

ESTIMACIÓN DEL VOLUMEN DE MERCANCÍAS EN LOGÍSTICA MEDIANTE MÚLTIPLES CÁMARAS DE RANGO

Pablo Artaso y Gonzalo López-Nicolás

Instituto de Investigación en Ingeniería de Aragón - Universidad de Zaragoza, España

artaso_pablo@hotmail.com, gonlopez@unizar.es

Resumen

En este artículo se presenta un sistema para la estimación del volumen de la mercancía transportada por una transpaleta. Para ello, se ha utilizado la información proporcionada por un conjunto de sensores de rango basados en distintas tecnologías como son la de luz estructurada y la de tiempo de vuelo. Se ha realizado un estudio de la precisión y exactitud de cada una de las cámaras para poder compararlas y conocer mejor su funcionamiento ante las variaciones de las condiciones de trabajo. Así mismo, se ha estudiado su comportamiento al trabajar conjuntamente con otros sensores de rango. La principal contribución de este trabajo es el desarrollo de los algoritmos que permiten estimar el volumen de la mercancía a partir de la información proporcionada por los diferentes dispositivos trabajando simultáneamente. La propuesta ha sido validada experimentalmente en un conjunto de situaciones reales, lo que ha permitido comprobar la validez del sistema diseñado.

Palabras clave: Sensores de rango, luz estructurada, tiempo de vuelo, RGB-D, volumen.

1 INTRODUCCIÓN

Una de las áreas más importantes de cualquier tipo de empresa, independientemente del sector en el que desarrolle su actividad, es la logística. Este campo ha cobrado especial importancia en las últimas décadas con el proceso de globalización experimentado. Por este motivo, se busca el ahorro de costes y de tiempo en cada uno de los procesos logísticos. Una consecuencia de esto es la progresiva incorporación de la tecnología.

En los últimos años han sido desarrollados sensores de rango de bajo coste que permiten su amplia utilización en el campo de la investigación entre otros, y cuya consecuencia ha sido el desarrollo de nuevas técnicas para el procesamiento de nubes de puntos. Algunos ejemplos de este tipo de dispositivos son los sensores RGB-D como la Asus Xtion Pro Live. Sin embargo, el uso de esta clase de sensores no está aún ampliamente extendido

en el campo de la industria por las carencias en cuanto a robustez y protección que presentan, al estar concebidos como controladores para juegos. Hay otro tipo de dispositivos que están siendo introducidos paulatinamente en algunos procesos industriales cuya finalidad es la misma, la de obtener información de rango, pero que lo hacen de una manera muy diferente. Algunos ejemplos de este segundo grupo son las cámaras Mesa SR 4000 e IFM PDM 3D sensor.



Figura 1: Imágenes de los dispositivos utilizados.

En este contexto, este artículo presenta un sistema que mediante la utilización simultánea de varios dispositivos de los ya mencionados, es capaz de calcular el volumen de la mercancía colocada encima de un palé que está siendo transportado por una transpaleta. Éste es un método de transporte ampliamente utilizado tanto para la carga y descarga de camiones o furgonetas como para transportar mercancía, por lo que la aplicación del método desarrollado puede suponer un ahorro de costes y de tiempo y un mejor control logístico. En la actualidad, existen algunos sistemas para calcular el volumen como los presentados en [6] o en [11]. Otro algoritmo que también permite calcular el volumen y que también ha sido evaluado, aunque obteniendo resultados poco satisfactorios, es [3]. Sin embargo, hasta ahora no se han desarrollado sistemas similares al diseñado en este trabajo capaces de estimar volúmenes de carga a partir de diferentes tipos de sensores de rango trabajando simultáneamente. Los dispositivos con los que se ha trabajado son una Asus Xtion Pro, una Asus Xtion Pro Live, dos Mesa SR 4000 y dos IFM PMD 3D (fig. 1). Los dos primeros dispositivos están basados en luz estructurada. Los dos últimos tipos de sensores se basan en la tecnología de tiempo de vuelo [10]. Al tratarse de sensores que obtienen la información a partir de tecnologías

diferentes, los datos disponibles para trabajar son muy distintos.

Sin pérdida de generalidad, se ha considerado que la carga cuyo volumen se quiere estimar está formada por cajas de cartón de diferentes tamaños, ya que es el modo habitual de almacenamiento. También se ha supuesto por motivos de seguridad y por las limitaciones de los sensores que no existen huecos internos entre las mercancías ni elementos en voladizo. En la figura 2 se muestra una imagen tomada durante una prueba del sistema. En ella se puede observar la colocación de los dispositivos en el entorno de trabajo y, en la pantalla situada lateralmente, la interfaz de usuario desarrollada mostrando los resultados así como lo que registra cada dispositivo en tiempo real.

