones que presenta el programa es la posibilidad de ver una sección del objeto. Para que este apartado funcione correctamente, se ha de haber creado la malla uniforme como paso previo. Este apartado permite ver una sección del objeto de la malla tridimensional. Para ello, pide que se introduzca el eje por donde se realizará perpendicularmente la sección: “x”, “y” o “z” (correspondientes a las tres dimensiones del modelo) y, posteriormente, pide el punto de corte por donde se realizará la sección a mostrar. Esta parte funciona fijando una de las dimensiones de la matriz (según el eje indicado) y mostrando la malla bidimensional resultante como una sección. También se calcula el número de puntos de la sección y el número de puntos del volumen total, dando un porcentaje aproximado del “peso” de la sección dentro del volumen total, expresado como el porcentaje de los puntos que están en la sección respecto a los puntos totales a 1 (“de objeto”) en la malla tridimensional.

El método explicado anteriormente muestra la sección tal como está en la matriz de puntos. Sin embargo, es posible que se quiera ver la sección rellena. En este caso, se realizan los pasos anteriores pero, antes de mostrar por pantalla la sección, se rellena su interior con todos 1, o sea, se rellena todo el hueco que hay dentro de la sección según se ve en en una de las dos dimensiones resultantes (se solicita si se quiere rellenar en horizontal o en vertical). En este caso no calcula el porcentaje respecto al total ya que se han añadido puntos al relleno.

El relleno se realiza usando el siguiente método: dependiendo de si se pide relleno en vertical u horizontal, se selecciona una columna o fila, se buscan el primer y último valores no cero y se rellena en su interior. El proceso se repite hasta completar todas las columnas o filas. A continuación, se muestran ejemplos de secciones, primero sin relleno y después rellenas. Se ha usado un intervalo de paso de magnitud 3 entre las muestras posibles de la malla, en la práctica tiene poca precisión pero se ha utilizado para que se vea claramente los puntos de la malla:

