

Análisis de Datos 3D Para Generación Automática de Modelos BIM de Interiores Habitados

A. Adán^{a*}, D. Huber^b

^a Departamento de Ingeniería Eléctrica, Electrónica y Automática. Universidad de Castilla La Mancha, C/ Paseo de la Universidad 4, 13071 Ciudad Real, España

^b The Robotics Institute, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA, Estados Unidos.

Resumen

Hasta la fecha, el procesamiento de la información proporcionada por escáners de media distancia en entornos de construcciones civiles y edificios se ha limitado, en la mayoría de los casos, a tareas de registro o generación manual de modelos tridimensionales CAD. En este artículo se da un paso más allá, acometiendo soluciones para problemas de segmentación automática y reconocimiento de partes representativas del entorno como uno de los pasos esenciales hacia la generación automática de modelos BIM (*Building Information Models*) en entornos habitados. Específicamente, se propone un procedimiento para identificar partes esenciales de la estructura de interiores de edificios en entornos altamente desordenados y con un alto componente de oclusión (Figura 1). La dificultad en el tratamiento de millones de puntos inconexos en un espacio no estructurado con fines de inteligencia artificial, hace especialmente atractiva esta línea de investigación, aun no desarrollada en la comunidad científica. El artículo expone una solución a través de etiquetado dinámico y aprendizaje en varias fases que finaliza con la reconstrucción de la superficie básica de interiores (paredes, suelo y techo) y la identificación de partes importantes en el modelo BIM en interiores (puertas, ventanas, armarios, etc). La técnica presentada en este artículo se ha experimentado con éxito sobre datos 3D de edificios proporcionados por empresas profesionales en digitalización con láser escáners de media distancia. Copyright ©2011 CEA. Publicado por Elsevier España, S.L. Todos los derechos reservados.

Palabras Clave: Procesamiento de datos 3D, Visión por Computador, Digitalización, Modelado 3D, Reconocimiento.



Figura 1. Interiores de edificios con un alto grado de desorden y oclusión

* Autor en correspondencia.

Correos electrónicos: Antonio.Adan@uclm.es (A. Adán),
dhuber@cs.cmu.edu (D. Huber)

1. Introducción

La información proporcionada por sensores 3D (*range data* o *points-cloud*) es utilizada en una amplia variedad de entornos. Además de síntesis y modelado de objetos, manufacturado e ingeniería inversa, visión por computador y aplicaciones robóticas, los datos 3D (originalmente proporcionados como millones de puntos, no necesariamente conexos en el espacio) son extensamente utilizados en ámbitos de ingeniería civil y arquitectura. En este contexto, las líneas de trabajo se han centrado, principalmente durante la última década, en modelado de gran escala y evaluación de calidad en construcción de edificios.

La obtención de modelos 3D de edificios durante o después de haber sido construidos (en inglés *as-built models*) es habitualmente realizado por usuarios que manejan programas interactivos de diseño gráfico. Esto supone un tedioso e impreciso proceso, máxime cuando se trabaja en entornos parcialmente ocluidos, en donde no toda la información 3D de los elementos esenciales de la construcción está disponible. En los últimos años han surgido líneas de investigación que pretenden aportar soluciones automáticas para la obtención de Modelos de Información de Edificios (los llamados modelos BIM: *Building Information Models* en nomenclatura anglosajona) en el contexto de ingeniería inversa, utilizando principalmente sensores laser escáner. Desafortunadamente, los sensores laser escáner proporcionan datos en forma de nubes de puntos y es necesario generar una información de alto nivel; como por ejemplo identificación, localización y forma de los muros, puertas y ventanas pertenecientes al edificio. En la práctica, esta información de alto nivel es representada a través de los modelos BIM.

El artículo aquí presentado se sitúa en el contexto de generación de modelos BIM en entornos habitados con varios objetivos. El primer objetivo pretende obtener automáticamente los modelos BIM de las partes esenciales en interiores de edificios (paredes, suelo y techo). En teoría, esta tarea puede parecer simple si se dispone de la nube de puntos del entorno. Sin embargo, en un entorno habitado, donde sólo parte de la superficie de los muros es adquirida por el escáner, el proceso es altamente complejo. Notar que los muros aparecen además habitualmente intercalados con ventanas y puertas, que algunas partes del muro pueden ser ocluidas por muebles u otros objetos (estanterías, relojes, cuadros, etc.) y que multitud de objetos pueden ocluir el punto de vista del sensor (sillas, mesas, etc). Por lo tanto, a pesar de disponer de datos tomados desde diferentes puntos de vista en el interior a modelar, la información de las paredes puede ser significativamente incompleta. La figura 1 muestra entornos complejos de interiores donde nuestra investigación está siendo llevada a cabo.

Tomando en consideración todo lo anterior, el segundo objetivo de nuestra investigación se centrará en aportar soluciones a problemas de oclusión en nubes de puntos, razonando su impacto sobre la generación automática del modelo del interior. Esto conllevará además el desarrollo de métodos de reconocimiento automático de puertas, ventanas y huecos en la pared bajo circunstancias de oclusión parcial.

2. Antecedentes y Contribuciones del Método

En general, el problema de la oclusión puede aparecer ligado a varias áreas de investigación relacionadas con visión desde diferentes puntos de vistas. Por ejemplo, la oclusión en el campo de visualización 3D y gráficos incide sobre la aparición, acceso y relación espacial de objetos en la escena, constituyendo un factor esencial a tener en cuenta. El trabajo publicado por Elmqvist et al. (Elmqvist 2008) define una taxonomía de oclusión 3D en entornos gráficos que puede ser extensible a otros entornos de trabajo en donde se formaliza una terminología común en un marco teórico. En el campo de visión por computador, los problemas de oclusión surgen en un sin fin de aplicaciones con objetivos de segmentación, reconocimiento, análisis e interpretación de escenas, registro de datos, etc. Así mismo, para trabajos de reconstrucción tridimensional utilizando tecnologías LiDAR (*Light Detection and Ranging*), la oclusión es considerada como una de las principales fuentes de incertidumbre y, por consiguiente, es primordial desarrollar técnicas robustas ante oclusión.

Hasta la fecha existen escasos trabajos que acometan el problema de oclusión en imágenes de rango para reconstrucción de interiores, si bien es cierto que hay multitud de artículos que tratan problemas parciales. Así, podemos encontrar soluciones a problemas de detección de zonas planas en nubes de puntos, segmentación 3D, reconstrucción de superficies de objetos ocluidos, tratamiento de oclusiones puntuales o técnicas de rellenado. Sin embargo, ninguna de ellas proporciona una solución integrada en escenas complejas cubriendo todos los aspectos anteriores. En lo que sigue, se muestra una revisión de técnicas actuales sobre oclusión y reconstrucción con datos de rango en entornos de edificios.

La reconstrucción de superficies planas bajo oclusión han sido tratada por F. Dell'Acqua et. al en (Dell 2002). El método está basado en la detección de discontinuidades en imágenes de profundidad obtenidos a partir de los datos de rango. Se realizan fases consecutivas de segmentación de regiones y reconexión, estudiando las posibles zonas que son ocluidas. Este trabajo es extendido posteriormente sobre superficies de fondo esférico y cilíndrico que son parcialmente ocluidas por otros objetos. Aunque se tratan casos simples de oclusión, las superficies ocluidas son finalmente reconstruidas a través de algoritmos de rellenado puramente geométricos.

La mayoría de las soluciones de identificación de regiones ocluidas se basan en procesos de segmentación previos. Por ejemplo, Sappa (Sappa 2002), analiza y recupera escenas formadas por superficies planas y cilíndricas – propias de interiores en fábricas – donde se asume que el modelo de la escena ha sido previamente segmentado. A través de grafos de conectividad entre regiones y distancias al sensor se logra identificar regiones ocluidas y nichos en la escena. El artículo presenta tan solo dos ejemplos en un escenario reducido. F. Han et. al (Han 2002) presentan un procedimiento de segmentación de datos que posteriormente son clasificados según un conjunto de formas. La segmentación es optimizada mediante cadenas de Markov en un entorno de redes Bayesianas. En este trabajo se trabaja con los datos 3D correspondientes a imágenes panorámicas en entornos desordenados, por lo tanto, es uno de los trabajos más próximos al desarrollado en este artículo. Sin

embargo, aunque los autores reconocen que la identificación y reconstrucción de partes ocluidas en los muros podría ser acometido como una extensión del trabajo, no resuelven el problema, exponiendo solamente algunas vagas ideas. En entornos de laboratorio podemos igualmente encontrar algoritmos que tratan de reconstruir un objeto o partes de escenas simples formados con un conjunto reducido de objetos, todo ello a partir de los datos de la escena obtenidos por un escáner. Un trabajo interesante en esta línea es el de U. Castellani et. al. (Castellani 2002) donde se reconstruye objetos poliédricos en escenas con oclusión. El procedimiento está basado en identificar los bordes y esquinas visibles de los objetos e inferir las partes ocluidas, reconstruyendo finalmente las piezas poligonales que conforman cada objeto.

