

Método óptico digital de extracción de información tridimensional de objetos. Aplicación a los modelos del Centro de Estudios de Puertos y Costas

IGNACIO IGLESIAS CASARRUBIOS (*)

RESUMEN. Se describe un sistema óptico digital de extracción de información tridimensional de objetos. Se usan las técnicas de codificación de grises y muestreo de fase para formar imágenes de alta resolución en las que se codifican las tres coordenadas del objeto. Se trata de un sistema robusto en cuanto que es capaz de tratar las zonas de ocultación y de proceso rápido. Se describe una posible aplicación en Ingeniería Civil en particular, a la determinación de superficies en los modelos de dinámica de puertos y costas.

OPTICAL DIGITAL METHOD FOR OBTAINING THREE-DIMENSIONAL INFORMATION FROM OBJECTS

ABSTRACT. A description is given of an optical digital system for obtaining three-dimensional information from objects. Grey coding and phase sampling techniques are used to form high resolution images, in which the three object coordinates are codified. It is a powerful and rapid system, that is also capable of dealing with hidden zones. A potential application in the field of Civil Engineering is also described, which will be of use for determining the surfaces of the dynamic simulation models of harbours and coasts.

1. INTRODUCCION

Las técnicas de obtención de imágenes de profundidad, es decir, la codificación de la información tridimensional de los objetos en imágenes tiene un interés considerable en cualquier proceso de modelización o determinación de superficies. En Ingeniería Civil existen campos en los que es interesante contar con sistemas de extracción de información tridimensional. En particular, se ha tratado el caso del problema del estudio de superficies en los modelos utilizados en el Centro de Estudios de Puertos y Costas (CEPYC) del Centro de Estudios y Experimentación de Obras Públicas. En este caso, los objetos son reproducciones a escala de zonas de estudio que son sometidas a la acción de oleaje simulado con el fin de estudiar su evolución en el tiempo. Este tipo de objetos

se caracterizan por su extensión y complejidad topológica, por lo que se requiere un método rápido en la evaluación y robusto en cuanto a que tiene que ser capaz de tratar zonas de dificultad morfológica. Hasta ahora los sistemas de medida utilizados, palpadores, métodos fotográficos, etc., o bien son difíciles de utilizar o no proporcionan la resolución deseada.

Los métodos ópticos, en general, poseen la ventaja de interaccionar con los objetos sólo a través de la luz por lo que se descarta cualquier modificación estructural en el proceso de medida, de ahí su denominación general de «métodos no destructivos»; además, permiten la obtención de gran resolución (1). Esta es la razón por la que se ha realizado un estudio para la utilización de un método de este tipo, que en líneas generales consta de una parte óptica, involucrada en la formación de una imagen, y otra digital, encargada del análisis de la imagen obtenida, hasta llegar a una imagen que contiene información tridimensional del objeto. A partir de aquí, se pueden elaborar otro tipo de representaciones: cortes, curvas de nivel, etc.

(*) Licenciado en Ciencias Físicas. Sector de Técnicas Físicas del Centro de Estudios de Técnicas Aplicadas. CEDEX.

MÉTODOS

Existe gran variedad de métodos de obtención de información tridimensional de objetos (1-2). Entre ellos los más interesantes son los que utilizan una cámara de televisión, ya que en un único proceso de medida se pueden obtener, en principio, tantos valores (x, y, z) como la resolución que se tenga en el sistema de adquisición. Dentro de este grupo se encuentra una amplia gama de técnicas que involucran alguna característica de la imagen producida por un objeto tridimensional. Así se puede encontrar sistemas basados en visión estéreo, análisis de texturas, sombras, etc. (3).

El problema general consiste en que la transformación de un objeto tridimensional en una imagen bidimensional no es invertible. Un píxel en la imagen corresponde a la proyección de una línea infinita en tres dimensiones, consecuentemente, en la transformación se pierde información. La óptica geométrica sólo garantiza que un elemento de imagen corresponde a una línea infinita en el espacio tridimensional, línea que se puede obtener realizando una transformación inversa sobre la imagen. En los sistemas de visión estéreo (4), la solución consiste en contar con otra imagen adquirida con otra cámara desplazada cierto ángulo de la anterior y obtener una segunda línea para el mismo punto sobre el objeto. El punto de corte entre las dos proporcionará un punto en el espacio del que se obtendrá la coordenada de profundidad. Este proceso se denomina «triangulación» y no es más que resolver el triángulo formado por las dos líneas y la recta que une los dos puntos en ambas imágenes. Sin embargo, el problema principal de cualquier algoritmo basado en un sistema de este tipo, es la identificación de puntos entre las dos imágenes, es decir, puntos que se correspondan sobre el objeto tridimensional (2). Aunque existen métodos para identificar puntos comunes en las dos imágenes (3), cualquiera de ellos resultan computacionalmente costosos y en general poco fiables. La razón fundamental es que las dos imágenes pueden diferir considerablemente debido a la transformación de la perspectiva, distinta para cada una de ellas. Además, el objeto, para cada uno de los puntos de vista, presenta diferentes zonas de ocultación. Esto puede resolverse en parte, disminuyendo el ángulo entre las cámaras, lo que implica pérdida de resolución.