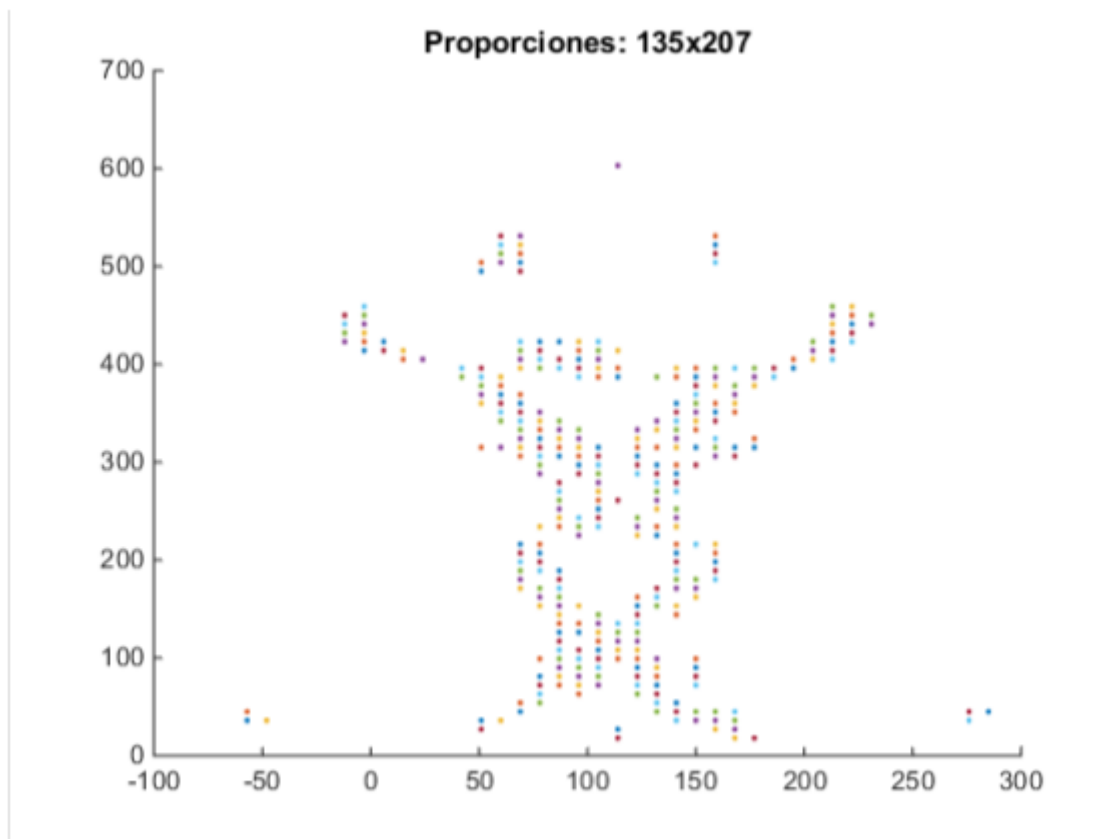


Ilustración 16: Sección sin relleno. Plano 'x', posición 0

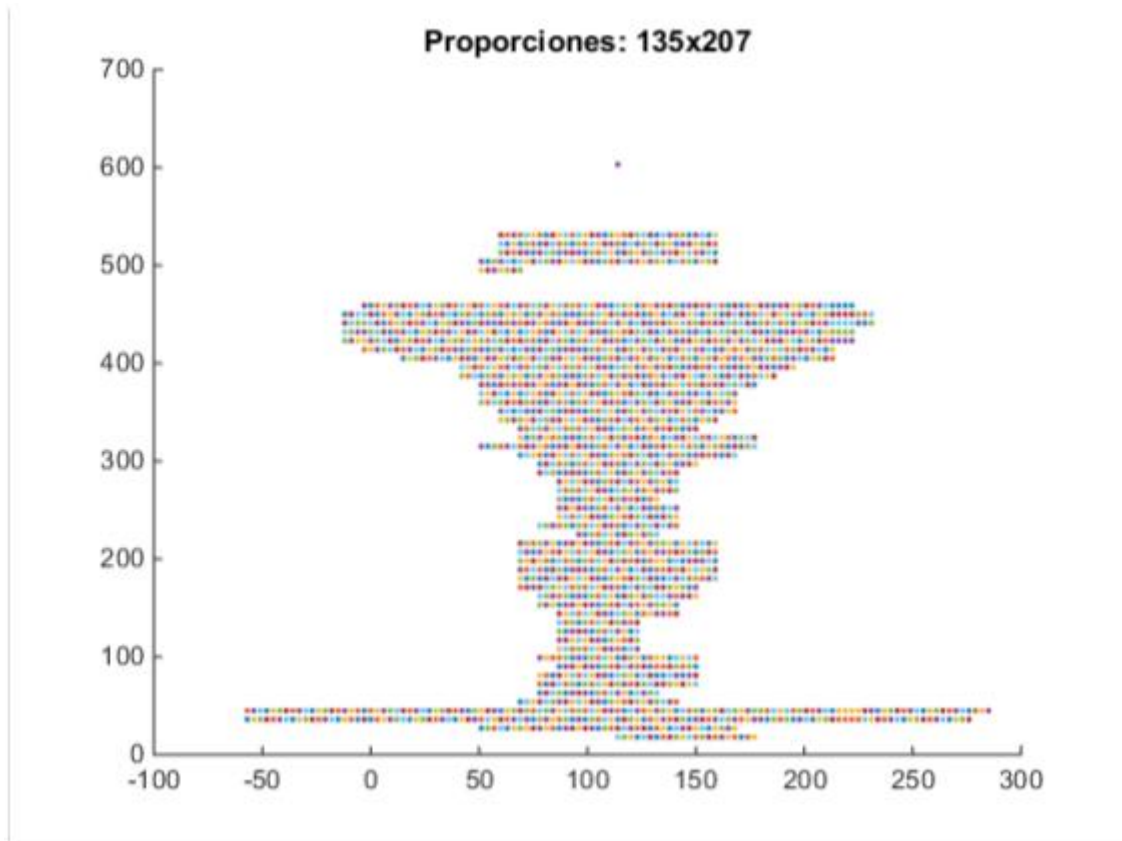


Ilustración 17: Sección rellena en horizontal (rellenado en dirección correcta). Plano 'x', posición 0

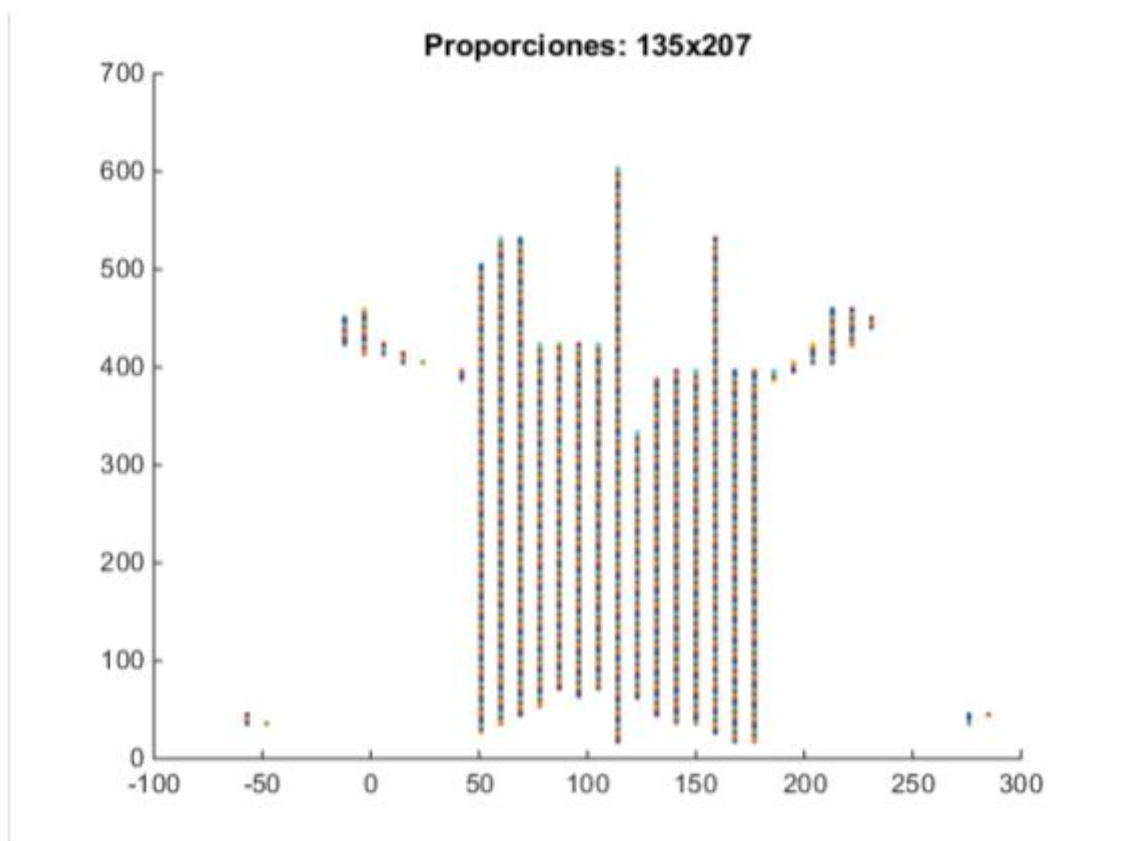


Ilustración 18: Sección rellena en vertical (rellenado en dirección incorrecta). Plano 'x', posición 0

3.4 LIMPIEZA DEL MODELO

Los modelos cargados en el entorno pueden tener muestras de ruido. Estas muestras son aquellas que no deberían aparecer en un caso ideal, pero que aparecen en la nube de puntos resultado. En algunas circunstancias es conveniente limpiar el modelo, eliminando este ruido.

