

# Reconstrucción 3D basada en estimación de desplazamiento

Federico Cristina, Sebastián H. Dapoto, Verónica Artola, Claudia Russo  
Instituto de Investigaciones en Informática LIDI. Facultad de Informática. UNLP  
{fcristina, sdapoto, vartola, crusso} @lidi.info.unlp.edu.ar

**Abstract.** Los dispositivos de escaneo tridimensional permiten obtener modelos 3D utilizando distintas técnicas de captura. Una técnica muy utilizada consiste en realizar la reconstrucción mediante el uso de estereovisión y luz estructurada. De esta forma, proyectando un patrón de luz conocido sobre una escena es posible determinar la correspondencia de puntos entre las imágenes. Por otro lado, las técnicas de estimación de movimiento pueden ser incorporadas a fin de encontrar dichas correspondencias sin la necesidad de utilizar una fuente de luz. El presente trabajo plantea una solución al problema de reconstrucción de modelos tridimensionales utilizando este tipo de técnicas, logrando así un prototipo simple, económico y de fácil utilización.

**Keywords:** estereovisión, escáner tridimensional, reconstrucción 3D, estimación de movimiento, block matching

## 1 Introducción

Los sistemas de digitalización 3D de visión activa sin contacto son los que hacen intervenir una fuente de luz específica para determinar las coordenadas tridimensionales de los puntos a medir [1]. Constan como mínimo de un emisor de luz y uno o más receptores. Conociendo el ángulo entre el emisor de luz y su receptor, se puede calcular trigonómicamente la profundidad del punto inspeccionado. Ejemplos de esta técnica son los sistemas que usan algún tipo de proyección de luz estructurada [3] o iluminación láser [4] sobre la superficie que se intenta reconstruir.

Los sistemas de digitalización 3D de visión pasiva se basan en utilizar dos o más puntos de vista de un mismo objeto para encontrar las coordenadas tridimensionales [2]. La principal complejidad de este método es la correspondencia de los puntos entre cada una de las imágenes monoculares. Las técnicas encargadas de resolver esta tarea presentan un elevado costo computacional, debido principalmente a la necesidad de realizar una búsqueda exhaustiva a fin de determinar la correspondencia de puntos.

El presente trabajo describe el desarrollo de un dispositivo de captura y reconstrucción de modelos 3D, el cual se encuentra dentro del grupo de visión pasiva. Se detallan los aspectos relevantes del software, en el cual se buscó minimizar los tiempos de procesamiento. Se presentan además los resultados obtenidos hasta el momento. Por último se presentan las conclusiones y el trabajo a futuro.

## 2 Estado del arte

Todos los sistemas de reconstrucción cuentan con un método de captura de la escena y de cálculo de correspondencia con el fin de adquirir las coordenadas 3D del objeto que se intenta reconstruir. Gran parte de los mismos pertenecen al grupo de visión activa, es decir, utilizan una fuente lumínica para determinar la correspondencia de puntos.

En la tabla 1 se presenta un resumen de trabajos categorizados según el método de reconstrucción y los dispositivos de hardware utilizados.

**Tabla 1.** Resumen de trabajos relacionados.

Hardware necesario	Técnica utilizada
Una cámara y un láser [4, 5, 6]	Realiza un barrido láser sobre el objeto a adquirir.
Una o más cámaras y un proyector de luz estructurada [3, 9]	Realiza una proyección de luz sobre el objeto a adquirir.
Una cámara, un láser y una estructura de referencia (doble cuadro) [8]	Realiza un barrido láser sobre el objeto a adquirir.
Dos cámaras, 1 láser [7, 12]	Realiza un barrido láser sobre el objeto a adquirir.
Una cámara. [10, 11, 15]	Utiliza un conjunto de imágenes del objeto a adquirir.
Dos cámaras [13]	Utiliza mapas de disparidad.
Una cámara y un patrón de puntos [14]	Utiliza un conjunto de imágenes del objeto junto al patrón. Reconoce el patrón de puntos situado como base del objeto a reconstruir y lo usa para generar el modelo.

El sistema desarrollado en el presente trabajo intenta minimizar el costo del prototipo de captura prescindiendo de un proyector de luz estructurada o láser, y haciendo uso de técnicas de estimación de desplazamiento para calcular las correspondencias entre los puntos obtenidos con el sistema de cámaras estéreo.

## 3 Fundamentos

Como se explicó previamente, una alternativa a la utilización de luz estructurada es el uso de técnicas de estimación de desplazamiento a fin de lograr la correspondencia de puntos.