Para solventar el problema de la identificación de puntos entre imágenes existe una técnica alternativa que se denomina «codificación de grises» (5). Se trata de la sustitución de una de las cámaras por un sistema de iluminación con luz espacialmente estructurada.

La iluminación consiste en un sistema de proyección de imágenes binarias, es decir, conteniendo elementos de transmitancia cero o uno, de tal manera que se codifiquen las diferentes direcciones de iluminación del objeto en la propia estructura de la luz. De esta forma las imágenes adquiridas con los objetos así iluminados contienen, junto con la proyección del objeto sobre el plano

de la cámara, información sobre la dirección de iluminación.

Además de las técnicas mencionadas existe otra denominada «muestreo de fase» (6). Consiste en el análisis de las imágenes de medidas de fase de la luz reflejada del objeto. En la fase está incluida la información tridimensional del objeto junto con otros términos. En este caso la iluminación también debe hacerse con luz espacialmente estructurada.

Estas dos técnicas, muestreo de fase y codificación de grises, se pueden combinar para proporcionar imágenes de gran resolución.

CODIFICACIÓN DE GRISES

Se supone el plano de proyección coplanar con el plano imagen de la cámara (Fig. 1). Cualquier punto (x, y) de este plano se transforma a través de la superficie en un punto (x', y') sobre el plano de la cámara.

Con esta configuración, la coordenada y coincidirá con y' ; así un punto en el plano imagen, junto con el correspondiente en el plano de la cámara, definen una línea denominada línea equipolar paralela al eje x y al x' . Por tanto, cualquier punto de la superficie con altura z_s queda determinado por la diferencia $\Delta_s = z_s - z_p$ sobre la línea equipolar. Una vez conocido Δ_s y los parámetros fijos del sistema (distancias focales, ángulos y distancia entre focos), se calcula z_s mediante triangulación (7).

En la fórmula de la diferencia Δ_s , el valor de z_p es conocido para cualquier punto de la imagen de la superficie.

Ahora el problema de la identificación entre puntos de la imagen de proyección y la imagen adquirida es el mismo que encontrar la coordenada x_s del punto del plano del proyector que ha sido transformado a través de la superficie sobre el plano de la cámara.

Para esto se proyectan sobre el objeto n patrones de manera que correspondan a la binarización de la coordenada x del plano de proyección para cada valor de $y = y'$. Los patrones consisten en redes binarias de período $\frac{p}{2^n}$ ($n = 0, 1, 2, \dots$). La imagen de cada uno de los

patrones es a su vez binarizada y almacenada sucesivamente en planos de bits consecutivos formando así un nivel de gris para cada píxel de la imagen. Así se consigue codificar en un punto visto por la cámara el valor x_s (la coordenada x del plano de proyección) en el nivel de gris de ese punto.

Como ejemplo en el caso de generar tres patrones se podrían codificar ocho valores de la coordenada x (Fig. 2).

Considerando un plano como objeto, los patrones serán tomados por la cámara sin distorsión, se formará así una imagen con una gradación uniforme en la dirección horizontal (imagen 1). La resolución del método de codificación de grises depende fundamentalmente del