Un aspecto importante es la integración de toda la información proporcionada por los distintos tipos de dispositivos. Para ello se ha trabajado con el concepto de nube de puntos, ya que ofrece una gran versatilidad para trabajar con datos en 3D. La herramienta escogida para trabajar con las nubes de puntos es la librería de uso libre “Point Cloud Library” (PCL)[12]. También se ha utilizado la librería de uso libre OpenCV para el procesamiento de imágenes 2D.



Figura 2: Entorno en el que se han realizado las pruebas del sistema desarrollado.

2 ANÁLISIS Y COMPARACIÓN DE LOS DISPOSITIVOS

Al trabajar con diferentes tipos de sensores, uno de los aspectos claves ha sido el conocer y poder comparar cómo es su funcionamiento al variar algunas condiciones como el material con el que se trabaja o la distancia de trabajo. Así mismo, al utilizar los dispositivos simultáneamente, es importante conocer si existen interferencias. El estudio realizado de la precisión y exactitud de los sensores ha permitido conocer en mayor detalle sus prestaciones y ha resultado de gran ayuda para saber cómo deben ser colocados espacialmente para obtener unos mejores resultados.

Para la realización del estudio se han empleado distintos materiales como una caja de cartón, otra de plástico, un palé de madera, otro de plástico, papeles y cartulinas de diferentes colores, un plástico de embalaje, otro de invernadero o una pantalla de portátil. Además, se han realizado las pruebas a dos distancias diferentes para ver cómo varían los datos obtenidos. En concreto, dichas distancias son de 1,5 y de 3 metros. Mediante el estudio de la precisión se ha podido conocer la repetibilidad de los dispositivos, mientras que mediante el de exactitud se ha podido observar el error cometido al estimar la distancia a la que se encuentran los objetos. A modo de ejemplo, en la figura 3 se muestra un diagrama de barras con el error cometido al estimar la distancia a la que se encuentran algunos de los materiales comentados anteriormente cuando están a 1,5 metros. Una recopilación exhaustiva de los resultados obtenidos puede encontrarse en [1].

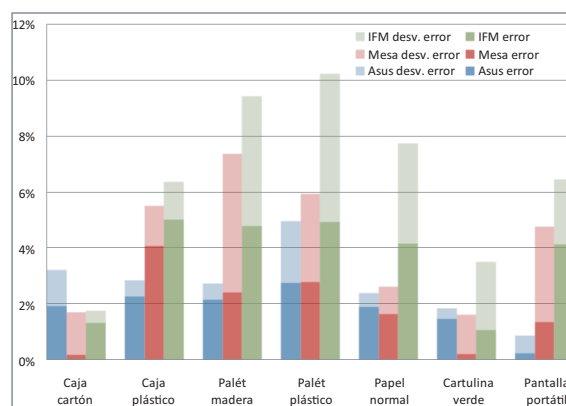


Figura 3: Error de estimación de distancia de los elementos empleados en el estudio cuando la distancia de trabajo es de 1,5 metros.

Como se puede observar, la Asus proporciona resultados muy similares independientemente del objeto utilizado. Esto mismo se ha observado también al analizar la precisión, de lo que se deduce que la tecnología de luz estructurada se ve menos afectada por los cambios en los materiales de trabajo. Sin embargo, las Asus se ven afectadas en mayor medida ante el aumento de la distancia de trabajo. De hecho, los datos tomados mediante la Mesa y la IFM mejoran a la distancia de tres metros respecto de la de un metro y medio. Una de las razones de este comportamiento al aumentar la distancia es que la tecnología de tiempo de vuelo se basa fundamentalmente en medir el tiempo que tarda el rayo emitido en regresar al sensor. Actualmente, se puede medir este tiempo con mucha exactitud, siendo incluso más fácil calcularlo cuando este tiempo es mayor. Otro motivo es el objetivo tan diferente para el que fueron concebidos estos dispositivos. La cámara Asus está

pensada inicialmente como un controlador de juegos, para lo cual no es primordial la calidad de los datos obtenidos. Sin embargo, las otros dos son cámaras industriales, por lo que garantizan una mayor precisión y exactitud para las aplicaciones en que son normalmente utilizadas.



Figura 4: A la izquierda la nube de puntos que representa una cartulina vista desde la Asus. En las otras dos imágenes se muestran resultados de la Asus al registrar la cartulina trabajando simultáneamente con la IFM.