Existen trabajos de restauración automática de edificios que detectan zonas ocluidas mediante procesamiento conjunto de varias imágenes 2D combinados, en ocasiones, con información de rango. En la mayoría de los casos, el objetivo se centra en proporcionar un modelo reconstruido de exteriores. Los artículos (Pospisilova 2007), (Dick 2001) y (Werner 2002) trabajan exclusivamente con imágenes de intensidad, persiguiendo obtener ortofotos simples o panorámicas, por lo que se sitúan en un contexto de investigación distinto al aquí tratado. Detectan partes no visibles (y por lo tanto ocluidas) del edificio mediante técnicas de acoplamiento de imágenes utilizando además una base de conocimiento de propiedades de objetos que aparecen en la escena (Pospisilova 2007). Habitualmente se tienen modelos de partes del edificio que son utilizados para reponer o reconstruir las partes identificadas como ocluidas o no registradas por el sensor. Por ejemplo en (Dick 2001) y (Werner 2002) se reponen ventanas de la misma forma en toda la fachada de un edificio. En otros casos, los métodos son semiautomáticos, donde la región a restaurar es identificada por el usuario, realizándose entonces una mera búsqueda del objeto que va a ser sustituido en la fachada en una base de modelos (Böhm 2007). En todos estos casos, el escenario es simple, con estructuras arquitectónicas repetitivas en todo el edificio.

El problema de oclusión cuando se trabaja con láser escáner está definitivamente presente en la etapa de toma de datos del edificio, donde es necesario situar al sensor en varias posiciones. Como es bien sabido, la colocación óptima del escáner para la próxima vista (NBV, *Next Best View*) es un problema no resuelto aun. El método propuesto en este artículo guarda cierta relación con técnicas NBV en dos sentidos. En primer lugar, nuestro método introduce un modelo de representación discretizado 3D con objetivos de etiquetado del espacio. El segundo punto en común es que el proceso de etiquetado de vóxeles se realiza mediante una integración secuencial de varias tomas del escáner. Este tipo de estrategias pueden encontrarse en trabajos con escáner en (Sanchiz 1999), (Blaer 2007) y (Low 2006-2).

La detección de zonas ocluidas puede ser una de los primeras etapas en algoritmos de rellenados de huecos en entornos tridimensionales. Normalmente este grupo de técnicas son aplicadas sobre modelos de mallas, en aplicaciones de rellenado en esculturas y restauración de piezas arqueológicas o museísticas. Sin embargo, existen algunos trabajos que se desarrollan en escenas de gran dimensión como el presentado por Wang, et.al en (Wang 2002). Se parte de una representación completa de malla con textura de la escena. En la última etapa se

detectan zonas ocluidas sobre objetos de la escena, los cuales han sido previamente segmentados y reconocidos, y se lleva a cabo un algoritmo de rellenado automático solamente para objetos con simetría axial. En ese sentido, la técnica tiene claras limitaciones prácticas.

En comparación con trabajos existentes relacionados con la reconstrucción 3D bajo oclusión, el presente artículo tiene las siguientes contribuciones. En primer lugar, el método es aplicado sobre tomas panorámicas y en entornos habitados. Esto tiene dos implicaciones. La primera es que se cubre todo el espacio alrededor del escáner, por lo que el volumen de información es el máximo y se llega a reconstruir el modelo completo de un interior. La segunda es que, como consecuencia de ser un entorno habitado, la identificación de las zonas ocluidas en los muros requiere de procesos de segmentación y reconocimiento con datos 3D muy complejos. En resumen, la escena tratada, así como los procesos a llevar a cabo sobre la nube de puntos de la escena, son de mayor complejidad que los de entornos de laboratorio, escenarios reducidos y escenarios no habitados.

En segundo lugar, los métodos referenciados tratan el problema de oclusión sobre una sola imagen de rango, mientras que el método presentado en este artículo es acometido sobre un conjunto de ellas, las cuales son integradas en un espacio de representación volumétrico. Hasta ahora, los algoritmos que tratan integración de vistas con escáner se circunscriben solo a contextos NBV, pero no existen antecedentes para problemas de reconstrucción de modelos de edificios en contextos *as-built*.

La tercera aportación del método se basa en generalizar los procedimientos específicos de reconocimiento de objetos existentes en edificios como ventanas, puertas o armarios empotrados. La solución propuesta en este artículo incorpora un concepto general que llamaremos “hueco”, entendido como un espacio vacío que forma parte de los muros y, por tanto, debe ser identificado en el modelo final. El reconocimiento de huecos permitirá, a su vez, identificar y rellenar las zonas ocluidas existentes en los muros.

3. Resumen del método

Supongamos que el objetivo perseguido es la reconstrucción de una determinada sala de un edificio donde existen multitud de objetos desconocidos y desordenados, de la que se han tomado varias tomas panorámicas desde distintas posiciones. Por lo tanto, todos o gran parte de los objetos esenciales del habitáculo (paredes, suelo, techo, ventanas, puertas y armarios) son ocluidos por otros objetos que forman parte del escenario (cuadros, estanterías, cortinas, sillas, mesas, lámparas, etc). En lenguaje anglosajón estas circunstancias se definen como “*clutter and occlusion*”. El objetivo de esta breve sección es dar una visión general del procedimiento seguido; desde que se dispone de los datos 3D hasta que se obtiene el modelo reconstruido de la sala. La figura 2 muestra un esquema del procedimiento dividido en las siguientes fases.

- I Fase de pre-procesamiento. Se realiza un filtrado de datos erróneos y se transforman todas las imágenes de rango obtenidas por el sensor en un sistema de referencia común. Este alineamiento de datos 3D es un proceso que se realiza

manualmente con ayuda de software que el propio escáner posee mediante inserción de patrones que incluyen un código propio. Así, todas las nubes de puntos quedan referenciadas en un espacio 3D que denotamos Γ .

- II Fase de detección de muros. Se introduce una representación volumétrica de vóxeles Ω y se obtienen en ese espacio los planos de vóxeles correspondientes a los muros del interior. En general llamaremos *muro* a las regiones que delimitan el espacio de un interior; es decir paredes, techo y suelo. Cada uno de esos planos se le llama Zona 0.
- III Fase de etiquetado. En esta etapa se lleva a cabo el etiquetado de la Zona 0 desde cada posición del escáner y se aplica un algoritmo de integración de etiquetas en el espacio de vóxeles. Este etiquetado es posteriormente llevado al espacio de puntos Γ , generando una imagen I de cada muro donde cada pixel es originalmente asociado a una etiqueta “vacío”, “ocupado” y “ocluido”.
- IV Fase de segmentación de zonas ocluidas. Se lleva a cabo un algoritmo de aprendizaje SVM (*Super Vector Machine*) para reconocer regiones denominadas “*huevo*” en muros. Reconocidos los *huecos*, se actualiza el etiquetado de la

imagen I acometiendo la segmentación de las zonas etiquetadas como *ocluidas*. Se delimitan las regiones que deben ser restituidas.

- V Fase de reconstrucción. En esta fase final, se realiza el relleno denso de las zonas a restituir y se recupera el modelo original de cada muro y, finalmente, del interior completo.

Con objeto de hacer más comprensible el método seguido, se han introducido diferentes colores en la Figura 2. Cada color representa un espacio diferente de representación (espacio de puntos Γ , espacio de vóxel Ω , espacio de imagen I) que se está utilizando en cada momento a lo largo de proceso.

El resto del artículo está dedicado a presentar cada una de las citadas fases. La sección 4 muestra el proceso de segmentación de los datos en un espacio voxelizado, el etiquetado y la integración de etiquetas desde las diferentes posiciones del sensor. Por su parte, la sección 5 profundiza en el algoritmo de segmentación de zonas ocluidas y en el reconocimiento de *huecos* en el muro. Los resultados experimentales son expuestos en la sección 6 y la sección 7 presenta las conclusiones del trabajo y los futuros avances en la investigación.

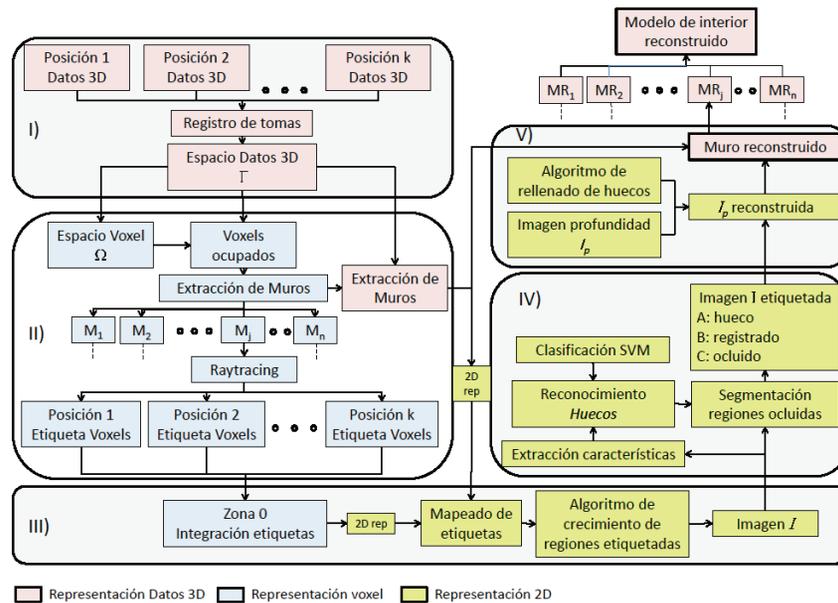


Figura 2. Esquema general de algoritmo de reconstrucción de modelos de interiores

4. Segmentación y Etiquetado de Muros

4.1. Segmentación de datos en un espacio voxelizado

Como se explicó anteriormente, todos los datos disponibles del entorno a reconstruir son inicialmente registrados (alineados) en un único sistema de referencia. En ese momento, los millones de puntos disponibles del escenario

aparecen como simples *items* aislados en el espacio tridimensional Γ . Con objeto de trabajar en un entorno donde los datos estén sometidos a una topología definida, se define un espacio de representación tridimensional volumétrico: el espacio de vóxeles Ω . Se crea así un espacio discretizado en minúsculos cubos, donde se pueden establecer asignaciones de etiquetas.