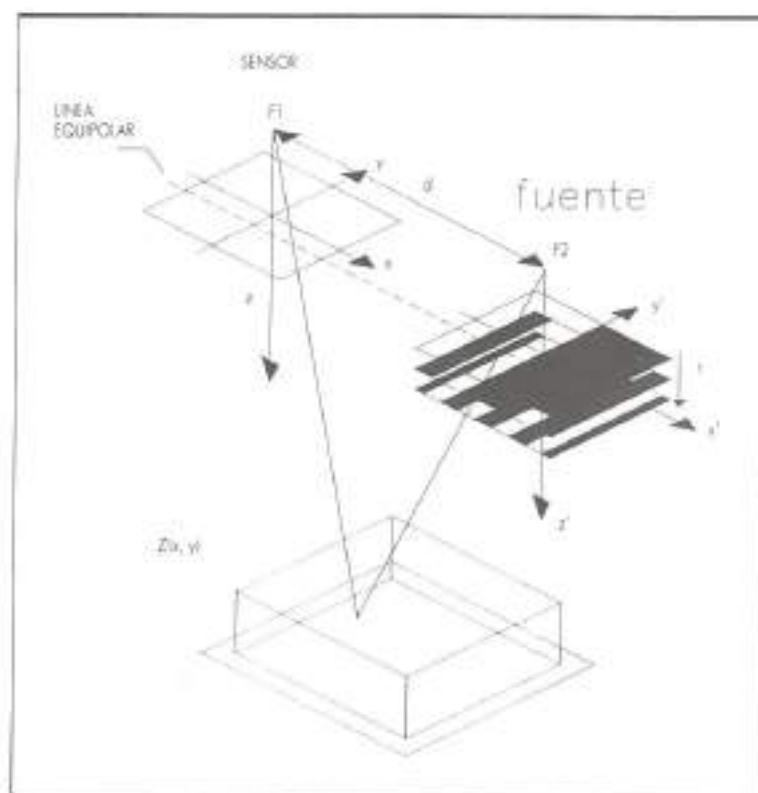


FIGURA 1. Configuración del sistema para la formación de imágenes de código de gris.

número de patrones binarios que se puedan proyectar. Sin embargo, existe la posibilidad de duplicar la resolución de una forma sencilla combinando esta técnica con la de muestreo de fase.

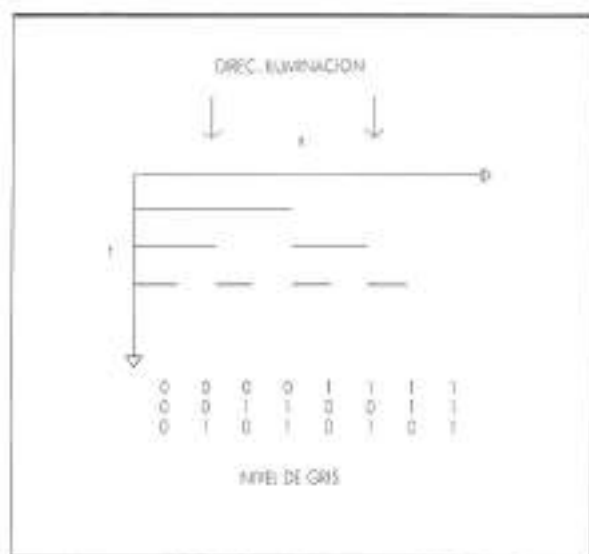


FIGURA 2. Ejemplo de formación de imagen de código de gris con tres patrones.



FOTO 1. Imagen de código de gris para un plano, valor máximo 64.

MUESTREO DE FASE

Existen muchos sistemas que efectúan un muestreo de fase (6); en este caso se realiza mediante la proyección sobre el objeto de una red de paso p . La imagen captada presentará una distorsión geométrica de acuerdo con la forma del objeto (8). Una red se puede escribir como una distribución periódica de una función de transmisión escalón:

$$T(x, y) = \begin{cases} 1 & \text{si } -p/2 < x \leq p/2 \\ 0 & \text{en otro caso} \end{cases}$$

Si sólo se considera el término más significativo de desarrollo de Fourier de una función de este tipo se tiene:

$$T(x, y) = \frac{1}{2} + \frac{1}{2} \cos \left(\frac{2 \pi x}{p} \right)$$

Así el patrón de intensidad sobre el objeto es en primera aproximación:

$$I(x, y) = I_0 + I_1 \cos \left(\frac{2 \pi x}{p} \right)$$

La imagen tomada por la cámara se puede expresar como:

$$I(x, y) = A(x, y) + B(x, y) \cos (\Psi(x, y) + \Phi)$$

con:

$$\Psi(x, y) = \left(\frac{2 \pi z(x, y) (\tan \alpha + \tan \beta)}{p} \right) + f(x, y).$$

$I(x, y)$ es la imagen captada en el detector. El problema consiste en extraer la fase de esta imagen que está enmascarada por $A(x, y)$ que representa la intensidad del fondo y $B(x, y)$ es el contraste de las franjas. Ambos tienen variación espacial y se desconocen en principio. El primer término de fase es el que contiene la función que describe el objeto. En $f(x, y)$ se engloban términos de mayor frecuencia espacial que el término principal que se desprecian. Los ángulos α y β corresponden a los formados entre la perpendicular al objeto y el eje de la cámara y la perpendicular al objeto y el eje del sistema de proyección, respectivamente.