Una de las conclusiones extraídas de la realización del estudio es que el dispositivo de IFM afecta en gran medida a la repetibilidad de la Mesa, pero que esto no se produce a la inversa. Esto puede ser solventado mediante una correcta sincronización de las cámaras. Sin embargo, también se han detectado fallos en el comportamiento de las Asus al trabajar a la vez que las IFM. Un ejemplo se muestra en la figura 4. Además, la IFM no presenta un mejor comportamiento en cuanto a exactitud y precisión que los otros dispositivos. A esto se le une que su resolución es considerablemente menor, ya que una nube de puntos tomada mediante ella presenta sólo 3.200 puntos frente a los 307.200 puntos de las Asus. Por todos estos motivos se ha decidido no utilizar las dos IFM PMD 3D en el sistema diseñado.

2.1 CONFIGURACIÓN ESPACIAL DE LOS SENSORES

Un punto importante del sistema a desarrollar es la colocación de los sensores. Esto determina aspectos claves como la distancia a la que trabajan los dispositivos, cómo son las nubes de puntos adquiridas o cómo es la información que se representa en ellas. También hay que tener en cuenta la necesidad de que se produzca un cierto solapamiento entre los elementos situados dentro del campo de vista de cada dispositivo para poder realizar el registro 3D.

Algunas de las decisiones a la hora de cómo colocar las cámaras han venido condicionadas por las conclusiones extraídas del estudio de precisión y exactitud. En dicho estudio se ha comprobado cómo la distancia a la cual trabaja cada dispositivo afecta a la calidad de los datos obtenidos. En concreto, la Asus proporciona mejores resultados cuando menor es la distancia de trabajo. Sin embargo, la Mesa proporciona mejores datos desde el punto de vista de la precisión y la exactitud cuando trabaja a mayores distancias que la Asus.

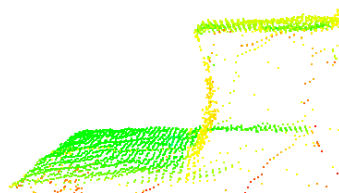


Figura 5: Defecto debido a múltiples reflexiones en una nube de puntos de la Mesa colocada con una orientación inadecuada. Nótese como el lateral de la caja parece ser curvo cuando debería ser plano. Además, la mayoría de los puntos que representan dicho lateral se concentran en la parte inferior.

Por estos motivos, al tener que colocarse todos los dispositivos en una estructura metálica en forma de arco adecuado para su uso en un pasillo, se ha decidido colocar las Asus en los laterales de dicha estructura y la Mesa en la parte superior. Al estar los laterales más cercanos a la carga que el techo, se favorece que las cámaras trabajen a unas distancias de trabajo óptimas.



Figura 6: Configuración escogida de los dispositivos. Las dos Asus en los laterales y la Mesa SR 4000 en la parte superior. Las Asus registran la información de los laterales exclusivamente y la Mesa proporciona una vista de la parte superior de la escena, siendo su eje z perpendicular al plano del suelo. Las cámaras han sido señaladas en la imagen con círculos de color verde.

Una vez que se ha determinado la colocación de los dispositivos, falta establecer la orientación con la que enfocan a la carga. En el caso de la Mesa se ha decidido enfrentarla al suelo, es decir, que su eje z se corresponda con la normal al plano del suelo. Esto facilita algunos aspectos del algoritmo desarrollado ya que proporciona una vista cenital de la mercancía y además su eje z se corresponde

directamente con la altura de los objetos. Cabe decir que el propio fabricante recomienda orientar la cámara perpendicularmente a los elementos a medir ya que en caso contrario se pueden dar problemas de reflexiones múltiples como se muestra en la figura 5. Las cámaras Asus se encuentran orientadas hacia los laterales de la mercancía pero ligeramente inclinadas ya que se encuentran situadas a una altura mayor (fig. 6).

El único inconveniente que presenta esta forma de colocar las cámaras, es que no se dispone de información ni de la parte delantera ni de la trasera de los elementos a calcular el volumen. Sin embargo, esto no es un problema ya que es suficiente la información de los laterales y de la parte superior del conjunto para poder calcular su volumen.

3 SEGMENTACIÓN DE ESCENA

En primer lugar, es necesario unificar los datos bajo un único formato. El escogido es el *pcd*, que es el que utiliza la librería PCL para trabajar con nubes de puntos, lo cual facilita su manipulación. La propia librería permite manejar las cámaras Asus y a su vez se encarga de la conversión de los datos a este formato. Sin embargo, hay que implementar el código tanto para controlar el dispositivo de Mesa como para realizar la transformación de los datos obtenidos al formato anterior.