Estas técnicas permiten determinar, por ejemplo, la distancia recorrida de un punto entre dos cuadros consecutivos. En la figura 1 se pueden observar dos imágenes pertenecientes a una secuencia de video, y una tercera en la que se muestran los vectores de desplazamiento de los bloques correspondientes entre ambas imágenes.



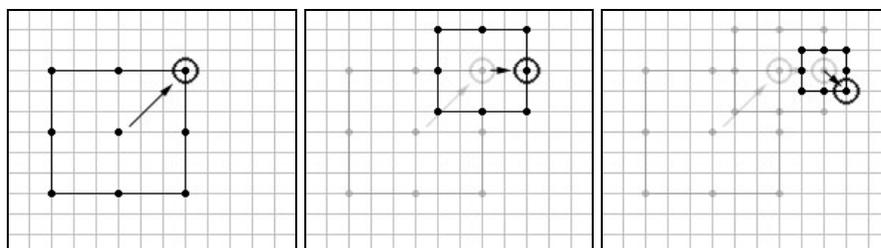
**Fig 1. a) y b)** Dos cuadros consecutivos pertenecientes a un video. **c)** Vectores de desplazamiento entre los respectivos cuadros.

Una de las técnicas más utilizadas para detección de movimiento es la denominada Block Matching [18]. Básicamente, una imagen  $Im1$  (por ejemplo imagen izquierda) se divide en un conjunto de bloques, y se determina la nueva ubicación de este bloque en la imagen  $Im2$  (imagen derecha).

Uno de los algoritmos más empleados a fin de determinar el lugar donde el bloque se ha movido en la imagen  $Im2$ , es la denominada Three Step Search. La misma consta de los siguientes pasos, los cuales son realizados iterativamente a lo largo de toda la imagen:

1. Se determina el tamaño  $T$  del bloque central  $C$  de la imagen  $Im1$ . Se toman 8 bloques de la imagen  $Im2$  para realizar la comparación. Éstos se encuentran a una distancia  $T$  con respecto al centro del bloque central  $C$ .
2. El bloque central  $C$  pasa a ser el que resulte de mayor similitud. El tamaño  $T$  es dividido a la mitad.
3. Los pasos 1 y 2 son repetidos hasta que  $T$  sea menor que 1.

A fin de determinar el bloque de mayor similitud, se realiza la comparación entre el central y cada uno de los candidatos mediante la función Mean Absolute Error. La comparación que devuelva el menor MAE indica que éste es el candidato de mayor similitud, con lo cual el vector deberá ser dirigido hacia dicha dirección. La figura 2 muestra un ejemplo de los pasos que comprende la técnica.



**Fig 2.** Pasos en la técnica Three Step Search para determinar la correspondencia de puntos

Al llevar esta técnica al plano de la reconstrucción de modelos tridimensionales, la determinación de los vectores de desplazamiento entre dos imágenes estereó de un objeto, hace posible la obtención de las correspondencias estereó de las mismas, aumentando en este caso la granularidad por bloque a granularidad por píxel.

Iterando este proceso en toda la imagen, se determinan las correspondencias de cada uno de los puntos que pertenecen al objeto a reconstruir, para luego obtener la nube de puntos 3D mediante el uso de estereovisión [16].

A fin de adaptar esta técnica en la solución implementada, fue necesario modificar la misma con el objetivo de maximizar la calidad de los resultados. Es por esto que en lugar de utilizar 8 puntos de comparación, y a fin de conocer la ubicación exacta del bloque en la imagen Im2, la implementación desarrollada realiza una búsqueda completa alrededor del bloque en análisis. Esta es una búsqueda exhaustiva – denomina Full Search – y aunque requiere un alto costo computacional, es la que brinda los mejores resultados en cuanto a la calidad lograda.

## 4. Detalle del procedimiento

La solución desarrollada consta de tres etapas principales: calibración y adquisición de las imágenes, estimación de las correspondencias y generación de la nube de puntos. Cada una de éstas se detalla a continuación.

### 4.1 Calibración y adquisición de las imágenes

Con el fin de lograr una correcta triangulación de los puntos en el espacio, es necesario primeramente conocer las características del sistema de cámaras a utilizar [17]. Utilizando una grilla de calibración es posible obtener los parámetros intrínsecos y extrínsecos del sistema de adquisición.

Si la disposición de las cámaras no es modificada, el proceso de calibración no deberá ser repetido nuevamente en futuras adquisiciones. En el caso que la posición de las cámaras fuera alterada (por ejemplo para adquirir objetos de diferentes tamaños), será necesario recalibrar extrínsecamente el sistema.

