



Soda Stéreo

Introducción

En este trabajo práctico estamos interesados en técnicas que nos permitan realizar digitalizaciones 3D invirtiendo poco dinero, es decir, utilizando solamente cámaras tradicionales. En particular, vamos a desarrollar una aplicación simplificada para la digitalización de objetos 3D basándonos en trabajos anteriores de la temática [2]. Utilizaremos la técnica de *fotometría estéreo* que nos permite *medir* la posición y distancia de los objetos con respecto al sensor (cámara), es decir, obtener las coordenadas x, y, z basándonos únicamente en las propiedades reflectivas de la luz, para objetos no especulares y en un ambiente con iluminación controlada.

El foco de este trabajo práctico será el análisis de los métodos numéricos para la resolución de los sistemas lineales obtenidos planteados.

Fotometría estéreo

Cuando miramos una fotografía o vemos la televisión, no vemos una imagen en 2D, sino que entendemos lo que vemos como un objeto en 3 dimensiones. Es decir, podemos inferir la forma 3D de los objetos a partir de imágenes 2D. La intensidad registrada en un determinado píxel depende de un número de factores: las condiciones de iluminación de la escena, la orientación del objeto, y las propiedades del material que componen su superficie. El brillo de un objeto de superficie Lambertiana [1] depende de la orientación de la superficie con respecto a la fuente de luz.

La intensidad de luz incidente se denota I_0 y el ángulo entre la luz incidente y el vector normal a la superficie se denomina ϕ . En la Figura 1 podemos ver un ejemplo del modelo de reconstrucción planteado.

Luego, la intensidad registrada en cada punto (x, y) de la imagen está dada por:

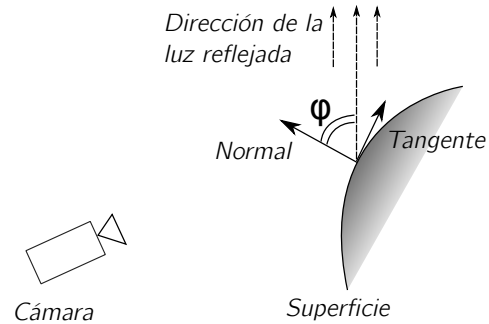


Figura 1: Relación entre la superficie del objeto y la fuente de luz. La cantidad de luz reflejada en una localidad de la superficie es proporcional al coseno del ángulo entre la fuente de luz y la normal a la superficie.

$$I(u, v) = I_0 \rho \cos(\phi) = I_0 \rho (n_x s_x + n_y s_y + n_z s_z) \quad (1)$$

donde la normal \mathbf{n} , y la dirección de la fuente de luz \mathbf{s} , están dadas por:

$$\mathbf{n} = \begin{pmatrix} n_x \\ n_y \\ n_z \end{pmatrix}, \quad \mathbf{s} = \begin{pmatrix} s_x \\ s_y \\ s_z \end{pmatrix} \quad (2)$$

Para inferir la orientación de la superficie necesitamos al menos tres imágenes cuya fuente de iluminación difiera entre sí. Éstas pueden ser tres posiciones espaciales distintas de la fuente lumínica. Supongamos \mathbf{s}^i tal que:

$$\mathbf{s}^1 = \begin{pmatrix} s_x^1 \\ s_y^1 \\ s_z^1 \end{pmatrix}, \mathbf{s}^2 = \begin{pmatrix} s_x^2 \\ s_y^2 \\ s_z^2 \end{pmatrix}, \mathbf{s}^3 = \begin{pmatrix} s_x^3 \\ s_y^3 \\ s_z^3 \end{pmatrix} \quad (3)$$

Cada una de las imágenes tendrá entonces distintas intensidades I_1 , I_2 e I_3 , en cada uno de los píxeles correspondientes. Como supusimos una superficie de tipo Lambertiana, las tres intensidades pueden relacionarse con la normal a la superficie del objeto en cada punto y la dirección de la luz incidente correspondiente:

$$I_i = I_0 \rho (n_x s_x^i + n_y s_y^i + n_z s_z^i), \quad \forall i = 1, 2, 3. \quad (4)$$