Para este apartado se hace una suposición: las muestras ruidosas se encuentran aisladas del resto de muestras válidas del objeto, y a una distancia suficientemente grande. Esto no tiene por qué ser siempre así, pero en las pruebas realizadas con el método descrito a continuación se han observado buenos resultados.

El método de eliminación de ruido consiste en la división del espacio tridimensional de la malla de puntos uniforme en pequeños subespacios, de forma que el prisma formado por la matriz que representa la malla uniforme se analice en pequeños subprismas. Por cada uno de estos pequeños prismas se analizará también un prisma algo mayor, centrado en la misma posición y que contiene todas las posiciones de su homólogo más pequeño (en el caso de los extremos puede que no esté centrado o que no tenga más muestras).

Se analiza el contenido de cada una de estas matrices tridimensionales. Si el contenido de muestras a uno de la submatriz pequeña es mucho menor al de la submatriz grande, entonces se considerará que está en una zona del objeto real, no aislada, y se mantendrá. En el caso contrario, significará que en un conjunto del volumen alrededor del prisma pequeño no existirá apenas muestras, por lo que cabe pensar que es una región aislada, y por tanto es ruido.

Para controlar el funcionamiento del algoritmo se puede elegir el tamaño de lado del subprisma pequeño y del subprisma grande. Es importante especificar, entre otros, una diferencia de estos valores que sea adecuada para la limpieza deseada.

A continuación se muestra un ejemplo del funcionamiento del algoritmo, en el caso del quinqué, que en principio tiene muestras ruidosas (en especial por la zona del mango), y tras la ejecución del algoritmo muchas de esas muestras desaparecen.

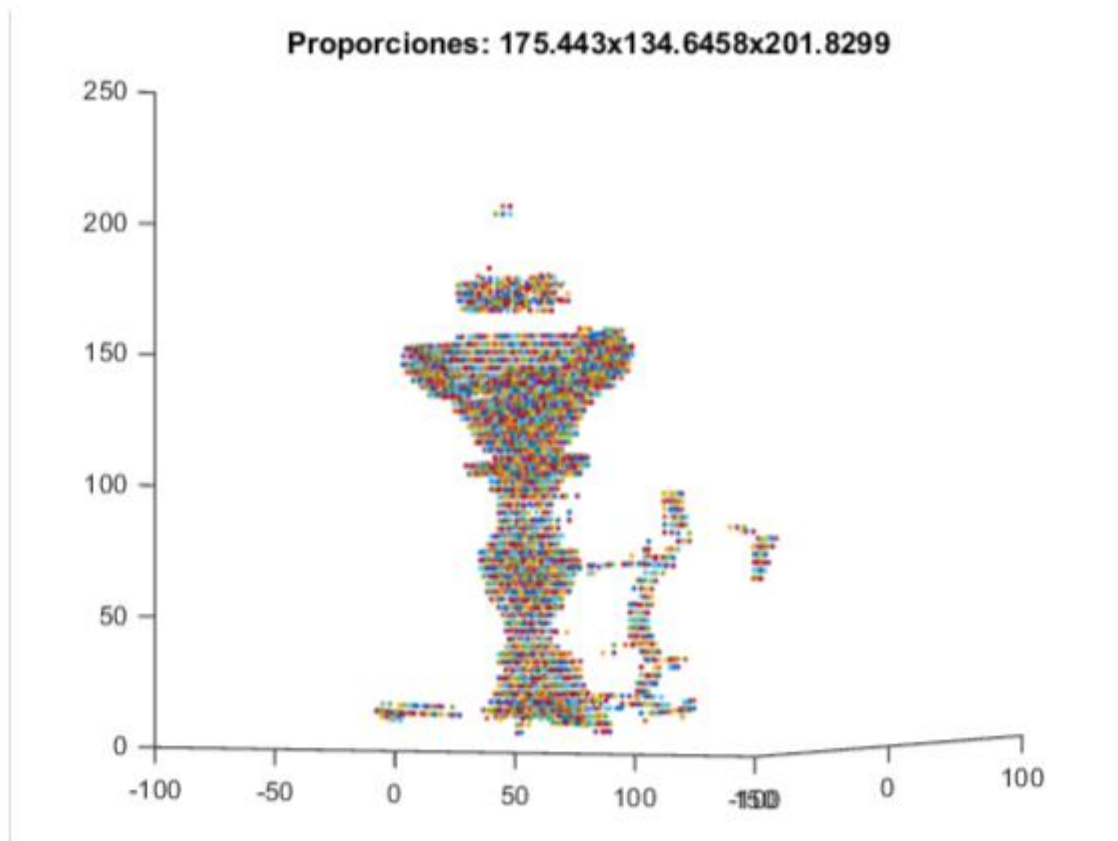
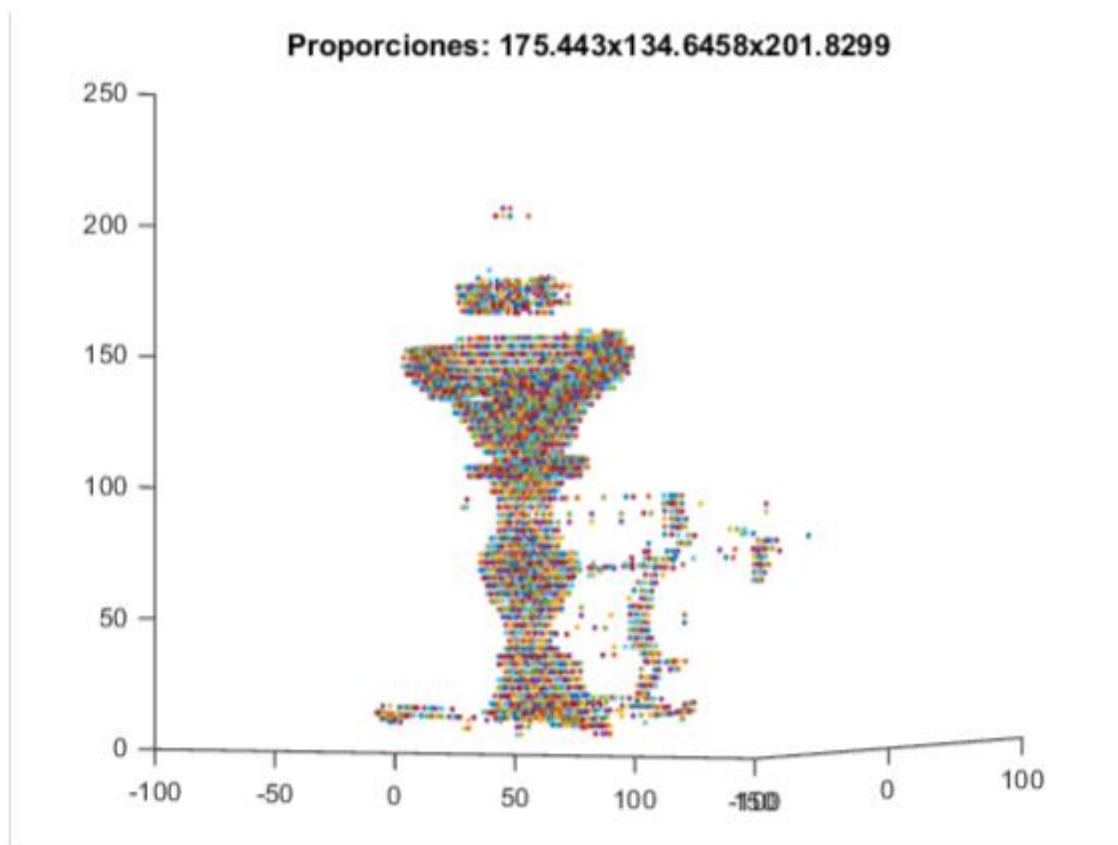