Pese a que todas las nubes con las que se trabaja están bajo el mismo formato, éstas son muy diferentes ya que los dispositivos de las que provienen toman los datos de diferente manera. Un ejemplo de esto es el tipo de información que proporcionan las cámaras. En el caso de la Asus, se dispone de un mapa de profundidad y de una imagen de color de la escena. Superponiendo ambas imágenes, se obtiene la nube de puntos 3D en color con la que se va a trabajar. Sin embargo, la Mesa no lleva incorporada cámara RGB, por lo que no proporciona información de color. En su caso, se dispone de un mapa de profundidad y de una imagen de intensidad, obtenida a partir de la diferencia de energía entre el rayo infrarrojo emitido y el recibido al haber sido reflejado en los distintos elementos situados dentro de su campo de visión. Mediante la unión de ambas informaciones se obtiene una nube de puntos 3D con información de intensidad.

Como ya se ha comentado, el objetivo es realizar el cálculo del volumen de una mercancía que está siendo transportada por una transpaleta. Sin embargo, en la escena puede haber mucha información que no es útil para llevar a cabo este cometido: suelo, la propia transpaleta, el palé, etc. Por tanto, resulta necesario el procesar las nubes de puntos para eliminar todos estos elementos.

3.1 EXTRACCIÓN DEL SUELO

Para extraer el suelo se ha empleado un algoritmo ampliamente utilizado como es el RANdom SAMple Consensus (RANSAC) [7]. Sin embargo, en vez de aplicar este algoritmo cada vez que una nube de puntos es tomada por cualquiera de los dispositivos, se ha decidido *calibrar* el plano del suelo mediante el RANSAC la primera vez que son colocadas las cámaras en el emplazamiento de trabajo. De este modo, se puede asegurar que no haya ningún objeto en la escena que no sea estático, es decir, que sólo aparezca el suelo y las paredes. Así se logra que el número de puntos espurios durante el cálculo de la ecuación del plano sea muy bajo y que el plano principal de la escena sea el suelo. Posteriormente, cuando se desee extraer el plano del suelo para calcular un volumen durante el funcionamiento normal del algoritmo, únicamente es necesario eliminar todos aquellos puntos que pertenezcan al plano previamente calculado. Esta forma de eliminar el suelo tiene sentido ya que, una vez colocadas las cámaras, el plano del suelo es siempre el mismo, por lo que no resultaría eficiente el aplicar el RANSAC cada vez que se desea extraer este elemento.

3.2 EXTRACCIÓN DE LA TRANSPALETA Y EL PALÉ

Aunque el suelo es fácilmente identificable, segmentar el resto de elementos puede resultar más complicado. Para poder separar el palé de la mercancía, se ha decidido utilizar la información adicional que proporcionan los dispositivos, la intensidad y el color. Esto ha sido posible gracias a las características del palé utilizado, que está formado por polímeros. Este tipo de palés están siendo cada vez más utilizados ya que son reciclables, con el consiguiente ahorro que esto conlleva. De este modo, características como el color son fáciles de modificar. En el caso del palé empleado, su color es negro, por lo que resulta sencillo identificarlo. A su vez presenta una superficie tratada con un alto grado de specularidad y muy lisa, lo que provoca que la intensidad reflejada en la Mesa sea muy baja en comparación con el resto de elementos.

Por tanto, para las nubes tomadas mediante la Asus de color basta con realizar un filtro de color para eliminar la mayor parte del palé. En el caso de la Mesa, el procedimiento es análogo pero sobre la información de intensidad. Al no disponer de información adicional la cámara Asus Xtion Pro, no se puede utilizar ningún filtro de los anteriores. Por lo que se ha decidido medir la altura a la que se encuentra el palé y eliminar todos los puntos que se encuentren a una altura inferior respecto del plano del suelo.

Una vez eliminado el palé, sólo quedan en las nubes de puntos ciertos elementos secundarios junto a la mercancía como parte de la transpaleta o del operario en caso de que apareciese. Para extraerlos, se ha decidido emplear una técnica muy utilizada en el procesamiento de imágenes 2D que es la búsqueda de “blobs” o agrupaciones pero aplicada a 3D [9]. De todas las agrupaciones encontradas, aquella con mayor número de puntos será la formada por la mercancía. En las figuras 7 y 8 se muestran dos ejemplos de procesamiento de nubes de puntos para las cámaras de Asus y de Mesa respectivamente.