Un vóxel es etiquetado inicialmente como *ocupado* o *vacío*, dependiendo de si acoge o no a un punto 3D de la escena. Este

proceso de etiquetado inicial lo llamaremos L_0 , donde un vóxel genérico v tiene asignada una etiqueta $l_0(v)$. La figura 3 a) muestra un ejemplo con los datos en los espacios de representación Γ y Ω . Por supuesto la representación Ω discretiza, y por lo tanto desvirtúa, la información original, pero sin embargo hace posible el desarrollo de algoritmos de procesamiento de datos en una representación estructurada y reduce altamente el costo computacional de procesos posteriores.

El primer proceso que se lleva a cabo concierne a la segmentación de los vóxeles que forman parte de los muros de la escena. Algunos autores, como (Sappa 2004) y (Huang 00), han desarrollado técnicas de extracción de regiones planas sobre datos. Nuestra propuesta está inspirada en la desarrollada por Okorn et. al en (Okorn-2010).

En nuestro algoritmo, el espacio de vóxeles ocupados es proyectado según un conjunto de direcciones, generándose para cada una de ellas una imagen z-buffer. Estas imágenes, que llamaremos $I_z(\delta, \lambda)$ - siendo δ, λ los ángulos de proyección - son realmente histogramas bidimensionales, donde el valor asignado a un pixel corresponde al número de vóxeles ocupados que se proyectan en él. La figura 3 b) muestra una visualización de imágenes $I_z(90,0)$ y $I_z(0,0)$.

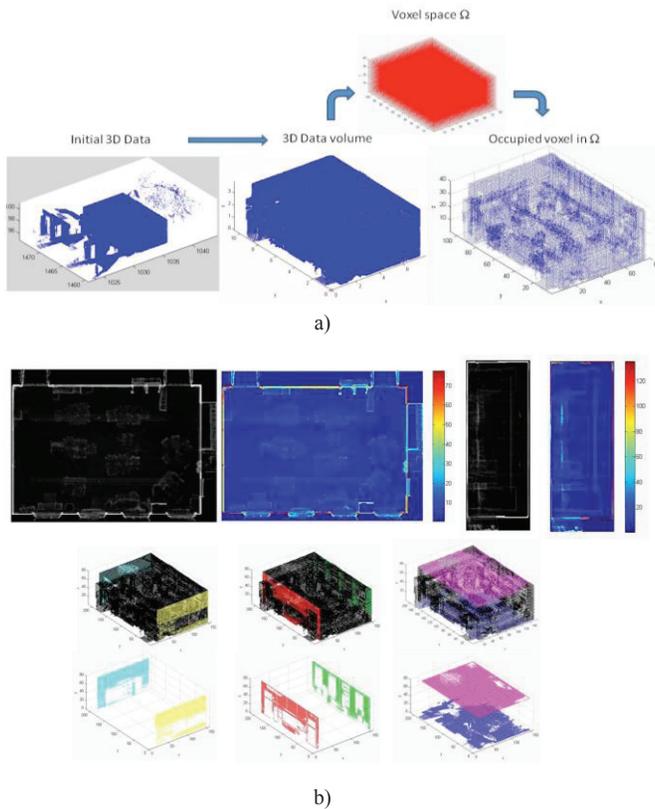


Figura 3. a) Espacio original de datos y espacio discretizado con los vóxeles ocupados b) Obtención de muros en el espacio voxelizado. Arriba) Representación de imágenes $I_z(90,0)$ y $I_z(0,0)$ en escala de grises y color. Abajo) Resultados de la extracción de Zonas 0.

A través de un sencillo proceso de detección de líneas y, posteriormente, de segmentos en imagen, se determinan los muros en la imagen y se deshace la transformación proyectiva, recuperando los correspondientes vóxeles en el espacio Ω y las regiones segmentadas para la sala (Ver figura 3 b). Cada región segmentada es llamada Zona 0 (Z_0), mientras que las regiones planares adyacentes delante y detrás de Z_0 se denotan como Z_{-1}, Z_2, \dots y Z_1, Z_2, \dots respectivamente. Notar que esta representación zonal permite analizar los datos contenidos delante y detrás de los muros en la habitación. La figura 4 presenta en distintos colores y para dos dimensiones de vóxel (5 cm y 10 cm) distintas zonas en la vecindad de un muro segmentado.

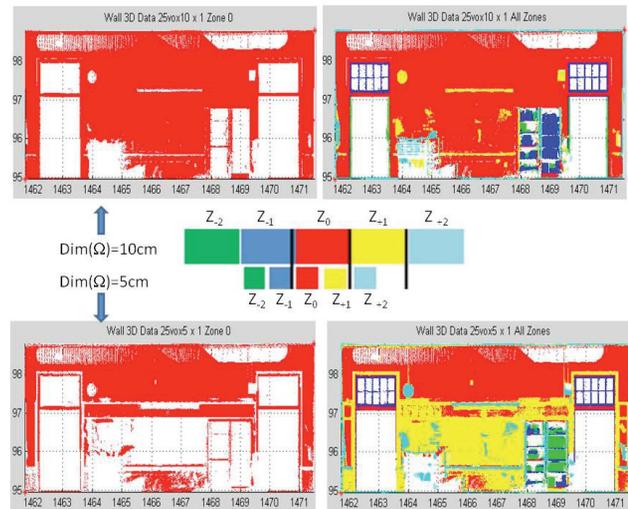


Figura 4. Visualización de zonas $Z_{-2}, Z_{-1}, Z_0, Z_1, Z_2$ para tamaños de vóxel de 10cm y 5 cm.

4.2. Etiquetado de muros.

Una vez delimitada la Zona 0, el objetivo es etiquetar sus vóxeles teniendo en cuenta las posiciones del sensor en la fase de toma de datos.

Sea o_j la j -ésimo posición del escáner en el espacio Ω y $q_1, q_2, q_3, \dots, q_n$ los vértices de cierto muro detectado. El volumen $S_j = \{o_j, q_1, q_2, \dots, q_n\}$ es etiquetado a través de un algoritmo de trazado de rayos (*ray tracing* en nomenclatura anglosajona) desde o_j a cualquier vóxel de Z_0 de la siguiente forma. Los vóxeles ocupados fueron ya etiquetados en la etapa de discretización del espacio (etiquetado L_0) y conservan su etiqueta en el proceso. Supongamos un rayo genérico Π que atraviesa el volumen S_j . Cualquier vóxel atravesado por Π entre o_j y un vóxel ocupado se etiqueta como vacío, mientras que los vóxeles detrás de un vóxel ocupado se etiquetan como ocluidos, incluidos los pertenecientes a Z_0 (ver detalle en Figura 5 a)).

Después de etiquetar un conjunto de volúmenes $S_j, j=1, \dots, k$, cada vóxel v de Z_0 tiene asignadas k etiquetas $\{l_1, l_2, \dots, l_k\}$. El

algoritmo de integración de etiquetas proporciona un etiquetado combinado, que llamaremos L , teniendo en cuenta el etiquetado global inicial L_0 , donde sólo los vóxeles *vacíos* pueden ser actualizados, cambiando a *ocluídos* siguiendo las reglas:

$$l(v) = \{\text{vacío}\} \text{ si } l_0(v) = \{\text{vacío}\} \ \& \ \exists j, l_j(v) = \{\text{vacío}\} \quad (1)$$

$$l(v) = \{\text{ocluído}\} \text{ si } l_0(v) = \{\text{vacío}\} \ \& \ \forall j = 1 \dots k, l_j(v) = \{\text{ocluído}\} \quad (2)$$

Notar que L se refiere al etiquetado conjunto de todo el muro, mientras que l se refiere a la etiqueta de un vóxel. La expresión (1) hace referencia a que un vóxel será *vacío* (en el sentido de que no formará parte de la pared) si, habiendo sido inicialmente etiquetado como vacío, en alguna vista se etiqueta también como *vacío*. Notar que no se necesita que todas las etiquetas sean *vacío*. Por ejemplo, podría suceder una etiqueta *ocluído* desde algún punto de vista. Los vóxeles en el espacio de una puerta son ejemplos de vóxeles *vacío*. En la expresión (2), se etiqueta un vóxel como *ocluído* si desde todos los puntos de vista es *ocluído*. Así, si existe algún punto de vista desde el cual el vóxel no es ocluido (por ejemplo *ocupado*), entonces no puede ser etiquetado genéricamente como *ocluído*.

La figura 5 b) muestra un etiquetado de un muro en el espacio Ω desde distintas posiciones del sensor. Se muestra la imagen de la pared que es vista desde cada posición del sensor y el etiquetado de los cinco volúmenes correspondientes a las mismas posiciones. Las etiquetas $\{A=\text{vacío}, B=\text{ocupado}, C=\text{ocluído}\}$ son visualizadas a través de colores asociados $\{\text{verde}, \text{rojo}, \text{azul}\}$ respectivamente. En la figura 5 c) se presentan, en los mismos colores, el etiquetado final de los vóxeles de la Zona 0 superpuestos a los datos 3D originales pintados en color negro.