Ilustración 19: Antes y después de la limpieza de ruido (las perspectivas pueden variar ligeramente)

3.5 HOJA DE DATOS DEL MODELO

Otra de las posibilidades que oferta el programa realizado es la opción de realizar una “hoja de datos” del modelo cargado. En esta hoja, impresa por pantalla, se pueden visualizar datos de interés siempre que haya un modelo cargado en la memoria. De esta forma se pueden comprobar algunos aspectos, tales como el número de puntos del modelo, que permitan discernir posibilidades antes posibles errores.

Se presenta a continuación un ejemplo de una salida de la opción:

```
HOJA DE DATOS DEL MODELO
*****
Nombre del modelo fichero cargado: cinco_escaner.ply
Numero de puntos del modelo original (no malla): 479047
Numero de nodos posibles de la malla: 4910688
Numero de nodos usados de la malla: 32257(0.65687 por ciento
del total)
```

Ejemplo de salida de hoja de datos, con factor de aumento de resolución 1 y malla de paso 1

La explicación de cada uno de los valores dados es la siguiente:

- Número de puntos del modelo original (no malla): Cantidad de puntos de la nube de puntos cargada.
- Número de nodos posibles de la malla: Posibles localizaciones candidatas para ser lugares del objeto. También se puede ver como lugares de posibles proyecciones tras el cuantificado tridimensional de la nube de puntos original.
- Número de nodos usados de la malla: Cantidad de posiciones de la malla que son puntos del objeto. Se da el porcentaje sobre el total de puntos de la malla.

3.6 COMPARACIÓN

Para terminar el programa, se incluye un apartado de comparación entre dos modelos, muy útil si se desean comparar modelos de un mismo objeto pero generados con dos métodos diferentes. Como se verá a continuación, este proceso es complejo y ha sido el que más trabajo y tiempo ha requerido para ser desarrollado.

Es necesario tener cargado un modelo para que esta opción funcione (paso primero). Una vez elegida la opción, pregunta por el modelo con el que comparar el que está ya previamente cargado.

Cuando se quiere trabajar con dos modelos tridimensionales a la vez aparecen dos problemas principales: la utilización de sistemas de referencia distintos y de resoluciones diferentes.

3.6.1 DIFERENTES REFERENCIAS

En este apartado, cuando se expresa que cada modelo viene en un sistema de referencia distinto, se refiere a que cada uno tiene:

- Diferente escala y origen de coordenadas. Cada algoritmo de escaneado tomará una referencia escalar diferente y un origen de coordenadas diferente. Incluso es posible que influya la distancia al objeto de captura, o la imposibilidad de tener una referencia absoluta de longitud. No es extraño que para un mismo objeto capturado por diferentes métodos, sus longitudes varíen en varios órdenes de magnitud. Por tanto, hay que unificar criterios.
- Diferente orientación. Otro problema que se presenta es la diferencia angular entre ambos modelos. Es posible que se tengan dos modelos con el mismo origen de coordenadas y el mismo tamaño, pero se encuentran girados en las dimensiones del espacio. Este aspecto también hay que corregirlo.

3.6.2 DIFERENTES RESOLUCIONES

Por otro lado se encuentra el problema de la resolución. No olvidemos que estamos trabajando con nubes de puntos; en concreto, en este apartado se trabaja con la forma original (conjunto desordenado de puntos).

Dada la misma naturaleza de escaneo, es posible que un método proporcione una resolución muy diferente al otro. Por tanto, hay que escoger una referencia que sea justa para comparar ambos modelos.

3.6.3 RESULTADO

El programa, una vez cargados los dos modelos, solicita datos sobre ellos. De una forma iterativa, a partir de los datos proporcionados, realiza el ajuste de los modelos. Es importante que la información que se le proporcione sea precisa, pues de ello dependerá la calidad del ajuste de ambos modelos. Para ello, además de estos datos, el algoritmo se basa en suposiciones previas acerca de la representación tridimensional de ellos.