La fase depende periódicamente de la intensidad lo que hace que se pueda determinar únicamente en módulo 2π y además, debido al carácter par del coseno no se puede determinar su signo con una sola medida de la intensidad (6).

Para extraer la fase se toman varios valores $I(x, y)$. Se obtienen mediante el desplazamiento de la red de proyección lo que corresponde a un incremento de fase:

$$\Phi_n = n \frac{2\pi}{N} \quad (n = 0, 1, 2, \dots)$$

En el caso $N=4$, es decir, tomando imágenes para cuatro incrementos (6):

$$\Psi(x, y) = \tan^{-1} \left(\frac{I_2 - I_4}{I_1 - I_3} \right) \quad [1]$$

en módulo 2π (Fig. 3).

Una vez extraída la fase y conociendo el valor de α , β y p se obtiene $z(x, y)$, es decir, la forma del objeto. En la foto 2 se muestra una imagen de fase para en el caso de una superficie plana.

En un caso general no todos los puntos de la imagen $I(x, y)$ pueden ser considerados como válidos (9). Reflexos indeseados, zonas de sombra, pueden dar lugar a

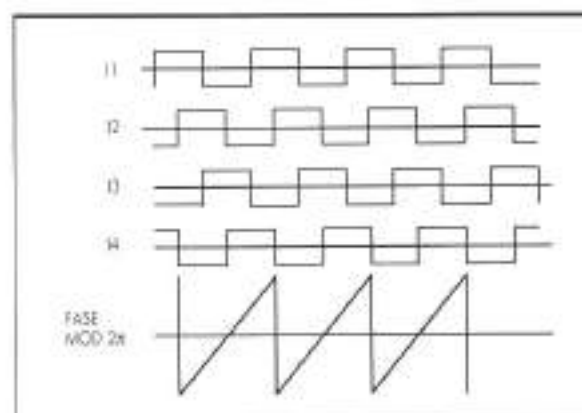


FIGURA 3. Formación de la imagen de fase con cuatro muestras.

pixels de baja modulación. Estos pixels se pueden caracterizar y aislar del proceso calculando su modulación. No es necesario nada más que calcular el término $B(x, y)$ en la fórmula de $I(x, y)$, que como en el caso de la fase involucra la adquisición de cuatro valores de intensidad para cuatro incrementos de fase (6):

$$B(x, y) = \frac{1}{2} \sqrt{(I_1(x, y) - I_3(x, y))^2 + (I_2(x, y) - I_4(x, y))^2} \quad [2]$$

Se establece un valor de umbral descartando todos los pixels (x, y) con $B(x, y) \leq B_0$.

Para obtener la fase real, es decir, suprimir la indeterminación módulo 2π , las técnicas existentes, demodulan cada pixel valiéndose de la información de la fase ya demodulada en pixels de otras zonas de la imagen (10). El problema surge cuando se intenta demodular zonas aisladas del resto por sombras o regiones de baja modulación. En estas zonas es imposible obtener el valor de la fase real ya que se desconoce el salto que se produce al atravesar una zona de baja modulación. Esto sucede cuando la superficie es geoméricamente complicada, es decir, presenta bordes, sombras o reflexiones indeseables.

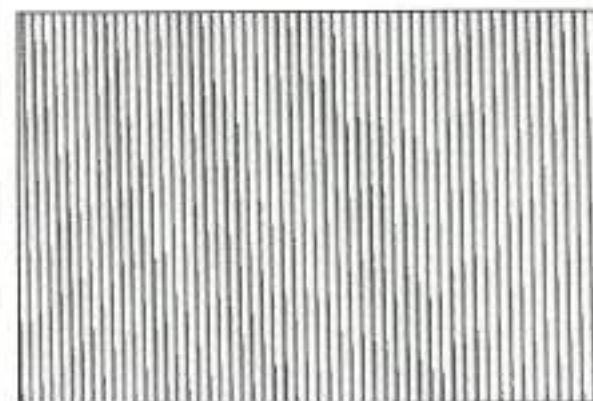


FOTO 2. Imagen de fase modulada para un plano.