Teniendo en cuenta el etiquetado del espacio voxelizado, se realiza un algoritmo de mapeado sobre los puntos originales captados por el sensor. Por lo tanto, se acomete una migración de etiquetado al espacio tridimensional Γ generando una imagen etiquetada de alta resolución I .

La imagen I es generada a través de un algoritmo de crecimiento de regiones con restricciones, utilizando los centros de los vóxeles como semillas de etiquetas en la imagen. Así, antes del algoritmo de crecimiento, la imagen I está formada en su mayor parte por píxeles de etiqueta desconocida y por aquellos que coinciden con los vóxeles etiquetados (A, B, C) , siendo estos últimos los que actuarán como semillas. Por lo tanto, el nombre de píxel '*desconocido*' hace referencia a que su etiqueta está por determinar, precisamente a través del algoritmo de crecimiento que se propone. Finalmente, estas etiquetas D se convierten en alguna de las etiquetas semilla $(A, B$ ó $C)$. En la figura 5 c) se puede apreciar los datos 3D, el conjunto de semillas en colores (que corresponden a las posiciones de los centros de los vóxeles etiquetados) y los espacios en blanco, que son los etiquetados inicialmente como $D=\text{desconocido}$. La figura 5 d) muestra la imagen I etiquetada generada. El algoritmo de crecimiento se describe a continuación.

Sea p_j un píxel de la región de crecimiento O para la

iteración j -ésima del algoritmo y sea $I(p_j)$ la etiqueta correspondiente. Se verifica que:

Para una región *ocluída*

$$\begin{aligned} p_0 \in O \text{ si } I(p_0) = C \\ p_j \in O \text{ si } I(p_j) \in \{C, D\} \end{aligned} \quad (3)$$

con la restricción:

$$\begin{aligned} \min \{ d(p_j, p), \forall I(p) = A \} > \alpha \\ \& \ \min \{ d(p_j, p), \forall I(p) = C \} < \beta \end{aligned} \quad (4)$$

Para una región *vacía*

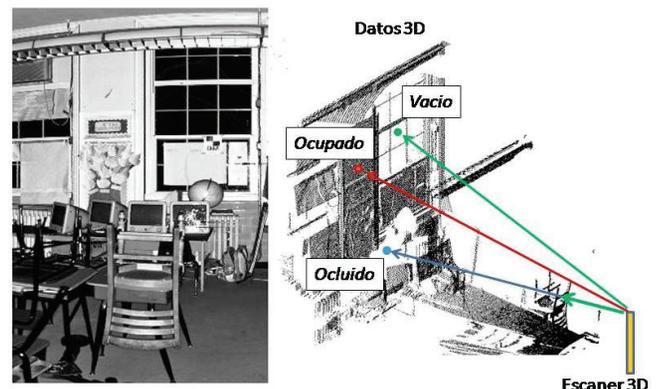
$$\begin{aligned} p_0 \in O \text{ si } I(p_0) = A \\ p_j \in O \text{ si } I(p_j) \in \{A, D\} \end{aligned} \quad (5)$$

con la restricción:

$$\begin{aligned} \min \{ d(p_j, p), \forall I(p) = C \} > \alpha \\ \& \ \min \{ d(p_j, p), \forall I(p) = A \} < \beta \end{aligned} \quad (6)$$

donde $d(p_j, p)$ es la distancia euclídea entre píxeles p_j y p en la imagen I . Los umbrales α y β han sido empíricamente establecidos en función de la resolución de la imagen I . La única relación que existe entre ambos umbrales es que $\beta > \alpha$, es decir la búsqueda del mismo tipo de etiqueta se realiza dentro de un radio máximo que es mayor que el radio mínimo a partir del cual se realiza la búsqueda de distinto tipo de etiqueta.

Como resultado del algoritmo de crecimiento, cada uno de los puntos 3D asociados a cada píxel de la imagen I posee una etiqueta definida $\{A=\text{vacío}, B=\text{ocupado}, C=\text{ocluído}\}$. Notar que la imagen I proporciona un etiquetado denso en el muro en función de la resolución deseada. En nuestro caso, la densidad fue de 2x2cm de superficie de pared por píxel de la imagen I .



a)

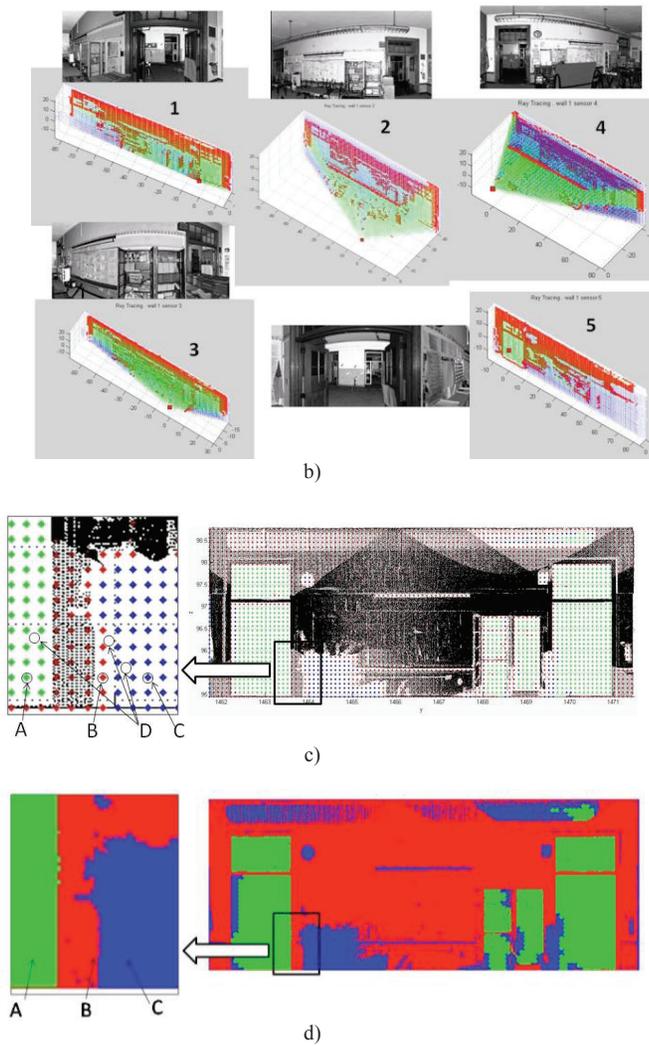


Figura 5. a) Detalle de raytracing en un sector de datos tomados desde una posición del sensor. b) Imágenes de un muro desde cinco posiciones del sensor y etiquetado del volúmenes S_j en colores, verde=vacio, rojo=ocupado, azul=ocluído c) Integración de etiquetas de vóxeles de la Zona 0 (puntos de colores) con los datos 3D superpuestos (puntos en color negro) d) Generación de la imagen etiquetada densa I mediante el algoritmo de crecimiento de regiones. Cada pixel equivale a una superficie de 2cmx2cm.

5. Segmentación de regiones ocluidas

5.1. Planteamiento del problema.

Con objeto de llevar a cabo la reconstrucción densa de la pared representada en la imagen etiquetada I , las zonas ocluidas deben ser segmentadas. El objetivo es identificar qué parte de una región *ocluída* (C) pertenece a la estructura de la pared y cuál no. En la Figura 6 se muestra un claro ejemplo del problema a resolver. Sobre la imagen etiquetada de la

pared se marca un recuadro donde se aprecia una región etiquetadas como *ocluída* (pintada en azul) que cubre parte del hueco de una ventana. La oclusión de la ventana se debe a un panel y un conjunto de papeles que están pegados a los cristales. Por supuesto, el hueco de la ventana no forma parte de la pared y, por consiguiente, es un espacio que no deberá ser rellenado al reconstruir el muro en la fase final. Como se demuestra en este ejemplo, la segmentación de zonas ocluidas en entornos tan desordenados puede ser un problema de difícil solución.

La estrategia de segmentación aquí propuesta se basa en reconocer y delimitar primero las aberturas existentes en un muro mediante un algoritmo de aprendizaje y después realizar la segmentación.

De aquí en adelante, cualquier tipo de abertura en un muro (ventana, puerta, armario, etc) se llamará genéricamente *hueco*. En un contexto más específico, la detección de ventanas o puertas ha sido un tópico común, mayoritariamente tratado para aplicaciones con robots móviles, como SLAM, navegación y posicionamiento autónomo ((Lee 04) (Muño 06) (Chen 08)). La mayoría de los trabajos existentes trabaja sobre imágenes 2D tomadas por cámaras portadas por el robot y raramente se consideran circunstancias de oclusión o entornos tan complejos. Uno de los trabajos recientes que utiliza datos de rango es el publicado por Ali et. al en (Ali 08) en el que se reconocen ventanas en la escena, pero en entornos sin oclusión.