Como resultado de este apartado se generan dos valores diferentes “r1” y “r2”, cuyos significados son:

“Dentro de un margen de tolerancia, expresado en distancia absoluta en unidades de matriz, qué porcentaje puntos del modelo indicado (1 o 2, respectivamente) están, dentro de un margen de tolerancia, cerca de puntos del otro modelo. Dicho de otra forma, qué porcentaje de puntos de una nube de puntos serían candidatos para encontrarse en la forma que representa la segunda nube de puntos.”

En el caso que tomamos como ejemplo, si comparamos a simple vista el modelo generado para el escáner 3D y el de SFM, nos damos cuenta que el escáner prácticamente sólo captura el objeto, mientras que SFM también captura puntos del entorno (superficie en la que se apoya el objeto, paredes, etc). Por tanto, cabe esperar que uno de estos valores (r1) sea alto, y el otro (r2) bastante más bajo, ya que los puntos del escáner se corresponden casi todos a lugares cercanos a lo de SFM, pero hay lugares de SFM que no se representan con el escáner 3D, el entorno p.ej.

Para terminar, comentar que el apartado concluye mostrando también por pantalla dos hojas de resultados, similares a las del apartado anterior, correspondientes a cada uno de los dos ficheros cargados; a continuación se muestra un ejemplo:

HOJA DE DATOS DEL MODELO PRIMERO

Nombre del modelo fichero cargado: libro_escaner.ply

Numero de puntos del modelo original (no malla): 1335900

Numero de nodos posibles de la malla: 25056

Numero de nodos usados de la malla: 2793(11.147 por ciento del total)

HOJA DE DATOS DEL MODELO SEGUNDO

Nombre del modelo fichero cargado: libro_sfm.ply

Numero de puntos del modelo original (no malla): 449628

Numero de nodos posibles de la malla: 1283854

Numero de nodos usados de la malla: 8324(0.64836 por ciento del total)

Se adjunta en el Capítulo 4 la salida de la ejecución de las pruebas de ejemplo junto con algunos comentarios.

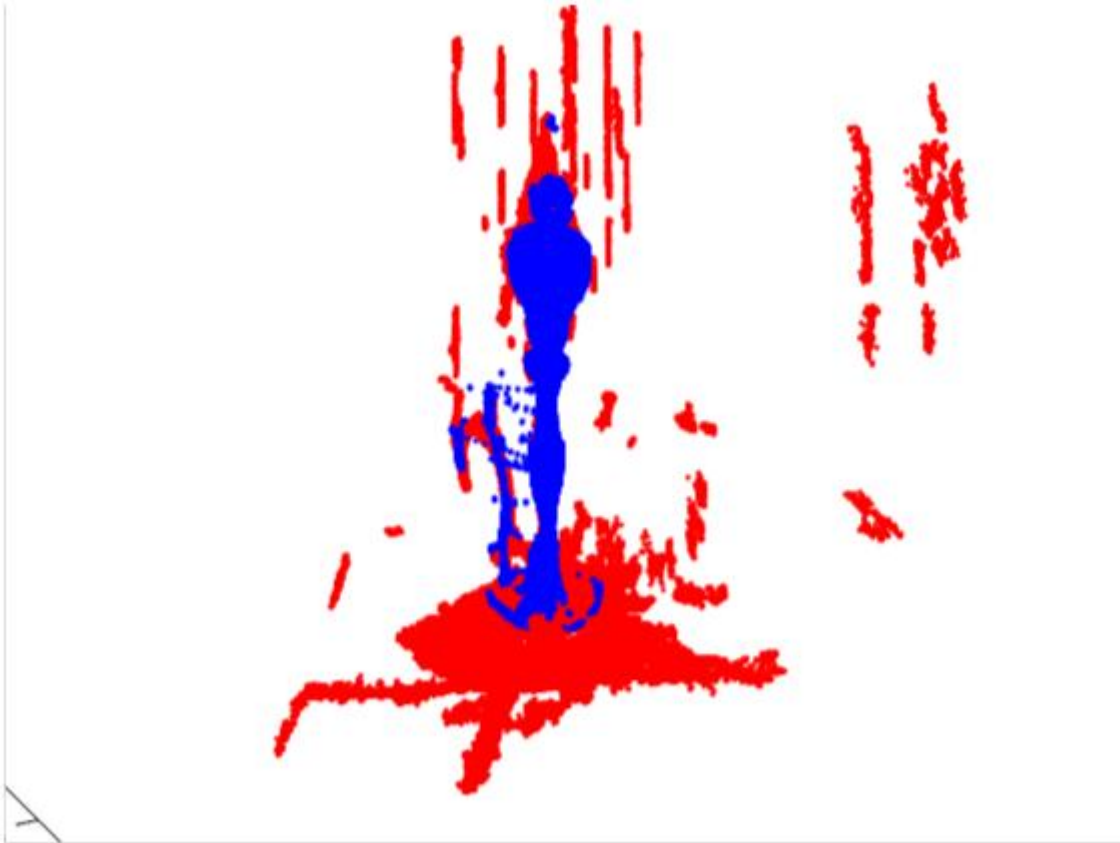


Ilustración 20: Superposición de los dos modelos generados (escáner 3D y SFM-MVS), tras ejecutar el algoritmo de ajuste de referencias

3.7 COMENTARIO FINAL

Como comentario adicional, decir que en algunos de los apartados del programa se permite “aumentar” la resolución a la hora de construir la malla uniforme. Esto es debido a que el lenguaje de programación no permite expresar índices de matrices que no sean naturales (mayores o iguales a 1), y en algunas ocasiones la escala de la nube de puntos es muy pequeña. Con este “aumento” de la resolución se expande la escala de la nube de puntos original antes de crear su malla uniforme, de tal modo que la malla tenga más resolución espacial donde muestrear la nube de puntos original.