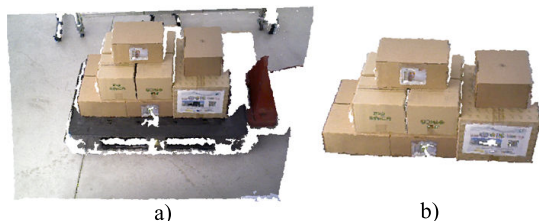


Figura 7: Ejemplo de procesamiento de una nube de puntos adquirida por la Asus Xtion Pro Live.

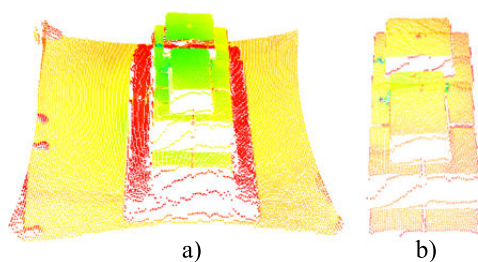


Figura 8: Ejemplo de procesamiento de una nube de puntos adquirida por la Mesa SR 4000.

4 REGISTRO 3D

Una vez que se dispone de todas las nubes procesadas, se tiene toda la información necesaria para calcular el volumen. El inconveniente es que el sistema de referencia de cada dispositivo es distinto, por lo que es necesario juntar todas las nubes bajo un único punto de vista, es decir, realizar un registro (reconstrucción) 3D. El fusionar la información se reduce a encontrar la matriz de transformación que permite cambiar del sistema de coordenadas de una cámara al de otra. Éste es un problema que ha sido ampliamente estudiado [13]. Los pasos a realizar pueden variar en función del tipo de nubes de puntos empleadas y de cómo sea la escena que éstas representan.

La solución adoptada para lograr el registro 3D de la escena sin que afecte a la robustez del sistema consiste en realizar una única calibración inicial de las cámaras. Es decir, cuando se instala por

primera vez el sistema, el propio instalador es el encargado de seleccionar los puntos característicos y las correspondencias a utilizar. A partir de éstas, se calcula una primera aproximación de la matriz de transformación. Para ello, se utiliza el algoritmo desarrollado por Arun en 1997 [2] basado en la descomposición en valores singulares.

La matriz de transformación así obtenida se utiliza como semilla del algoritmo “Iterative Closest Point” [4]. Al ser las correspondencias introducidas por el usuario correctas, la semilla utilizada está muy próxima a la solución final, lo que facilita que el ICP converja al mínimo error global y lo haga en muy pocas iteraciones.

De esta forma, se evita el tener que hacer todo el proceso típico de un registro 3D automático cada vez que se calcula el volumen, con el consiguiente ahorro en lo que a tiempo de cómputo se refiere. Esta manera de actuar es posible ya que una vez instaladas las cámaras, las posiciones relativas son siempre las mismas. Por lo tanto, las matrices de transformación que relacionan los diferentes puntos de vistas de las cámaras son constantes. De este modo, una vez que son calculadas sólo es necesario aplicar el cambio de referencia a las nubes de puntos adquiridas por cada uno de los dispositivos.

El calcular las matrices de transformación de forma automática cada vez que se toman nubes de puntos se ha desechado aparte de por el tiempo de cómputo necesario porque no ofrecía suficiente robustez. Esto se debe a que la escena a registrar está compuesta normalmente por cajas. Por lo tanto, la mayoría de los elementos geométricos existentes son planos perpendiculares y paralelos de aspecto muy similar. Además, tanto desde el punto de vista del color como de la intensidad registrada por la Mesa, resulta muy homogéneo todo el conjunto. Esto dificulta considerablemente la obtención de adecuados puntos característicos y de buenas correspondencias al no haber zonas con grandes gradientes o especialmente características. Otra de las razones es que se están utilizando sólo tres cámaras para lograr registrar toda la escena, por lo que el cambio de vista de un dispositivo a otro es grande. De hecho, se está intentado reconstruir un arco de 180° con sólo tres cámaras.

5 CÁLCULO DEL VOLUMEN

A continuación se describen los métodos desarrollados para el cálculo del volumen de la carga.

5.1 MÉTODO VOXEL

Inicialmente, se pensó en dividir toda la nube de puntos en una cuadrícula. De este modo, se

podrían dar tres tipos diferentes de cubos dentro de la cuadrícula: cubos del interior de los objetos a calcular el volumen que no presentan ningún punto en su interior, cubos que presentan puntos en su interior porque están situados en las superficies de los elementos y cubos situados fuera de la nube de puntos que no tienen ningún punto en su interior. De este modo, si se contabilizasen únicamente los cubos del interior y de la superficie, siendo conocidas sus dimensiones, el cálculo del volumen sería inmediato. Sin embargo, el poder diferenciar aquellos cubos del interior de los del exterior de la nube puede resultar complejo, ya que ninguno de ellos presenta puntos en su interior.