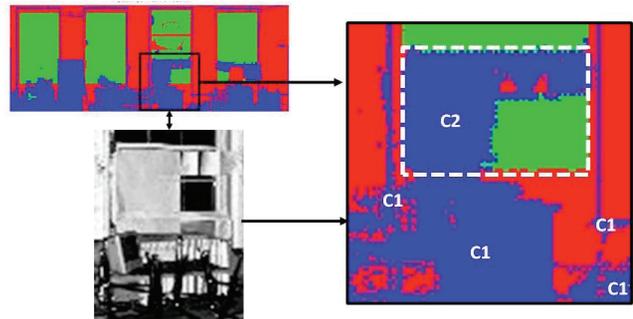


Figura 6. Detalle de regiones ocluidas en una pared. Las zonas etiquetas con C1 corresponderían a regiones ocluidas pertenecientes al muro, mientras que las etiquetadas con C2 son oclusiones sobre huecos en la pared.

5.2. Definición de características

Como hipótesis inicial sostenemos que la forma de los huecos existentes en los muros de un edificio son mayoritariamente rectangulares o quasi-rectangulares, habiendo en ocasiones pequeñas desviaciones. Por ejemplo, las esquinas podrían ser redondeadas o truncadas. En cualquier caso, los huecos podrían estar siempre circunscritos y ajustados razonablemente a rectángulos. En nuestro caso, consideraremos éste como el caso estándar, dejando para investigaciones de futuro otras formas de huecos menos usuales.

Con el objetivo de encontrar posibles *huecos* candidatos en un muro detectado y delimitado en fases previas, se ha definido un vector de características V sobre cualquier rectángulo extraído después de procesar la imagen de profundidad de la Z_0 (que denotamos como I_p). Recordar que Z_0 delimita los puntos pertenecientes a la pared. La imagen de profundidad I_p , es dual a la imagen etiquetada I y, por tanto, de igual dimensión.

Para un rectángulo R en la imagen I_p , las componentes del vector de características son definidos de la siguiente manera:

$$v_1 = \dim_1(R) * \dim_2(R) \quad (7)$$

$$v_2 = \frac{\dim_1(R)}{\dim_2(R)} \quad (8)$$

$$v_3 = \frac{\dim_1(R)}{\dim_1(I)}, v_4 = \frac{\dim_2(R)}{\dim_2(I)} \quad (9)$$

$$v_5 = x(R), v_6 = \dim_1(R) - x(R) \quad (10)$$

$$v_7 = y(R), v_8 = \dim_2(R) - y(R) \quad (11)$$

$$v_9 = \sqrt{\frac{\sum_{i=1}^n [d(p_i, \Pi)]^2}{n(n-1)}} \quad (12)$$

$$v_{10} = \frac{\sum I(A)}{v_1}, v_{11} = \frac{\sum I(B)}{v_1}, v_{12} = \frac{\sum I(C)}{v_1} \quad (13)$$

$$v_{13} = \arg\{R_j, R_j \subset R\} \quad (14)$$

En estas expresiones \dim significa dimensión. Las componentes v_1 a v_8 expresan magnitudes absolutas o relativas geométricas del rectángulo candidato, en concreto: área (v_1), la relación Anchura/Altura (v_2), dimensiones relativas con respecto a las dimensiones de la pared (v_3 y v_4) y distancias de los lados a los bordes de la pared (v_5 a v_8). La componente v_9 es un coeficiente de planitud (*flatness coefficient*) del área que encierra el rectángulo. Se define como el error cuadrático medio de las distancias de los puntos interiores del rectángulo al plano Π que mejor los ajusta (ver ejemplo en la Figura 7 a)). Las componentes v_{10} a v_{12} corresponden a la proporción de etiquetas A , B , y C que existen dentro del rectángulo en la imagen etiquetada de la pared. Por último, el componente v_{13} es el número de rectángulos interiores al rectángulo candidato. Debido a los gradientes de profundidad que suelen encontrarse en las proximidades de *huecos* - por ejemplo debidos a los marcos de puertas o ventanas - es posible que en el proceso de detección de líneas en la imagen I_p , se encuentren rectángulos contenidos. Esta circunstancia es tenida en cuenta por la componente v_{13} . La Figura 7 b) ilustra un ejemplo donde se han detectado dos rectángulos contenidos.

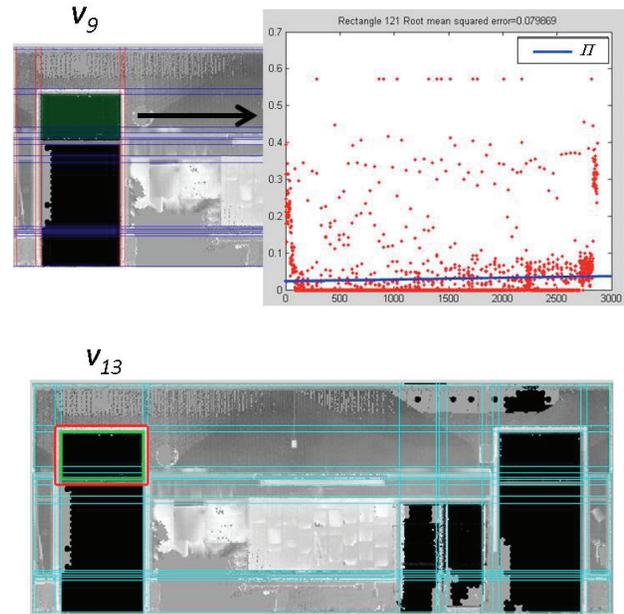


Figura 7. a) Ejemplo de un rectángulo candidato marcado sobre la imagen de profundidad de la zona 0, I_p . y gráfico de posición de los puntos interiores y plano de ajuste Π . b) Ejemplo de detección de rectángulos contenidos en la imagen.

5.3. Reconocimiento de huecos

El algoritmo de reconocimiento de huecos ha sido llevado a cabo a través de un algoritmo de clasificación SVM (Support Vector Machine) (Vapnick 1998) en el cual se utiliza el vector de características $V = \{v_1, v_2, \dots, v_{13}\}$. Como es sabido, SVM es una de las técnicas de aprendizaje supervisado y clasificación binaria que ofrece mejores resultados. En nuestro caso, se utilizó el *kernel* RBF (*Radial Basis Function*) para un modelo SVM binario. La base de aprendizaje se realizó sobre un conjunto de paredes del edificio donde se realizaron test posteriores del método. Ya que se trata de la reconstrucción del modelo *as-built* del edificio, los modelos del proyecto arquitectónico no pueden ser utilizados como información de referencia (*ground-truth* en la nomenclatura anglosajona) del algoritmo de aprendizaje. Notar que, precisamente uno de los objetivos en la extracción de los modelos *as-built* BIM es detectar las variaciones que han sufrido los edificios respecto de su modelo proyectado. Por lo tanto, el *ground-truth* se llevó a cabo marcando manualmente los *huecos* en un conjunto de salas de un edificio de tres plantas.

En las circunstancias de alto desorden y oclusión existentes en entornos habitados, los requerimientos exigidos a los rectángulos candidatos extraídos en la imagen I_p deben ser los mínimos. Por lo tanto, el proceso de imagen no tiene por objetivo detectar segmentos que puedan formar parte de los bordes de los hipotéticos huecos, ya que estos pueden estar parcial o incluso totalmente ocluidos. El procedimiento se basa en detectar líneas en la imagen utilizando el algoritmo de

Hough (Hough 1959), calcular todos los posibles rectángulos delimitados por dichas líneas y aplicar posteriormente procesos de clasificación y agrupamiento (*clustering*). De este modo, huecos con altos grados de oclusión donde incluso uno de sus bordes esté totalmente ocluido, podrá ser reconocido al final del proceso. Por lo tanto, implícitamente se está utilizando una estrategia de reconocimiento no basada estrictamente en propiedades geométricas, sino extendida a otras características existentes. En nuestro caso, las componentes v_9 a v_{13} del vector V de características, no corresponden a características geométricas.

Debido a la laxitud en la consideración de rectángulos candidatos, el algoritmo SVM proporciona aun un número alto de rectángulos clasificados como *huecos* en las primeras etapas del proceso de reconocimiento. Notar que, por ejemplo, en los marcos de ventanas o puertas existen discontinuidades de profundidad que generan conjuntos de líneas próximas en la imagen I_p . A su vez, esta circunstancia lleva a que se detecten varios rectángulos reconocidos en las proximidades del hueco. La Figura 8 muestra un detalle de los rectángulos clasificados positivamente alrededor de una ventana superior y puerta.

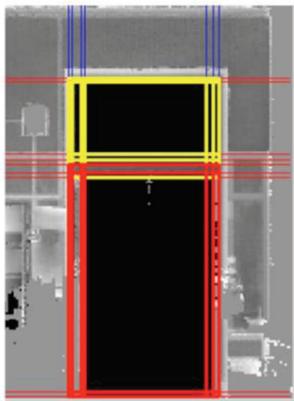


Figura 8. Detalle de múltiples rectángulos clasificados como *hueco*.

A través de un algoritmo de agrupamiento, en nuestro caso un algoritmo *k-means* sobre los centroides de los *huecos* con restricciones de tamaño, se determinan grupos de rectángulos alrededor del definitivo *hueco*. El número óptimo de grupos es calculado mediante un algoritmo iterativo que encuentra un mínimo global de la suma de distancias en los componentes del grupo. Finalmente, para cada grupo obtenido se calcula un rectángulo prototipo que será considerado definitivamente como la región *hueco*. La Figura 9 b) muestra un ejemplo en la ejecución de la fase de agrupamiento donde aparecen grupos de rectángulos en el mismo color. La Figura 9 c) muestra las regiones *hueco*.