4 RESULTADOS DE LA COMPARATIVA CON EL CONJUNTO DE EJEMPLO

Tras haber realizado todo el proyecto, se valida el funcionamiento de todo el procedimiento empleado, desde la captura de los modelos con la cámara fotográfica del dispositivo móvil o del escáner hasta la comparación de los modelos por el programa informático realizado. Para ello, se sigue el proceso completo de trabajo con tres objetos diferentes: el libro, el quinqué y la plancha antigua.

Todos ellos fueron fotografiados y reconstruidos con las técnicas SFM-MVS y capturados con el escáner de superficie. Para comprobar el funcionamiento del programa realizado, se probaron todas las opciones, incluida la de la comparativa que fue ejecutada con los pares de modelos correspondientes. Los resultados mostrados a continuación son el resultado de la salida del algoritmo de ajuste de referencias y comparación a partir de los pares de modelos introducidos.

En la tabla comparativa que se muestra a continuación se expresa, para cada objeto, los siguientes valores:

- Número de fotografías empleadas para la reconstrucción SFM-MVS
- Puntos nube SFM-MVS: Tras haber ejecutado los algoritmos correspondientes, cantidad bruta de puntos existente en el fichero PLY tomado como resultado.
- Puntos nube escáner 3D: Una vez realizado el escaneado en tres dimensiones del objeto, cantidad bruta de puntos existente en el fichero PLY tomado como resultado.
- Puntos malla uniforme SFM-MVS/escáner 3D: Cantidad de posibles posiciones en el espacio de proyección tridimensional de la malla uniforme. Entre corchetes, cuántos de esos puntos están finalmente activos (representando al objeto)
- Tolerancia: valor de tolerancia para el que se expresan los valores r_1 y r_2 . Ver más información en el subapartado correspondiente en las páginas siguientes.
- r_1 / r_2 : Porcentaje de puntos del modelo 1 ó 2 que son candidatos a formar parte del objeto que representa el modelo 2 ó 1, respectivamente. Es decir, dicho de forma muy relajada, porcentaje del objeto en 1 ó 2 que sería candidato a encajar en la representación del objeto en 2 ó 1, respectivamente. Para esta opción se toma un valor de tolerancia, expresado en unidades naturales de tamaño de la malla uniforme.

Si necesita más información acerca del significado de cada valor, acuda a los apartados 1 a 3 de esta memoria.

TABLA DE RESULTADOS

AVISO: Los resultados de esta comparativa no sólo dependen de la calidad de los algoritmos de reconstrucción de los modelos en tres dimensiones; también intervienen las condiciones de captura, la calidad de la cámara, la configuración de los algoritmos de captura y la precisión de la información proporcionada al programa de análisis, entre otros factores.

Nombre [tolerancia]	Número de fotografías	Puntos nube SFM-MVS	Puntos nube escáner 3D	Puntos malla uniforme SFM-MVS [puntos usados]	Puntos malla uniforme escáner 3D [puntos usados]	r1	r2
Libro* [1]	68	449628	1335900	1283854 [8324]	25056 [2793]	79.9499	24.4834
Libro* [2]						95.7393	30.6944
Libro* [3]						99.0333	34.1302
Quinqué [1]	49	448671	479047	252525 [2797]	8568 [567]	97.1781	29.1026
Quinqué [2]						98.2363	35.0375
Quinqué [3]						98.9418	38.3625
Plancha [1]	61	423276	842332	1332090 [4181]	9614 [1031]	85.8390	20.6888
Plancha [2]						95.4413	26.4769
Plancha [3]						99.7090	33.3413

- Para ver el significado de los parámetros, acuda a los apartados correspondientes.
- Aumento de resolución con escala 3 en todos los resultados y con un paso de la malla uniforme de una unidad.
- Tamaño de la malla uniforme tras la comparación entre ambos modelos del mismo objeto usando tolerancia unidad, y los parámetros del punto anterior. Al ajustar tamaños, los modelos de SFM-MVS cubren más volumen que los del escáner (alrededores del objeto), por eso es mayor la malla uniforme.

* En los modelos del libro se observa una leve deformación del objeto al apoyarlo sobre superficies diferentes en instantes distintos. Es posible que por eso se produzca un menor parecido de los modelos entre sí.

4.1 TOLERANCIA

Los valores r_1 y r_2 se expresan en base a una tolerancia, según se explicó en el apartado 3 de esta memoria. Este valor de tolerancia se expresa en unidades naturales de índices de la malla uniforme. Como cada modelo tiene un tamaño de malla uniforme diferente, se exponen a continuación valores relativos respecto a estas dimensiones (valores aproximados). Estos tamaños están tomados respecto a la diagonal tridimensional de la malla uniforme final tras el cálculo (esta malla cambia de tamaño en el proceso):

	1	2	3
Libro	1,98%	3,97%	5,95%
Quinqué	3,02%	6,04%	9,05%
Plancha	2,83%	5,66%	8,49%

5 CONCLUSIONES

Hasta aquí el desarrollo formal del trabajo. Han sido varios meses de esfuerzos que se plasman en este documento. Es complicado definir un conjunto de “conclusiones” o “lecciones aprendidas” que se puedan extraer del trabajo y es difícil sintetizarlos en unos cuantos puntos redactados. Tras realizar una síntesis general, las conclusiones que se pueden extraer del desarrollo de este Trabajo Fin de Máster son:

- Se ha hecho patente la necesidad de equipo informático relativamente potente para el despliegue de las técnicas expresadas en el trabajo. El equipamiento utilizado no poseía gran capacidad computacional, por lo que los tiempos de espera se alargaron hasta varias horas en algunas reconstrucciones. Como consecuencia de esto se presenta una cierta criticidad en los fallos, puesto que al repetir algunos procesos se demoraba mucho el avance del trabajo.
- Las tecnologías de reconstrucción de objetos en tres dimensiones tienen un interesante futuro por delante. Sin embargo, a día de hoy se presentan todavía problemas como la poca documentación disponible en algunos procesos, o en el elevado número de pasos necesarios en algunos casos. Este problema debe solucionarse con el tiempo, cuando la tecnología esté suficientemente probada muchos de estos problemas se mitigarán.
- El nivel conceptual necesario para comprender las técnicas es elevado. Constituye un gran esfuerzo intelectual llegar a entender la teoría detrás de la tecnología y poder darle explicación a los resultados obtenidos tras cada paso dado.