Una vez que las cámaras se encuentran correctamente calibradas, se procede a realizar la captura de una imagen del objeto a reconstruir con cada una de éstas.

## 4.2 Estimación de las correspondencias

Dado que el proceso de estimación de correspondencias es el más costoso computacionalmente se desarrolló una solución ad hoc basada en la técnica de Block Matching, explicada previamente.

A fin de minimizar los tiempos de respuesta, el algoritmo implementado permite definir una serie de parámetros de entrada que brindan la versatilidad suficiente para adaptarse a las características de la escena y del objeto a capturar; reduciendo al mínimo el área de búsqueda.

- Macrobloque (S): permite especificar el tamaño inicial del bloque.
- Búsqueda horizontal (BH) y búsqueda vertical (BV): sea  $(x,y)$  el punto de origen en la imagen izquierda. La búsqueda de su correspondencia en la imagen derecha se realizará desde el punto  $(x - BH, y - BV)$  hasta el punto  $(x + BH, y + BV)$ .
- Desplazamiento horizontal (DH) y desplazamiento vertical (DV): dependiendo de las características ópticas de las cámaras y del tamaño y ubicación del objeto a reconstruir, un píxel puede estar menor o mayormente desplazado con respecto a su correspondiente en la imagen derecha. Los valores DH y DV se utilizan para mejorar la efectividad del algoritmo realizando la búsqueda en la zona más acorde a las características de la escena, descartando escenarios inexistentes.
- Detalle malla (DM): mediante este parámetro es posible variar la densidad de la nube de puntos resultante. El valor DM indica el salto en píxeles entre dos puntos procesados. De esta forma, para  $DM=1$  se procesan las imágenes en su totalidad, mientras que para  $DM=N$  ( $N \neq 1$ ), se procesa 1 de cada N píxeles.

## 4.3 Generación de la nube de puntos

Una vez que se determinaron las correspondencias de los puntos del objeto entre la imagen izquierda y derecha, es posible calcular las coordenadas tridimensionales. Para esto es necesario utilizar la información extrínseca obtenida en la etapa de calibración y realizar las triangulaciones correspondientes [17].

A fin de mejorar los resultados finales, el algoritmo cuenta con dos procesos de mejoramiento de la nube de puntos resultante, las cuáles pueden ser habilitadas o no según sea necesario. Estos procesos son:

- Corrección de cálculos erróneos: si se habilita esta función, los puntos que se encuentren muy distantes a su entorno son clasificados como erróneos y por lo tanto eliminados de la nube de puntos.

- Erosión: esta función verifica que cada punto contenga una determinada cantidad de vecinos. Si la cantidad de vecinos de un punto no supera el umbral deseado, el mismo se elimina. Esto produce nubes de puntos de menor densidad.

Las nubes de puntos son almacenadas en formato WRL, con el fin de hacer posible su visualización en cualquier entorno VRML (Virtual Reality Modeling Language), por ejemplo, un navegador web que posea un plug-in para este formato. El lenguaje VRML permite representar gráficos tridimensionales de una forma simple a través de un archivo de texto. Para cada punto se almacena su coordenada 3-D y su textura relevada a partir de las imágenes de entrada.

## 5. Resultados

Las pruebas presentadas en este trabajo fueron realizadas en un ambiente controlado, a fin de analizar la posibilidad de su implementación en un escenario real.

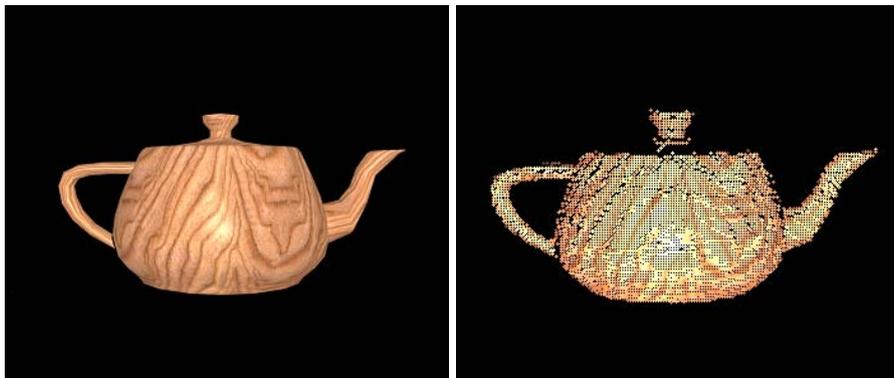
Para esto, se realizó un modelo virtual mediante un programa de creación de gráficos y animación 3D, en el cual se hicieron pruebas con diferentes posicionamiento y orientación de las cámaras. Por otro lado, se utilizaron objetos de diversos tamaños, colores y texturas.