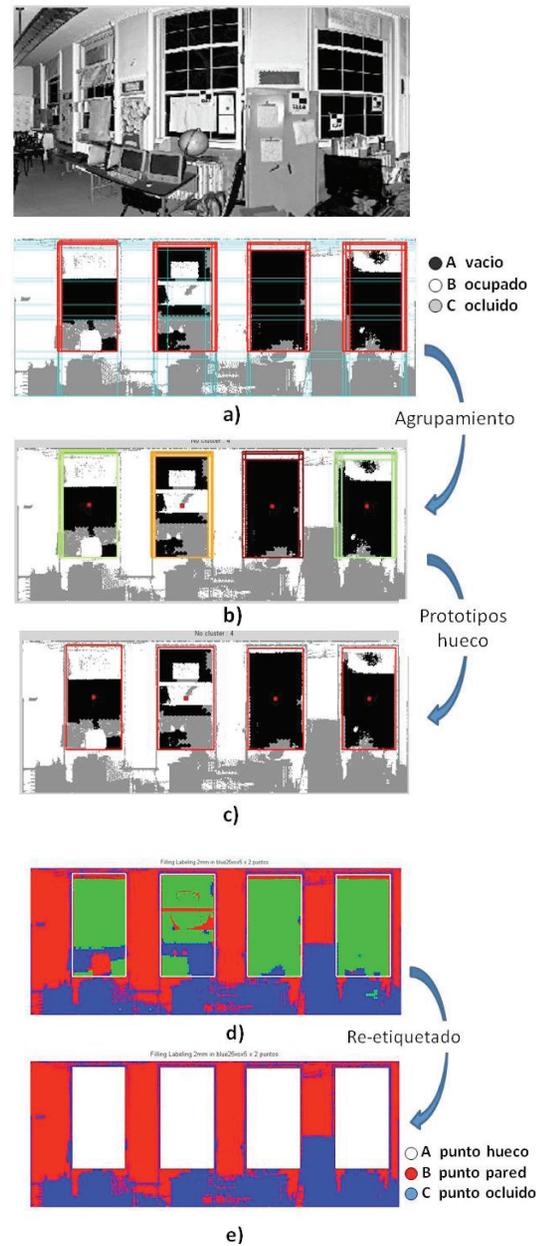


Figura 9. a) Imagen planar de intensidad del muro (arriba) y huecos candidatos dibujados sobre la imagen después de aplicar el algoritmo de clasificación SVM (abajo). b) Grupos de *huecos* obtenidos en el algoritmo de agrupamiento en distintos colores. c) Prototipo de regiones *hueco* finalmente definidas. d) Superposición de los bordes de las regiones *hueco* sobre la inicial imagen etiquetada. e) Etiquetado final.

5.4. Etiquetado final

Retomando el objetivo inicial de esta sección, la segmentación de las regiones ocluidas en la imagen etiquetada I_p puede ser fácilmente llevada a cabo después de identificar los *huecos* en el muro. La Figura 9 d) muestra los rectángulos *huecos* calculados en la fase de reconocimiento superpuestos

la imagen etiquetada *I*. Como consecuencia de la segmentación de la zona ocluida, se reestructura definitivamente el etiquetado de un punto en tres casos. La etiqueta A corresponde a los espacios pertenecientes al *hueco* obtenido en la fase de reconocimiento. Los puntos de la región A los llamaremos definitivamente *puntos hueco*. La etiqueta B conserva su sentido inicial aunque con un matiz. No solamente son puntos captados por el sensor, sino puntos que pertenecen a la estructura de la pared, por eso ahora los llamaremos *puntos pared*. Finalmente, la etiqueta C también conserva su sentido de espacios ocluidos pero, al igual que en el caso anterior, son pertenecientes a la estructura de la pared. Por lo tanto, estos espacios son los que deberán ser reconstruidos para generar la estructura completa de la pared y los llamaremos *puntos ocluidos*.

6. Resultados Experimentales del Método

El método de análisis y reconstrucción de interiores presentado en este artículo ha sido desarrollado utilizando datos 3D proporcionados por una empresa dedicada a escaneado de construcciones civiles. La toma de datos, se realizó en un edificio habitado de tres plantas. Las imágenes presentadas en la figura 10 muestran la alta complejidad de las escenas tratadas. La información proporcionada en cada toma consistió en una imagen planar panorámica de la escena, la correspondiente imagen de rango y las coordenadas tridimensionales de la nube de puntos. En esta investigación, la imagen de reflectancia y la imagen de rango (que proporciona la profundidad desde la posición del sensor) se han desechado para el propósito que se perseguía. Sin embargo, se contempla que puedan ser de utilidad para otras estrategias de solución en el futuro. Por consiguiente, solamente se utilizaron las nubes de puntos captadas por el escáner desde diferentes posiciones en la escena de interior.



Figura 10. Imágenes planares de distintas salas sobre las que se ha aplicado el método de reconstrucción de interiores. En la última fila se señalan ejemplos de patrones codificados que son utilizados para registrar distintas vistas.

El alineamiento de datos 3D es un proceso que se realizó manualmente utilizando de patrones que son fácilmente detectados en la imagen planar de la escena (Figura 10). Para una toma dada, se identifica un grupo de patrones en la imagen planar y el usuario marca sus centros obteniéndose sus

coordenadas tridimensionales (referidas al sistema de coordenadas del sensor). Para dos posiciones del sensor, se establecen manualmente un conjunto de correspondencias entre estos puntos centro y se calcula la matriz de transformación entre los dos sistemas de referencia. Este mismo procedimiento se sigue para las siguientes posiciones del sensor, habiendo tomado como sistema universal la primera posición del sensor. De esta manera, todos los datos quedan registrados en un sistema de referencia común y las posiciones de los sensores son igualmente conocidas. Este proceso de registro es uno de los habituales en el campo de la digitalización de interiores o exteriores en edificios.

La densidad promedio del láser escáner proporciona tomas de 50 millones de puntos por toma. Debido al inasumible costo computacional y requerimientos de memoria con tan ingente resolución espacial, los algoritmos utilizados en nuestro método trabajaron con una nube de puntos muestreada a 1/25. En cualquier caso, se demostró durante el procesamiento de datos que existían zonas de las paredes con muy baja densidad de puntos o zonas no convenientemente registradas, a pesar de disponer de varias tomas en la escena. Esta circunstancia de irregularidad en la densidad de los datos influyó notablemente en la calidad de los resultados finales.

El espacio fue muestreado para varios tamaños de vóxeles. Aquí presentamos los resultados para vóxel de 5 y 10 centímetros de arista y para espacios de 400 m³, lo que supone disponer de aproximadamente 340.000 y 2.700.000 vóxeles respectivamente por sala tipo. El tamaño real de los muros está dentro del intervalo 25 m² a 78 m², con una resolución de 2 cm/píxel en la imagen etiquetada *I*, por lo que la reconstrucción volumétrica del muro fue de celdas de 2cm³.

La Figura 11 a) y e) muestran ejemplos de los datos 3D originales y la reconstrucción de los muros, para tamaños de vóxeles de 5 y 10 cm. Como puede apreciarse, las partes del muro que son ocluidas por muebles u otros objetos han sido reconstruidas. La parte b) muestran el resultado del etiquetado final dibujando en blanco los *huecos* reconocidos en la pared. En c) se muestran los puntos originales de la pared en negro y los puntos rellenados en las zonas *ocuidas* en color rojo. En d) se muestra una representación renderizada del muro reconstruido sin considerar los huecos calculados y en e) después de haber extraído las regiones de huecos. Notar que el tamaño del vóxel puede tener una influencia en el resultado final. Por ejemplo una discretización en vóxeles muy pequeños puede proporcionar peores resultados, ya que algunos puntos del muro pueden caer fuera de la Zona 0 y, por lo tanto, no son tenidos en cuenta en el resto del proceso. Esto puede ocurrir para muros con deficiente planitud. Por el contrario, para tamaños demasiado grandes, la Zona 0 puede contener puntos que no pertenecen a la misma estructura del muro, como paneles o cuadros colgados en la pared. Por lo tanto, la optimización del tamaño y origen del espacio discretizado Ω es una problema que deberá ser resuelto en el futuro con objeto de mejorar el método propuesto.

La tabla 1 presenta información concerniente a datos de discretización para un conjunto de test realizado sobre una de las salas del edificio. Se incluye el número original de puntos disponibles en la Zona 0 (*N3D*), el tamaño del muro obtenido

en m^2 (SW), la precisión en cm de la imagen etiquetada I (AI), sus dimensiones (DIM) y el total de los rayos trazados para etiquetar el muro en la etapa de *raytracing* (NR).