Por todo ello, tras terminar el proyecto, se concluye que se ha producido un gran avance en las técnicas de reconstrucción tridimensional, y las bases teóricas y las primeras aplicaciones prácticas funcionales ya están en marcha. Sin embargo, queda todavía mucho camino que recorrer en este campo.

6 GLOSARIO

- Alzado (vista): Imagen que se forma de un objeto cuando se ve desde uno de sus laterales.
- ASCII: Formato de representación de datos en el que la información se almacena en grupos de 7 bits (8 bits para ASCII extendido), lo que da hasta 128 valores posibles (256 para ASCII extendido). En esta codificación entran tanto letras, números y otros caracteres especiales. [29]
- Binario: Formato de representación de datos que permite sólo dos tipos de valores diferentes, el 0 y el 1. Una unidad de información almacenada en binario se denomina bit.
- Calibración: Procedimiento por el cuál se prepara un dispositivo para entrar en funcionamiento mediante el ajuste de algunos de sus parámetros. Estos parámetros pueden ser propios del dispositivo, externos a él o una mezcla de ambos.
- CPU: Unidad central de procesamiento, dispositivo encontrado en un ordenador moderno, comúnmente llamado “procesador”, que se encarga de realizar las operaciones matemáticas con los datos que se le proporciona.
- Distancia focal: Distancia, expresada normalmente en milímetros, entre el centro óptico de un objetivo fotográfico y el plano de formación de la imagen.
- Enfoque: Ajuste de la(s) lente(s) para conseguir que la imagen de interés se forme en el dispositivo de captura de forma nítida. Esto se consigue generalmente variando la distancia focal de la cámara para que en el punto de captura el plano donde se ubica el objeto de interés se represente con los círculos de confusión suficientemente pequeños.
- Escáner en 3 dimensiones: Un escáner en tres dimensiones adquiere la información de volumen acerca del objeto a partir de una o diferentes fuentes físicas y la almacena en uno o varios ficheros informáticos.
- Escáner tradicional: Dispositivo que obtiene una representación en dos dimensiones de un objeto plano.

- ISO (sensibilidad): En una cámara digital, cantidad de ganancia aplicada a la salida física de los sensores. Generalmente un mayor valor ISO proporcionará una imagen más luminosa si las condiciones son de baja iluminación, a cambio de aumentar el ruido de imagen de los sensores.
- Megapíxel (Mpx): Unidad de conteo de píxeles correspondiente a un millón de píxeles. En algunas referencias también se interpreta como la unidad que define a 1024x1024 píxeles (1.048.576 píxeles). Se emplea para medir la definición de las cámaras digitales y de las fotografías digitales.
- Planta (vista): Imagen que se forma al ver un objeto desde una posición perpendicular a él desde su parte superior (“se ve desde arriba”)
- PLY (formato de archivo): Formato de archivo que almacena información de volumen acerca de un objeto, entre otros valores.
- Ruido de sensor: En una cámara digital, ruido que se genera en una imagen proveniente de los mecanismos de captación de imagen.
- SFM: Se conoce como técnicas de Structure From Motion (SFM) a aquellas que realizan una reconstrucción en tres dimensiones de un escenario a partir de un conjunto de fotografías.
- Simetría de revolución: Un objeto con simetría de revolución es aquel que puede ser generado al girar una superficie con una forma determinada alrededor de una línea que se llama “eje de revolución”.
- Tarjeta gráfica: Dispositivo informático de procesamiento especializado en realizar cálculos relacionados con aspectos gráficos y de proporcionar una señal de vídeo. En la actualidad el concepto se ha expandido, teniendo unas funciones mucho más amplias.
- Tiempo de exposición: Tiempo durante el cuál el mecanismo captador de la cámara está expuesto a la luz. Durante este tiempo se forma la imagen. Si el tiempo de exposición es demasiado elevado, la imagen se sobreexpondrá (muy blanca); por el contrario, si es muy bajo se subexpondrá (muy oscura).

REFERENCIAS

- [1] Paul Bourke. “Automatic 3D reconstruction: An exploration of the state of the art”. GSTF Journal on Computing, October 2012, Volume 2, Number 3, pp 35-40, ISSN 2251-3043
- [2] Prof. Ashutosh Saxena and Prof. Andrew Y. Ng. “Make 3D”. Cornell University <http://make3d.cs.cornell.edu/>
- [3] Changchang Wu, "VisualSFM: A Visual Structure from Motion System", <http://ccwu.me/vsfm/>, 2011
- [4] https://es.wikipedia.org/wiki/Structure_from_motion, consultado por última vez el 17/11/2016
- [5] “The Structure from Motion Pipeline” <https://www.youtube.com/watch?v=i7ierVkXYa8>, consultado por última vez el 19/11/2016
- [6] M.J. Westoby, J. Brasington, N.F. Glasser, M.J. Hambrey, J.M. Reynolds “‘Structure-from-Motion’ photogrammetry: A low-cost, effective tool for geoscience applications”. *Geomorphology*, Volume 179, 15 December 2012, Pages 300–314
- [7] “3D reconstruction from multiples images”. Wikipedia en ingles. https://en.wikipedia.org/wiki/3D_reconstruction_from_multiple_images Consultado por última vez el 18/11/2016
- [8] Sylvain Paris. Presentación “Methods for 3D Reconstruction from Multiple Images” MIT CSAIL
- [9] Shawn McCann “3D Reconstruction from Multiple Images”
- [10] Savil Srivastava, Ashutosh Saxena, Christian Theobalt, Sebastian Thrun, Andrew Y. Ng. “i23 - Rapid Interactive 3D Reconstruction from a Single Image”. VMV 2009
- [11] Carlos Hernandez, George Vogiatzis and Yasutaka Furukawa “3d Shape Reconstruction from Photographs: a Multi-View Stereo Approach” Half-day tutorial at CVPR 2010 Monday, June 14, 2010
- [12] Yasutaka Furukawa, Carlos Hernández “Multi-View Stereo: A Tutorial” *Foundations and Trends® in Computer Graphics and Vision*, vol. 9, no. 1-2, pp. 1–148, 2013.
- [13] MathWorks Documentation. “What Is Camera Calibration?” <https://es.mathworks.com/help/vision/ug/camera-calibration.html?requestedDomain=www.mathworks.com> Consultado el 18/11/2016