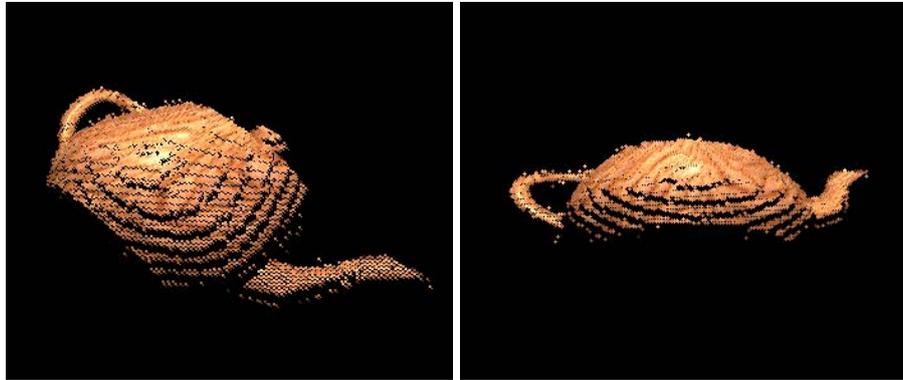
Cabe destacar que en las pruebas efectuadas, la variación en la disposición y orientación de las cámaras incidieron considerablemente en los resultados obtenidos.

Bajo una configuración de cámaras en las que éstas se encuentran demasiado separadas (dependiendo ésto del objeto a capturar), existirán problemas de oclusión. Esto es: sectores del objeto que son alcanzados por la cámara derecha pero no por la cámara izquierda y viceversa.

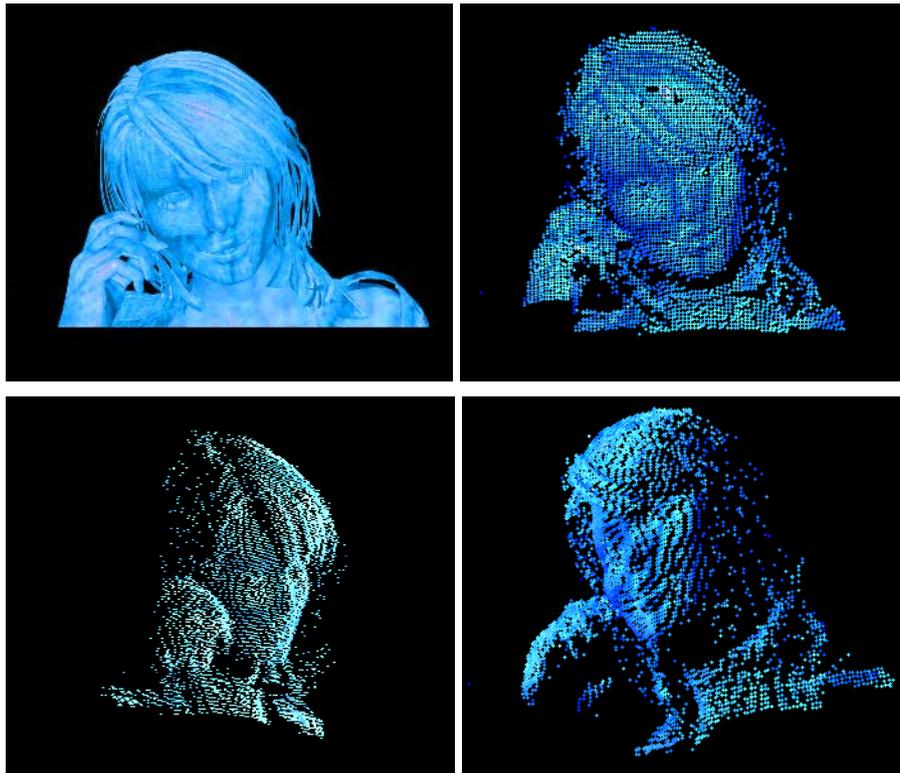
De la misma forma, si las cámaras se ubican de manera tal que los ejes de captura de cada cámara se intersecten, las vistas obtenidas por las mismas son suficientemente diferentes como para afectar considerablemente la búsqueda de correspondencias, degradando así los resultados obtenidos.

En las figuras 3 y 4 se pueden observar dos ejemplos de los resultados obtenidos mediante el software desarrollado.