Como datos conocidos tenemos I_1 , I_2 e I_3 , y las direcciones de la fuente de iluminación \mathbf{s}^1 , \mathbf{s}^2 y \mathbf{s}^3 de cada imagen. No conocemos la intensidad de la fuente I_0 , el parámetro ρ y la normal \mathbf{n} . Podemos sin embargo juntar las incógnitas por cada coordenada en una nueva variable \mathbf{m} tal que $m_x = I_0 \rho n_x$, $m_y = I_0 \rho n_y$ y $m_z = I_0 \rho n_z$. De esta forma tenemos que resolver el siguiente sistema de ecuaciones para cada uno de los píxeles de la imagen y de aquí obtener \mathbf{n} :

$$\begin{pmatrix} s_x^1 & s_y^1 & s_z^1 \\ s_x^2 & s_y^2 & s_z^2 \\ s_x^3 & s_y^3 & s_z^3 \end{pmatrix} \begin{pmatrix} m_x \\ m_y \\ m_z \end{pmatrix} = \begin{pmatrix} I_1 \\ I_2 \\ I_3 \end{pmatrix} \quad (5)$$

Luego, podemos obtener \mathbf{n} de la siguiente igualdad:

$$\|\mathbf{m}\| = |I_0 \rho| \|\mathbf{n}\|, \quad \|\mathbf{n}\| = 1 \quad (6)$$

Finalmente tenemos para cada píxel de la imagen (x, y, \mathbf{n}) . Ahora, necesitamos encontrar la coordenada z (profundidad) a partir de \mathbf{n} . Para ello sabemos que \mathbf{n} es la normal a la superficie en el punto (x, y) y si \mathbf{v}_1 y \mathbf{v}_2 son aproximaciones a los vectores tangente en las direcciones X y Y respectivamente (ver Figura 2), se puede inferir píxel a píxel que

$$\mathbf{v}_1 = (x+1, y, z_{x+1,y}) - (x, y, z_{x,y}) = (1, 0, z_{x+1,y} - z_{x,y}) \quad (7)$$

$$\mathbf{v}_2 = (x, y+1, z_{x,y+1}) - (x, y, z_{x,y}) = (0, 1, z_{x,y+1} - z_{x,y}) \quad (8)$$

Como \mathbf{n} debe ser normal a \mathbf{v}_1 y \mathbf{v}_2 , se debe cumplir lo siguiente:

$$\mathbf{n} \cdot \mathbf{v}_1 = 0 \quad (9)$$

$$\mathbf{n} \cdot \mathbf{v}_2 = 0 \quad (10)$$

Lo cual se puede reescribir usando las ecuaciones 7 y 8:

$$n_x + n_z(z_{x+1,y} - z_{x,y}) = 0 \quad (11)$$

$$n_y + n_z(z_{x,y+1} - z_{x,y}) = 0 \quad (12)$$

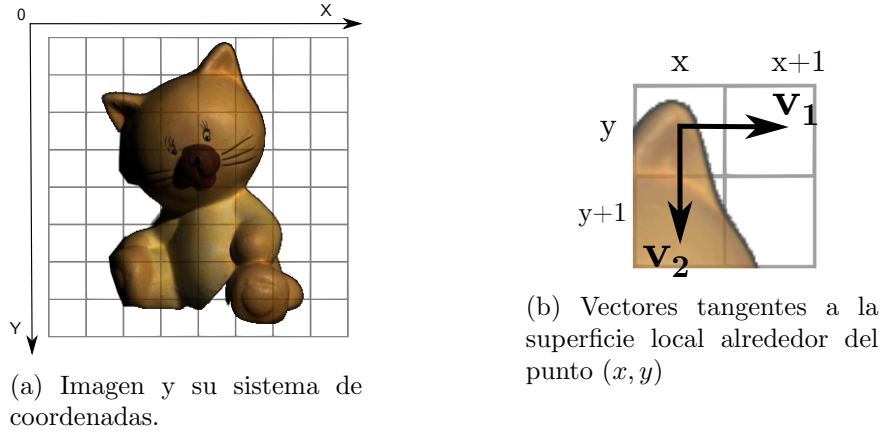


Figura 2: Diagrama de los sistemas de coordenadas usados.

Las ecuaciones 11 y 12, nos permiten construir un sistema de ecuaciones sobredeterminado para todo el espacio de píxeles (teniendo en cuenta los bordes donde no hay vecinos):

$$Mz = v \quad (13)$$

Que resolveremos de la siguiente manera¹:

$$M^t Mz = M^t v \quad (14)$$

De esta forma, podemos definir $A = M^t M$ y $b = M^t v$ y así resolver el siguiente sistema lineal mediante los métodos vistos en clase para obtener la profundidad z en cada píxel (x, y) .

$$Az = b \quad (15)$$

De esta forma, obtenemos la representación 3D del objeto, uniendo las coordenadas de los píxeles, la profundidad y su normal en cada una de ellas

Procedimiento

A continuación, describimos paso a paso el procedimiento de la aplicación de digitalización de objetos 3D mediante fotometría estéreo.