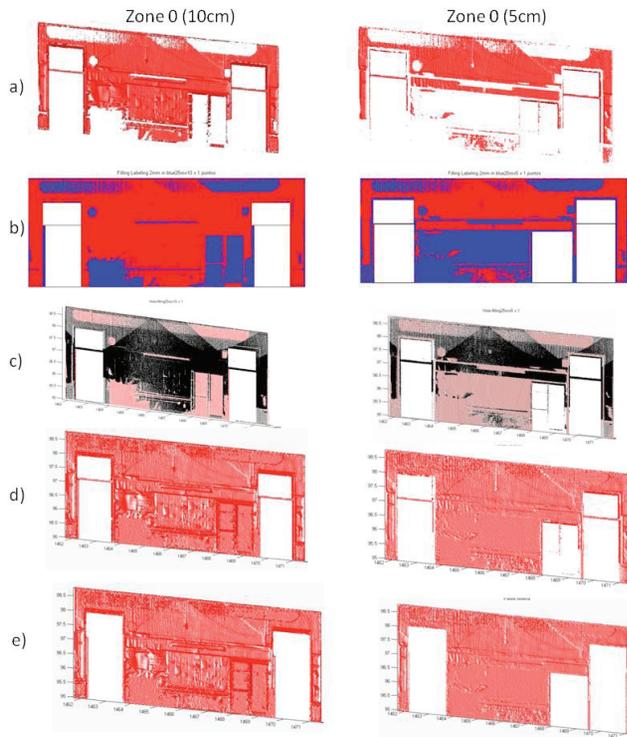


Figura 11. Reconstrucción del muro #1 para tamaños de vóxeles de 5 y 10 cm. a) Modelo renderizado inicial de la Zona 0 extraída. b) Imagen etiquetada final I . c) Visión de las regiones rellenadas en el muro en color rojo. d) Representación renderizada de la pared rellenada. e) Reconstrucción de la pared habiendo extraído los datos pertenecientes a los huecos reconocidos.

Para cuantificar el grado de éxito en la detección de oclusiones se contabilizó en número de regiones existente en cada muro y se contrastó con las presentes en la imagen etiquetada I . La Figura 12 a) muestra diversas partes de la imagen de reflectancia donde se enumeran las regiones que ocluyen el muro, mientras que en la Figura 12 b) se presenta la imagen etiquetada obtenida y se identifican las regiones anteriores. De esta comparación se obtiene el número de regiones ocluidas ciertas y falsas. La Tabla 2 muestra los resultados de oclusión obtenidos para los muros del interior considerado en nuestro test.

Tabla 1. Información de datos y etiquetado de muros para el conjunto de test.

Vox 10cm	$N3D$	SW	AI	DIM	NR
#1	352485	37.7508	0.02	489x193	19305
#2	165317	37.6736	0.02	488x193	19800
#3	152529	25.862	0.02	335x193	13600
#4	194265	25.9392	0.02	336x193	13600
#5	253832	66.1128	0.02	489x338	34155
#6	182694	78.0192	0.02	516x378	39520

Vox 5cm	$N3D$	SW	AI	DIM	NR
#1	191556	37.5552	0.02	489x192	75460
#2	140990	37.6736	0.02	488x193	76440
#3	152467	25.862	0.02	335x193	52650
#4	190525	25.9392	0.02	336x193	52650
#5	244926	65.6476	0.02	487x337	133280
#6	244926	65.6476	0.02	487x337	157320

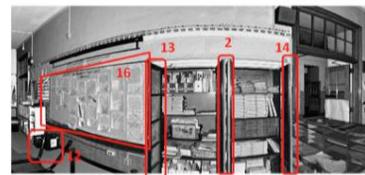
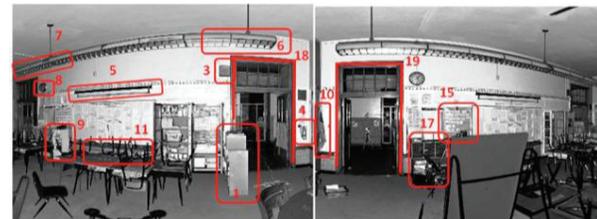


Figura 12. a) Objetos ocluyentes de la pared 1 señalados sobre la imagen de intensidad b) Regiones ocluidas correspondientes en la imagen etiquetada I .

Tabla 2. Resultados de oclusión en el conjunto de test. NT (Numero de regiones ocluidas), NTD (Numero de regiones ocluidas detectadas ciertas), NFD (Numero de regiones ocluidas detectadas falsas), C (porcentaje de superficie ocluida en la Zona 0)

Vox10cm	Vox 10 cm				Vox 5 cm			
	NT	NTD	NFD	C	NT	NTD	NFD	C
#1	17	12	2	9,9	17	18	1	29,02
#2	16	14	1	26,9	16	16	1	26,22
#3	5	5	1	38,6	5	5	1	43,72
#4	10	9	2	18,4	10	9	2	19,69
#5	-	-	-	2,6	-	-	-	2,64
#6	-	-	-	29,2	-	-	-	30,67

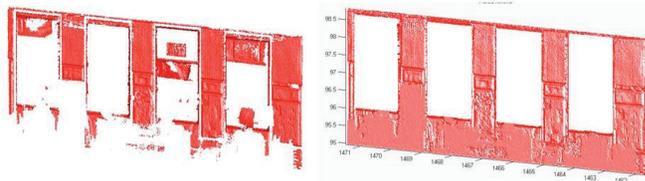
Los resultados de reconocimiento de huecos en los muros está resumida en la tabla 3. Se muestran en columnas el número de huecos existente, el número de verdaderos positivos, falsos positivos y falsos negativos. También se incluyen los porcentajes de las etiquetas finales en el muro.

A pesar de la dificultad en algunos puntos de la escena, el algoritmo funcionó correctamente en la determinación de huecos. Por ejemplo, nótese que en la figura 13 a) los huecos correspondientes a las ventanas son bien reconocidos a pesar de que existen zonas etiquetadas inicialmente como ocupadas en el interior de algunas ventanas. Estas zonas ocupadas

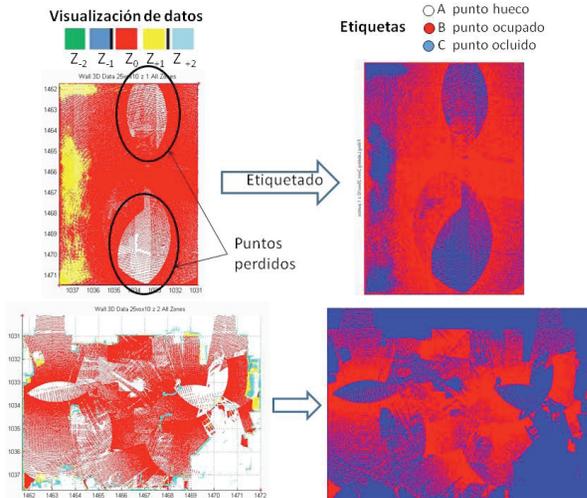
corresponden a persianas, un compresor de aire acondicionado embutido en el cristal o a papeles pegados en los cristales. Por otro lado, el algoritmo también funcionó para casos negativos (#5 (suelo) y #6(techo)) donde no se detectaron huecos a pesar de existir regiones inicialmente etiquetadas como vacías. Estas zonas aparecen como consecuencia de una baja densidad de puntos registrados en suelo y techo. La Figura 13 b) muestra a la izquierda una representación de las zonas Z_{-2} , Z_{-1} , Z_0 , Z_1 , Z_2 para el suelo y techo. Notar que las zonas blancas corresponden a ausencias de datos, lo que puede ser debido a oclusiones o a problemas de acceso del escáner. En el techo se marcan dos zonas donde claramente no hay puntos registrados por el sensor. En la parte derecha de la figura se muestra el etiquetado final donde se aprecia la ausencia de huecos.

Tabla 3. Detección de huecos y etiquetado final: NG (Numero de huecos), NTP (Numero de verdaderos positivos), NFP (Numero de falsos positivos), NTN (Numero de ciertos negativos). Los valores de A, B y C corresponden a los porcentajes de superficie etiquetada

Vox 10cm	NG	NTP	NFP	NTN	A	B	C
#1	5	4	0	1	21,0	57,0	21,0
#2	4	4	0	0	39,0	36,0	23,0
#3	0	0	0	0	0	50,7	49,3
#4	2	2	0	0	30,5	49,5	19,9
#5	0	0	0	0	0	68,9	31,1
#6	0	0	0	0	0	46,1	53,9
Vox 5cm	NG	NTP	NFP	NTN	A	B	C
#1	5	5	0	0	13,5	35,9	32,5
#2	4	4	0	0	41,5	34,6	24,3
#3	0	0	0	0	0	50,5	49,5
#4	2	2	0	0	30,6	49,4	19,8
#5	0	0	0	0	0	68,8	31,1
#6	0	0	0	0	0	44,1	55,8



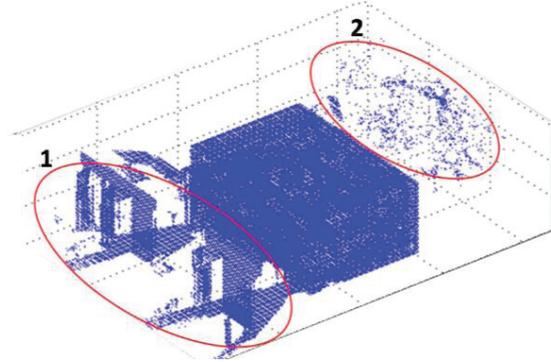
a)



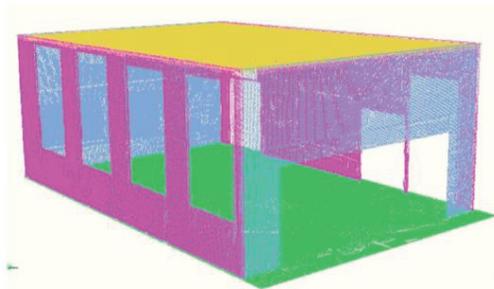
b)

Figura 13. a) Casos de reconocimiento de huecos que contienen puntos 3D no pertenecientes a la estructura de la pared. Los datos proceden de objetos de pequeño grosor (papeles o cortinas) cercanos al plano de la pared. Estos objetos son registrados por el escáner y asignados a la Zona 0 en el proceso de detección de paredes. b) Casos de ausencia de huecos. A pesar de la existencia de zonas con baja densidad de datos (que podrían ser pensadas como hipotéticos huecos), el algoritmo funciona correctamente. A la derecha aparece el etiquetado final en donde solo existen puntos pared y puntos ocluidos.