**Fig 3.** Modelo 3D de una tetera y la nube de puntos obtenida desde distintas perspectivas.



**Fig 4.** Modelo 3D de un rostro femenino y la nube de puntos obtenida desde distintas perspectivas.

## 6. Prototipo

El prototipo de hardware para el sistema de captura se encuentra en etapa de desarrollo y contará con dos cámaras soportadas mediante una estructura diseñada a medida (figura 5).

Para lograr una calidad aceptable en los resultados, las cámaras deberán adquirir imágenes con una resolución mínima de 640x480 píxeles.

La estructura deberá ser simple y rígida, a fin de mantener las cámaras en la posición inicial dispuesta. La misma permitirá una serie de posicionamientos diferentes, de forma tal que las cámaras puedan ser ubicadas a menor o mayor distancia entre sí, dependiendo del tamaño del objeto a capturar.



Fig 5. Prototipo de captura

## 7. Conclusiones

Aunque el proyecto se encuentra en etapa de desarrollo, el estado actual del mismo permite prever resultados más que aceptables con la tecnología de hardware que se va a utilizar en conjunto con el software desarrollado.

Las pruebas realizadas en ambiente controlado aportan un balance positivo en cuanto a la relación costo/beneficio, dado que el software sólo utiliza un par de imágenes estereó como entrada y permite trabajar con objetos de diversos tamaños.

El escenario a futuro prevé la finalización del prototipo de hardware, a fin de realizar pruebas en un ambiente real. Por otro lado, se estudiará la alternativa del desarrollo de un prototipo de hardware con una única cámara móvil.

## 8. Referencias

- [1] Digitalización.  
<http://plutarco.disca.upv.es/~jcperez/doctorado/SV2D3DPI/trabajos/Calzado/digitalizacion.html>
- [2] Surface Reconstruction by Propagating 3D Stereo Data in Multiple 2D Images  
Gang Zeng, Sylvain Paris, Long Quan, and Maxime Lhuillier
- [3] Interactive 3-D Media with Structured Light Scanning  
Nelson L. Chang - Imaging Systems Laboratory  
HP Laboratories Palo Alto - HPL-2003-112 - May 29th , 2003
- [4] A Simple 3D Scanner. Steve Baker.  
[http://www.sjbaker.org/wiki/index.php?title=A\\_Simple\\_3D\\_Scanner](http://www.sjbaker.org/wiki/index.php?title=A_Simple_3D_Scanner)
- [5] Low-Cost Laser Range Scanner and Fast Surface Registration Approach  
Simon Winkelbach, Sven Molkenstruck, Friedrich M. Wahl
- [6] High Speed 3D Scanner with Real-Time 3D Processing  
Joseph P. Lavelle, Stefan R. Schuet, Daniel J. Schuet
- [7] Interactive 3D Scanning Without Tracking  
Matthew J. Leotta, Austin Vandergon, Gabriel Taubin
- [8] A paint-brush laser range scanner  
Lyubomir Zagorchev, A. Ardeshir Goshtasby
- [9] A low cost 3D scanner based on structured light  
C. Rocchini, P. Cignoni, C. Montani, P. Pingi, R. Scopigno
- [10] An automatic method for acquiring 3d models from photographs: applications to an archaeological site  
M. Pollefeys, R. Koch, M. Vergauwen, L. Van Gool
- [11] From Images to 3D Models: How computers can automatically build realistic 3D models from images acquired with a hand-held camera  
M. Pollefeys, L. Van Gool
- [12] Desarrollo de un escáner 3D mediante cámaras estereoscópicas e iluminación láser  
Federico Cristina, Sebastián H. Dapoto, Javier Vegas, Verónica Artola, Claudia Russo, Abásolo María José
- [13] 3D reconstruction from stereo images for interactions between real and virtual objects  
Hansung Kim, Kwanghoon Sohn
- [14] 3D Software Object Modeller  
<http://www.3dsom.com/>
- [15] Reality : an Interactive Reconstruction Tool of 3D Objects from Photographs  
S. Dedieu, P. Guitton, C. Schlick, P. Reuter
- [16] Introductory Techniques for 3D Computer Vision  
Emanuele Trucco, Alessandro Verri
- [17] An Invitation to 3-D Vision: From Images to Geometric Models  
Yi Ma, Stefano Soatto, Jana Kosecka, S. Shankar Sastry
- [18] Search Algorithms for Block-Matching in Motion Estimation.  
Deepak Turaga, Mohamed Alkanhal.  
[http://www.ece.cmu.edu/~ee899/project/deepak\\_mid.htm](http://www.ece.cmu.edu/~ee899/project/deepak_mid.htm)