El algoritmo implementado finalmente parte de la idea de dividir la nube de puntos en cubos, pero, en vez de contabilizar aquellos que pertenecen a la nube, lo que se hace es sumar prismas, resultando en una “integración” del volumen. Partiendo del plano del suelo y de la nube de puntos que representa un registro 3D de los elementos que se desea calcular el volumen, los pasos a realizar para calcular el volumen V se exponen a continuación (consultar [1] para más detalle):

- Alinear el eje z de la nube con la normal al plano del suelo.
- Buscar en toda la nube de puntos el punto cuya distancia al plano del suelo sea mínima ($DistMin$), que se corresponde con la cara del palé donde se apoyan las cajas.
- Dividir la nube en celdas (*voxels*), quedando definida la nube únicamente por los centroides de dichas celdas ($LC = LadoCelda$).
- Mientras no se hayan procesado todos los puntos de la nube de puntos original, ejecutar el siguiente bucle:
 1. Se busca el punto de la nube con mayor distancia al plano del suelo que no haya sido procesado aún ($APA = AlturaPuntoActual$).
 2. Se añade al volumen el prisma correspondiente a ese punto:

$$V+ = LC^2 * (APA - DistMin); \quad (1)$$
 3. Se marca como procesada la celda a la que pertenece este punto, así como a todas las celdas situadas debajo.

5.2 MÉTODO PROYECCIÓN DE PLANOS

El método anterior es válido para cualquier tipo de carga. Sin embargo, uno de los modos más habituales de transportar la mercancía es mediante cajas. Este es el motivo por el que todas las pruebas

experimentales han sido realizadas con cajas. Si se realiza la hipótesis de que toda la carga a transportar está empaquetada de esta forma, se puede hacer uso de esta información para desarrollar un nuevo método para calcular el volumen. Al estar formadas las cajas exclusivamente por planos, se puede calcular el volumen mediante su proyección vertical. En concreto, el volumen del conjunto se corresponde con lo que hay debajo de los planos paralelos al suelo, como se observa en la figura 9. Por lo tanto, sólo hace falta conocer todos los planos de la nube y, de los que sean paralelos al suelo, calcular su área. Siendo conocida su altura respecto del palé, se puede computar fácilmente el volumen. Los pasos a realizar para calcular el volumen V mediante este método son:

- Buscar en el registro 3D el punto cuya distancia al plano del suelo sea mínima ($DistMin$), que se corresponde con la cara del palé donde se apoyan las cajas.
- Buscar todos los planos de la nube de puntos de la Mesa SR 4000, que representa la vista cenital.
- Ordenar todos los planos en función de su distancia al origen (la propia cámara).
- Mientras no se hayan procesado todos los planos encontrados, se ejecuta el siguiente bucle:
 1. Extraer el plano (plano referencia) de la nube situado más cercano al origen que no haya sido computado aún ($APR = Altura Plano Referencia$).
 2. Calcular su área ($A = Area$).
 3. Si hay algún plano (plano actual) por debajo del de referencia, calcular el volumen hasta este plano inferior ($APA = Altura Plano Actual$).

$$V+ = A * (APR - APA); \quad (2)$$

En caso de que no queden más planos por debajo, calcular el volumen hasta la base:

$$V+ = A * (APR - DistMin); \quad (3)$$

Una dificultad de este algoritmo es el cálculo del área, ya que en ocasiones los planos pueden presentar irregularidades y zonas con concavidades. Existen fórmulas para calcular el área de polígonos irregulares a partir de los vértices si están ordenados. Sin embargo, esta información no es conocida a priori.

La solución adoptada consiste en transformar la nube de puntos 3D en una imagen 2D. Una vez que

se dispone de ella, es procesada mediante técnicas de visión por computador. En primer lugar, se le aplican algoritmos de “eroding” y “dilating”. La segunda coloca puntos nuevos alrededor de los ya existentes originalmente, de modo que la variación en el interior de la imagen es mínima. Sin embargo, los contornos de la imagen se ven desplazados (es como si la imagen se dilatara, de ahí su nombre). El efecto producido por la otra función es justo el contrario. Por lo que finalmente se obtiene la misma imagen que la original, salvo que haya agujeros o zonas no conexas, los cuales se habrán reducido o incluso habrán desaparecido. La utilización de estas funciones permite que la detección de los contornos realizada posteriormente sea mucho más robusta y precisa.