El trabajo presentado en este artículo está centrado exclusivamente en la tarea de etiquetado y reconstrucción de muros en interiores de edificios. La Figura 14 puede representar la entrada y la salida final del proceso. La entrada está proporcionada por la información 3D que capta el escáner desde distintas posiciones. La Figura 14 a) muestra las nubes de puntos alineadas en un mismo sistema de referencia. Notar que, ya que las puertas de la sala están abiertas, el sensor capta información de zonas externas a la sala (región 1). Igualmente se aprecia cómo el escáner recoge puntos a través de los cristales de las ventanas (región 2). En la parte b) de la figura se muestra el modelo geométrico denso recuperado de la sala después de aplicar la propuesta aquí presentada.



a)



b)

Figura 14. Ilustración de la entrada y salida del proceso de obtención de modelos de interiores de edificios. a) Imagen de reflectancia de la sala y datos captados por el escáner de media distancia. b) Modelo geométrico de la sala después de integrar los muros recuperados.

7. Conclusiones y Futuros Avances

La reconstrucción automática de modelos de información geométrica BIM es uno de las líneas de investigación más actuales en el campo de visión por computador 3D con láser escáner densos. El trabajo presentado supone uno de los pasos esenciales hacia la automatización en la ingeniería inversa de construcciones. Específicamente, se pretende aportar soluciones de reconstrucción de modelos densos de interiores en entornos habitados, lo que implica una alta complejidad en las escenas a tratar.

El método propuesto se desarrolla en varias etapas de etiquetado utilizando como entrada el conjunto de nubes de puntos que el sensor adquiere desde varias posiciones. Debido a las circunstancias de oclusión y desorden de la escena, se proponen soluciones de segmentación de datos 3D y reconocimiento de regiones, llegando a identificar lo que se han llamado *puntos hueco*. Finalmente, el método es capaz de delimitar las regiones ocluidas en la estructura construida de los muros, aplicar un algoritmo de rellenado sobre las mismas y generar finalmente el modelo denso del interior.

Los avances inmediatos de esta investigación están dirigidos en varios sentidos. Por una parte que quiere ampliar la semántica de los modelos reconstruidos, incorporando la clasificación de huecos detectados en entidades menores (por ejemplo ventanas, puertas, armarios, etc) así como reconocimiento de estructuras de detalle en las construcciones (marcos o molduras). Por otra parte, con objeto de mejorar la eficiencia y robustez del método, se estudia la posibilidad de incorporar la información de reflectancia.

El problema de zonas con baja densidad o zonas desprovistas de datos, puede ser un inconveniente en este tipo de trabajos. Una posible solución de futuro iría encaminada a realizar un muestreo no uniforme en la fase de pre-procesado. Se podría fijar, para cada vóxel, la densidad mínima de puntos contenidos y realizar un muestro adaptado a dicha densidad. El principal problema sería el enorme costo computacional y de memoria que este proceso supondría

La obtención del modelo de representación de alto nivel del edificio completo en entornos *as-built*, así como la evaluación de la disparidad entre el modelo recuperado comparada con el modelo 3D proyectado y la detección de taras detectadas en la construcción del edificio, serán igualmente objeto de futuros desarrollos y publicaciones.

English Summary

3D Data Processing: Towards the Automated BIM in Inhabited Indoors

Abstract

In this paper we present a method for the reconstruction of interiors using a set of panoramic range data in scenes with clutter and occlusion. Although the ultimate goal of our project is to obtain automated Building Information Models of

facilities, here we specifically deal with the reconstruction of simply-shaped wide areas (such as walls, ceilings and floors) behind furniture and facility pieces in interiors. Our approach is based on a sequential updating labeling strategy in different data representation spaces. A volumetric representation is used to permit the labeling of the 3D space for different range data and the fusion of all the scene's labels to obtain one single 2D labeling image for each of the simply-shaped wide areas of the room. Based on this labeling process, our method is able both to identify the occluded regions in the wall and, through a learning technique, to infer essential parts, such as doors, windows and closets. This method was tested under difficult clutter and occlusion conditions, yielding promising results. Scans were obtained using a state of the art laser scanner operated by a professional 3D scanning service provider.

Keywords: 3D Data Processing, Computer Vision, Digitizers, 3D Modeling, Recognition

Agradecimientos

Este trabajo ha sido realizado bajo financiación del National Science Foundation (NSF) de Estados Unidos en el proyecto "Automating the Creation of As-built Building Information Models" CMMI-0856558, desarrollado en la Universidad de Carnegie Mellon, Pittsburgh, PA, USA.

Referencias

- Ali H., Ahmed B. P. Gerhard, 2008. Robust Window Detection from 3D Laser Scanner Data. Congress on Image and Signal Processing, Vol 2, 115-118. Sayna/Hainan, China.
- Blaer P.S. and Allen P.K., 2007. Data Acquisition and View Planning for 3D Modeling Tasks. IEEE/RSJ International Conference on Intelligence Robots and Systems, (IROS'07). Vol 1, pp 417-422. San Diego, CA.
- Bohm J., Becker S., Haala N., 2007. Model Refinement by Integrated Processing of Laser Scanning and Photogrammetry. Institute for Photogrammetry, Universitaet Stuttgart.
- Castellani U., Livatino S., Fisher R., 2002. Improving Environment Modelling by Edge Occlusion Surface Completion. Proceeding of the Third International Symposium on 3D Data Processing, Visualization and Transmission (3DPVT'02), pp 672-675. Padova. Italy.
- Chen Z. and Birchfield S.T., 2008. Visual Detection of Lintel-Occluded Doors from a Single Image. First International Workshop on Internet Vision pp1-8. Anchorage, AK, USA, June.
- Dell'Acqua F., Fisher R., 2002. Reconstruction of Planar Surface Behind Occlusions in Range Images. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 24, 569-575.
- Dick A.R., Torr P.H., Ruffle S.J. and Cipolla R., 2001. Combining Single View Recognition and Multiple View Stereo for Architectural Scenes. International Conference Computer Vision, pp 268-280. Vancouver, BC, Canada.
- Elmqvist N. and Tsigas P., 2008. A Taxonomy of 3D occlusion Management for Visualization. IEEE Transactions of Visualization and Computer Graphics, Vol 14, No 5, 1095 -1109.
- Han F., Tu Z., Zhu S., 2002. A stochastic Algorithm for Scene Segmentation and Reconstruction. ECCV 2002. Lecture Notes in Computer Science 2352m, pp 502-51.
- Hough P.V.C., 1959. Machine Analysis of Bubble Chamber Pictures, Proceedings of High Energy Accelerators and Instrumentation.

- Huang H., Lee A. and Mumford D., 2000. Statistics of Range Images. IEEE International Conference on Computer Vision and Pattern Recognition pp 324-331 Hilton Head Island, SC, USA, June 2000.
- Lee J., Doh N.L., Chung W.K., You B.J., Youm Y.I., 2004. Door Detection Algorithm of Mobile Robot in Hallway Using PC Camera. Conf. ISARC 2004, Korea.
- Low K.L. and Lastra A., 2006. Efficient Constraint Evaluation Algorithms for Hierarchical Nex-Best-View Planning. Proceeding of the Third International Symposium on 3D Data Processing, Visualization and Transmission (3DPVT'06), pp 830 - 837. University of North Carolina, Chapel Hill, USA, June 14-June 16.
- Muñoz Salinas R., Aguirre E., Garcia-Silvente M., 2006. Detection of Doors Using a Genetic Visual Fuzzy System for Mobile Robots. *Autonomous Robot* (2006) 21: 123-141.
- Okorn B.E., Xiong X., Akinci B., and Huber D., 2010. Toward Automated Modeling of Floor Plans. Proceedings of the Symposium on 3D Data Processing, Visualization and Transmission. Espace Saint Martin, Paris, France, May 17-20.
- Pospisilova R., 2007. Occlusion Detection and Surface Completion in 3D Reconstruction of Man-Made Environments. 15th International Conference Central Europe on Computer Graphics, Visualization and Computer Vision, pp 133-140. University of West Bohemia, Campus Bory, Plzen - Bory, Czech Republic.
- Sanchiz J.M., Fisher R.B., 1999. A next-best-view algorithm for 3D scene recovery with 5 degrees of freedom. *BMVC99*, pp 163-172. Nottingham, UK.
- Sappa A.D., 2004. Automatic Extraction of Planar Projections from Panoramic Range Images. 2nd International Symposium on 3D Data Processing, Visualization and Transmission. pp 231-234.
- Vapnik V.N., 1998. *Statistical Learning Theory*. John Wiley & Sons, Inc.
- Wang J. and Oliveira M., 2002. Improved Scene Reconstruction from Range Images. *Computer Graphics Forum* Vol 31 (3), pp 521-530.
- Werner T., Zisserman A., 2002. Model Selection for Automated reconstruction from Multiple Views. *British Machine Vision Conference*, Cardiff, UK, pp 53-62.