-
- [14] Prof. Rob Fergus. Presentación “Lecture 6: Multi-view Stereo & Structure from Motion”
- [15] Yasutaka Furukawa, Jean Ponce. “Accurate Camera Calibration from Multi-View Stereo and Bundle Adjustment” CVPR 2008
- [16] Blog del fotógrafo. “Todo Lo Que Necesitas Saber Sobre La Distancia Focal De Tu Objetivo”
<http://www.blogdelfotografo.com/distancia-focal/#>
Consultado el 15/11/2016
- [17] “What is difference between multi view stereo (MVS) and structure from motion (SFM) methods in 3D surface reconstruction?” <https://www.quora.com/What-is-difference-between-multi-view-stereo-MVS-and-structure-from-motion-SFM-methods-in-3D-surface-reconstruction>. Consultado el 18/11/2016
- [18] “Using Un-Calibrated Cameras in Photogrammetry”
<http://info.photomodeler.com/blog/tip-61-using-un-calibrated-cameras-in-photogrammetry/>
Consultado el 18/11/2016
- [19] Xiaobing Sun ; Jizhou Sun ; Jiawan Zhang and Mingchu Li. "Simple zoom-lens digital camera calibration method based on EXIF", Proc. SPIE 5302, Three-Dimensional Image Capture and Applications VI, 79 (April 16, 2004)
- [20] Foro de Agisoft (AgisoftForum). Varios autores
<http://www.agisoft.com/forum/index.php?topic=3084.0>
Consultado el 18/11/2016
- [21] David G. Lowe, “Object Recognition from Local Scale-Invariant Features”. Proc. of the International Conference on Computer Vision, Corfu (Sept. 1999)
- [22] Yasutaka Furukawa (University of Illinois at Urbana-Champaign, University of Washington), Jean Ponce (University of Illinois at Urbana-Champaign, Ecole Normale Supérieure). “Patch-Based Multi-View Stereo Software (PMVS – Version 2)”. <http://www.di.ens.fr/pmvs/>
- [23] Yasutaka Furukawa. “Clustering View for MultiStereo (CMVS)”
<http://www.di.ens.fr/cmvs/>
- [24] Escáner 3D BQ Ciclop. <https://www.bq.com/es/ciclop>
Consultado por última vez el 19/11/2016
- [25] Fábrica Digital. Calle Manuel Yélamo Crespillo s/n, Jerez de la Frontera (Cádiz)
- [26] ply_read.m, por Pascal Getreuer.
- [27] “PLY (file format)” [https://en.wikipedia.org/wiki/PLY_\(file_format\)](https://en.wikipedia.org/wiki/PLY_(file_format))
Consultado por última vez el 19/11/2016

- [28] Meshlab. <http://meshlab.sourceforge.net/>
- [29] “Código ASCII” <http://es.ccm.net/contents/55-codigo-ascii>
Consultado por última vez el 19/11/2016
- [30] David G. Lowe, “Distinctive Image Features from Scale-Invariant Keypoints”.
International Journal of Computer Vision, 2004
- [31] David G. Lowe. “Local Feature View Clustering for 3D Object Recognition”. Proc.
of the IEEE Conference on Computer Vision and Pattern Recognition, Kauai,
Hawaii (December 2001)
- [32] Changchang Wu, "Towards Linear-time Incremental Structure From Motion", 3DV
2013
- [33] Changchang Wu, Sameer Agarwal, Brian Curless, and Steven M. Seitz, "Multicore
Bundle Adjustment", CVPR 2011
- [34] Changchang Wu, "SiftGPU: A GPU implementation of Scale Invariant Feature
Transform (SIFT)", 2007
- [35] Diccionario online de la Real Academia Española de la lengua: <http://www.rae.es/>
- [36] Marc Pollefeys “Multiview Stereo”
<http://www.cs.unc.edu/~marc/tutorial/node112.html>

7 AVISO LEGAL

En este Trabajo Fin de Máster se han descrito y empleado varios algoritmos científicos y varios programas informáticos (software's) y dispositivos (hardware's) diferentes. Cada uno puede tener su contrato de licencia y condiciones de uso, asegúrese de que cumple con ellos.

Durante la redacción de la memoria de este Trabajo Fin de Máster se han utilizado nombres de productos registrados y/o marcas registradas. Éstas les pertenecen a sus respectivos propietarios.

La versión empleada de "ply_read.m" viene licenciada bajo la licencia LGPL y es parte separada del programa desarrollado en el trabajo. Está prohibido, excepto previo permiso por escrito del autor, el uso y la distribución del código (programa informático) desarrollado en este trabajo y que emplea la citada versión de "ply_read.m".

No se permite el empleo u obra derivada de esta memoria excepto en los casos autorizados y bajo las condiciones aplicables.

Todo el contenido de esta memoria y del código fuente o de cualquier tipo que sea original de este trabajo se proporciona bajo sus respectivas condiciones aplicables y sin ningún tipo de garantía. Úselo bajo su responsabilidad.

Los resultados de las comparativas expuestas en este trabajo no sólo dependen de la calidad de los algoritmos de reconstrucción de los modelos en tres dimensiones; también intervienen las condiciones de captura, la calidad de la cámara, la configuración de los algoritmos de captura y la precisión de la información proporcionada al programa de análisis (si aplica), entre otros factores.