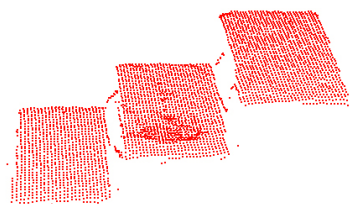


Figura 9: Nube de puntos tomada desde la parte superior de la carga. Representa una vista cenital.

Una vez que se tiene una imagen conexas, se utiliza un filtro de Canny [5] para obtener los bordes de la imagen y posteriormente ordenarlos. Finalmente, se contabilizan los píxeles existentes dentro de cada uno de los contornos mediante una función basada en el *teorema de Green* [8] para calcular el área. Dicha área se encuentra medida en píxeles y ya sólo queda transformarla a m^2 utilizando el escalado entre la imagen 2D y la nube de puntos 3D.

6 RESULTADOS EXPERIMENTALES

Una vez explicadas cada una de las etapas de las que consta el método desarrollado, falta evaluar la exactitud del volumen calculado. Para ello se han establecido quince situaciones diferentes a estudiar, en las que se ha variado la configuración de la carga desde situaciones muy sencillas, en las que hay una única caja, hasta otras en las que hay hasta once cajas y dispuestas en posiciones muy diversas. En cada una de estas situaciones se ha medido la exactitud del volumen estimado y el tiempo de cómputo necesario utilizando los métodos desarrollados. Los experimentos se han realizado en un ordenador de sobremesa de 64 bits con seis procesadores *Intel(R) Xeon(R) CPU X5650* (2,67 GHz) y con 24 GB de memoria RAM. Cabe decir que el código desarrollado no ha sido

optimizado para hacer uso de toda la capacidad computacional disponible en este ordenador.

Tabla 1: Resumen de los resultados obtenidos al aplicar cada método a los quince casos. El término error se refiere al error medio.

Método	Error	Desv. típica
“Convex hull” [3]	50,50%	35,25%
“Voxel grid”	7,40%	3,26%
“Proyección planos”	4,11%	3,48%

En la tabla 1 se muestra un resumen de los resultados obtenidos para los dos métodos explicados. En ella se puede observar como el error cometido al calcular el volumen mediante el método “voxel” es bastante aceptable. En concreto, el error medio es del 7,4 %, siendo la desviación del 3.26%, lo cual entra dentro de los límites admisibles para el sistema. El tiempo de cómputo empleado para calcular el volumen varía entre 1 y 5 segundos, en función de la cantidad de carga transportada. En caso de que se restrinja la mercancía a transportar cajas exclusivamente, el método de “proyección de planos” resulta más eficiente y preciso. Los tiempos de cómputo obtenidos mediante este método son los menores de todos (entre 0,4 y 2 segundos). Además, el error medio cometido es sólo del 4,1%.

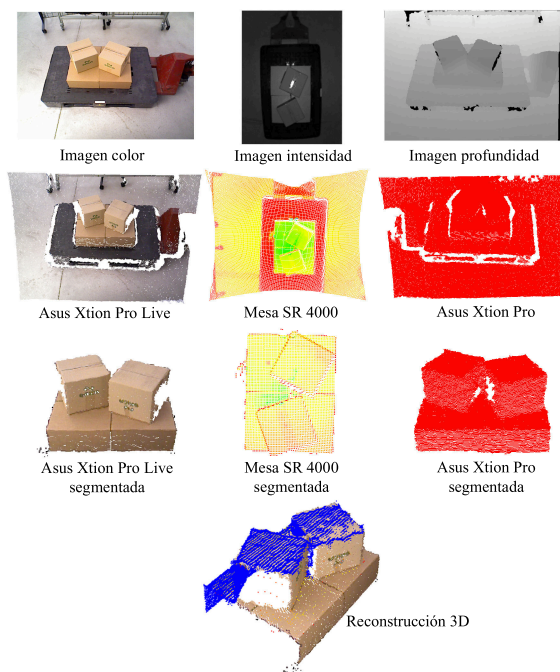


Figura 10: Resumen de los pasos realizados para calcular el volumen para una de las situaciones utilizadas para comprobar la validez del método.

A modo de ejemplo, en la figura 10 se muestra un resumen de los pasos intermedios realizados para

calcular el volumen en uno de los casos utilizados para evaluar el sistema. El volumen medido en esta configuración es de $67,5 \text{ dm}^3$, siendo la estimación mediante el método “voxel” de $72,5 \text{ dm}^3$ y de $69,3 \text{ dm}^3$ mediante el método de proyección de planos, lo que se corresponde con un error del 7,4% y del 2,7% respectivamente. En [1] se describe una relación más extensa de los experimentos realizados y de los resultados obtenidos.