- Calibración del sistema. Cálculo de las direcciones de iluminación \mathbf{s} .

La secuencia de imágenes de la esfera (mate) nos servirá para calibrar el sistema por **única vez** y así obtener las direcciones de iluminación en cada caso $\mathbf{s}^1, \mathbf{s}^2, \dots, \mathbf{s}^k$, con $k = 12$ para nuestro conjunto imágenes. Dado que se conoce la geometría de la superficie de la esfera, se conoce el valor de la normal en cualquier punto de la superficie. Por lo tanto, el punto más brillante sobre la imagen de la esfera será el que corresponderá a la dirección de iluminación buscada y en dirección opuesta al normal ese punto ².

¹Notar que estas son las ecuaciones normales que resuelven el problema por cuadrados mínimos. Este tema no es el foco del presente trabajo y será tratado más adelante en la materia.

²**Sugerencia:** este punto puede encontrarse buscando en la imagen el píxel de máximo valor en un entorno.

- Reconstrucción del modelo 3D de los objetos digitalizados.

Dadas las secuencias de imágenes de un objeto y las direcciones de iluminación en cada caso obtenidas en la calibración, queremos encontrar los puntos (x, y, z) y las normales en cada punto que representará al mismo:

1. Construcción del campo normal.

Cada coordenada (x, y) correspondiente a los píxeles de la imagen servirá para calcular la normal en ese punto a partir de la ecuación 5 y así obtener \mathbf{m} . Luego, con la ecuación 6 se despeja \mathbf{n} para obtener todo el campo normal. Aquí se deben elegir 3 direcciones de iluminación de todas las obtenidas en la calibración.

2. Estimación de la profundidad.

Por último debe plantearse y resolverse el sistema de ecuaciones lineales expresado en la ecuación 15 para encontrar cada z . El sistema a resolver cuya matriz es $A = M^t M$, tiene dimensión $2N \times 2N$ donde N de la cantidad de píxeles de la imagen.

En la Figura 3 se puede ver un ejemplo de estimación de normales y de la profundidad para un objeto dado.

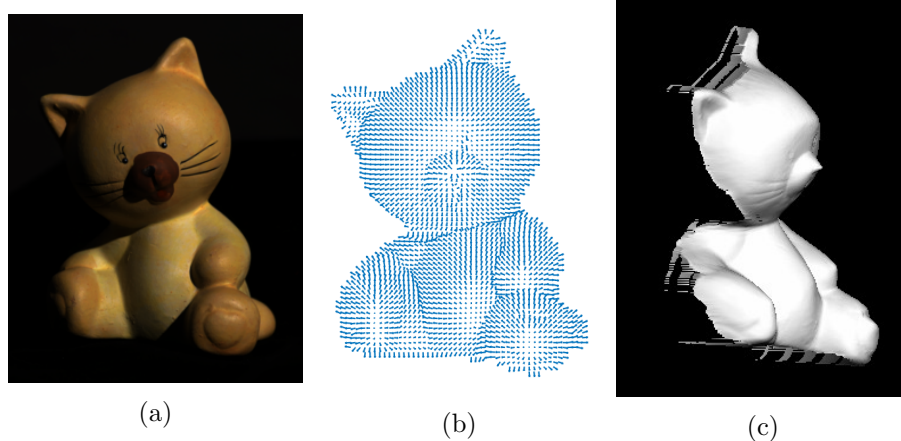


Figura 3: (a) Imagen de un objeto con iluminación direccional. (b) Normales obtenidas con el método. (c) Vista generada del modelo 3D estimado.

Enunciado

Se debe implementar un programa en C o C++ que realice el procedimiento de digitalización 3D mediante fotometría estéreo descrito en la Sección Procedimiento. Como parte obligatoria en relación a los métodos de resolución de sistemas de ecuaciones lineales se pide implementar lo siguiente:

1. El método de Eliminación Gaussiana clásico (EG)
2. El método de Factorización de Cholesky (CL). Justificar por qué a la matriz A de la ecuación 15 se le puede aplicar CL.

Además, se pide realizar un informe utilizando como guía las pautas de laboratorio de la materia³. Es importante incluir en el informe del trabajo práctico, en la sección desarrollo, aquellas decisiones tomadas en función de las estructuras de datos utilizadas y las alternativas consideradas y descartadas para los métodos utilizados.