7 CONCLUSIONES

En vista de los resultados, se puede afirmar que se ha logrado diseñar un sistema robusto y eficiente capaz de calcular el volumen de la mercancía transportada por una transpaleta encima de un palé. Además, se ha superado el reto que supone el trabajar con sensores basados en tecnologías diferentes como son la luz estructurada y la tecnología de tiempo de vuelo. El sistema desarrollado ha sido evaluado experimentalmente estimando el volumen en quince situaciones distintas que abarcan una amplia variedad de posibles configuraciones de mercancías. Los resultados muestran que el sistema diseñado es preciso, con un margen de error menor del 10% de media y con un tiempo de cómputo de unos segundos. Aunque este error puede ser reducido mejorando algunos aspectos concretos del sistema implementado, el margen de mejora es bastante reducido ya que este valor es muy próximo a los errores de medición inherentes a las condiciones de medida y a los propios dispositivos observados en el estudio de exactitud y precisión realizado.

Dados los buenos resultados obtenidos en las pruebas realizadas con el sistema diseñado, éste ha sido instalado en el Centro Demostrador TIC de Logística del Instituto Tecnológico de Aragón. La aplicación va a ser mostrada a las empresas que visiten estas instalaciones para que comprueben las posibilidades que ofrecen los sensores de rango utilizados para la realización de este trabajo.

Agradecimientos

Los autores agradecen a Guillermo López (ITA INNOVA) su ayuda en la realización de este trabajo. El trabajo ha sido parcialmente financiado por el Ministerio de Economía y Competitividad y la Unión Europea (DPI2012-31781), por la DGA-FSE (grupoT04) y por el Instituto Tecnológico de Aragón (beca ref. PFCLOGISTICA1-13-458).

Referencias

- [1] Artaso, P., (2014) “Tecnologías en cámaras 3D y procesamiento de nubes de puntos para su aplicación en logística”, Proyecto Fin de Carrera, Universidad de Zaragoza.
- [2] Arun, K. S., Huang, T. S., y Blostein, S. D., (1987) “Least-Squares Fitting of Two 3-D Point Sets”, *IEEE Trans. Pattern Anal. Mach. Intell.*, pp. 698–700.
- [3] Barber, C. B., Dobkin, D. P. y Huhdanpaa, H., (1996) “The Quickhull Algorithm for Convex Hulls”, *ACM Trans. Math. Softw.*, pp. 469–483.
- [4] Besl, P. J., y McKay, N. D., (1992) “A Method for Registration of 3-D Shapes”, *IEEE Trans. Pattern Anal. Mach. Intell.*, pp. 239–256.
- [5] Canny, J., (1986) “A Computational Approach to Edge Detection”, *IEEE Trans. Pattern Anal. Mach. Intell.*, pp. 679–698.
- [6] Dellen, B. y Rojas, I. A., (2013) “Volume measurement with a consumer depth camera based on structured infrared light” *16th Catalan Conference on Artificial Intelligence*.
- [7] Fischler, M. A., y Bolles, R. C., (1981) “Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography”, *Commun. ACM*, pp. 381–395.
- [8] Kaplan, W., (1991) “Green’s Theorem”, §5.5 in *Advanced Calculus, 4th ed. Reading, MA: Addison-Wesley*, pp. 286–291.
- [9] Kaufman, L., y Rousseeuw, P. J., (1990) “Finding Groups in Data: An Introduction to Cluster Analysis”, John Wiley and Sons.
- [10] Kolb, A., Barth, E., Koch, R., y Larsen, R., (2009) “Time-of-Flight Sensors in Computer Graphics”, *Proceedings of Eurographics*, pp. 119–134.
- [11] Lee, D. J., Eifert, J., Zhan, P., y Westhover, B., (2003) “Fast surface approximation for volume and surface area measurements using distance transform”, *Opt. Eng.*, pp. 2947–2955.
- [12] Rusu, R. B., y Cousins, S., (2011) “3D is here: Point Cloud Library (PCL)”, *IEEE International Conference on Robotics and Automation (ICRA)*, China.
- [13] Tam, G. K. L., Cheng, Z.-Q., Lai, Y.-K., Langbein, F. C., Liu, Y., Marshall, D., Martin, R. R., Sun, X.-F., y Rosin, P. L., (2013) “Registration of 3D Point Clouds and Meshes: A Survey from Rigid to Non-rigid”, *IEEE Transactions on Visualization and Computer Graphics*, pp. 1199-1217.