Experimentación

La experimentación será dividida en dos partes, cada una con sus respectivos ejes. En primer lugar, buscamos hacer una evaluación *cuantitativa* de los métodos de resolución de sistemas lineales considerados, i.e., EG y CL, para la resolución del sistema de la ecuación 15, en términos del tiempo de cómputo y el tamaño de los sistemas a resolver. Para conseguir sistemas más chicos, se pueden reducir las imágenes provistas.

Además, sabemos que existen casos donde el algoritmo EG no puede encontrar una solución. **Se pide** incluir en el desarrollo una justificación sobre por qué el algoritmo funciona correctamente en el caso del método.

Tener en cuenta que el sistema planteado en la ecuación 15 es un sistema muy grande. Para lidiar con posibles problemas de implementación, **se pide** realizar alguna de las siguientes opciones:

- Considerar versiones reducidas de las imágenes de entrada.
- Notar que la matriz A tiene gran parte de sus elementos igual a cero, es decir, es una matriz esparsa. Considerar este hecho para optimizar los cálculos.
- Utilizar la máscara suministrada para cada objeto para reducir el cómputo, con el fin de realizar los cálculos únicamente donde la máscara tiene el color blanco.

Por otro lado, el sistema de 3×3 planteado en la ecuación 5 debe ser resuelto tantas veces como píxeles tengamos con el fin de obtener todas las normales. Puede verse que los valores s_x^i, s_y^i, s_z^i no cambian píxel a píxel sino que solamente se modifica el valor de la intensidad en (x, y) para la imagen correspondiente I_i , y por lo tanto poseen la misma matriz pero con distinto término independiente. **Se pide**, utilizar estas propiedades para reducir los cálculos. Justificar.

La segunda parte de la experimentación se centra en un análisis *cualitativo* de los resultados obtenidos de la aplicación de reconstrucción 3D. Para lo cual, será responsabilidad de los integrantes del grupo plantearse preguntas en forma de hipótesis o conjeturas, que deberán intentar responder adecuadamente. Si bien muchos de estos interrogantes no podrán ser respondidos con argumentos teóricos, es necesario que el análisis de los mismos se realice mostrando resultados convincentes (gráficos, diagramas, tablas, etc), que refuercen y/o validen las respuestas propuestas por ustedes. **Se pide**, como mínimo analizar y responder las primeras tres preguntas siguiente:

1. Comparar las direcciones de iluminación obtenidas por el método de calibración con las provistas por la cátedra.

³<http://www-2.dc.uba.ar/materias/metnum/homepage.html>

2. ¿Cómo afecta la calibración del sistema en el resto de las etapas?
3. ¿Cómo impacta la elección de las 3 direcciones de iluminación para el cálculo de las normales?
4. Dado que cada imagen RGB tiene 3 canales (rojo, verde y azul), ¿Cómo pueden ser combinados para afectar lo menos posible los resultados?
5. ¿Cómo afectan la estimación de las profundidades el cálculo de las normales?
6. ¿Qué métodos de solución de los sistemas lineales arrojan mejores cálculos de profundidades? Justifique.

En todos los casos es **obligatorio** fundamentar los experimentos planteados, proveer los archivos e información necesarios para replicarlos, presentar los resultados de forma conveniente y clara, y analizar los mismos con el nivel de detalle apropiado. En caso de ser necesario, es posible también generar instancias artificiales con el fin de ejemplificar y mostrar un comportamiento determinado.

Fecha de entrega

- Formato Electrónico: Jueves 7 de Septiembre de 2017, hasta las 23:59 hs, enviando el trabajo (informe + código) a la dirección `metnum.lab@gmail.com`. El subject del email debe comenzar con el texto [TP1] seguido de la lista de apellidos de los integrantes del grupo.
- Formato físico: Viernes 8 de Septiembre de 2017, a las 18 hs. en la clase de laboratorio.
- Pautas de laboratorio: <http://www-2.dc.uba.ar/materias/metnum/homepage.html>

Importante: El horario es estricto. Los correos recibidos después de la hora indicada serán considerados re-entrega.

Referencias

- [1] Tinku Acharya and Ajoy K Ray. *Image processing: principles and applications*. John Wiley & Sons, 2005.
- [2] E Eric Kalu et al. *Numerical Methods with Applications: Abridged*. Lulu. com, 2008.