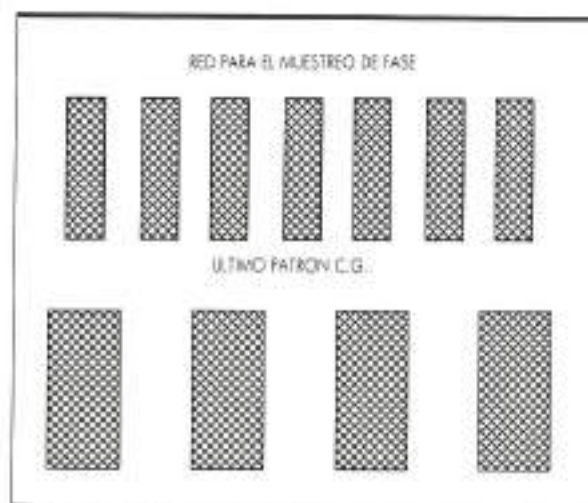


FIGURA 4. Configuración de la última red para la codificación de gris y la que se usará para el muestreo de fase.

Sin embargo, esta técnica produce la mayor precisión posible con una red, ya que, en principio, de todos los píxeles de la imagen se podría extraer información de la fase.

El proceso alternativo de demodulación que aquí se presenta consiste en, una vez capturado y procesado el último patrón de la codificación de grises (bit menos significativo), generar una red de período mitad que este patrón (Fig. 4) para realizar el muestreo de fase descrito. Se considera entonces que la imagen de muestreo de fase corresponde al byte menos significativo de una palabra de 16 bits, siendo entonces el byte de la codificación de grises el más significativo. Se obtiene así una imagen de dos bytes de precisión.

Todo el intervalo de variación entre 0 y 2π de la fase coincidiría con el paso del patrón correspondiente al bit menos significativo de la codificación de grises y, por lo tanto, al intervalo donde el nivel de gris es constante.

El hecho de juntar los dos bytes produce directamente la fase demodulada.

Si esta imagen se resta de otra imagen de fase demodulada, que corresponda a un plano de referencia, se obtiene la imagen del objeto junto con la coordenada z codificada en el nivel de gris de cada punto con dos bytes de precisión.

Como ejemplo para el caso de un plano como superficie y para 3 bits de codificación de grises y 2 de muestreo de fase, se obtiene una pendiente de escalones de 1 bit (Fig. 5).

Este método es insensible a zonas de baja modulación ya que el proceso de demodulación para un píxel es independiente de los demás. Estas zonas se suprimen de la imagen mediante la generación de una máscara asignando un umbral B_{m} , ya que no se puede extraer información de ellas.

IMPLEMENTACION

Adquisición y proceso de imágenes. Para la realización práctica del sistema se ha contado con un ordenador (PC), una tarjeta de procesamiento de imágenes y digitalizador y una cámara de vídeo (CCD). El sistema de adquisición es capaz de tratar imágenes de 512×512 píxeles con una profundidad de 8 bits. La tarjeta de procesamiento cuenta con memoria para trabajar con cuatro imágenes de este tipo; también es capaz de realizar operaciones lógicas y aritméticas entre imágenes.

Sistema de proyección. Los patrones se han generado mediante un sistema de proyección que incluía una pantalla de cristal líquido (LCD) con una resolución de 380 líneas lo cual permite la formación de seis patrones para

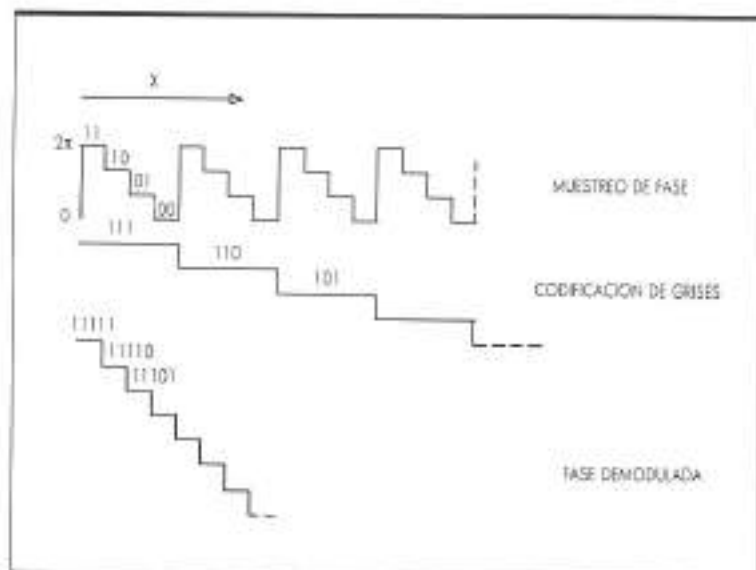


FIGURA 5. Ejemplo de unión de la imagen de fase (2 bits) y la de codificación de grises (3 bits).

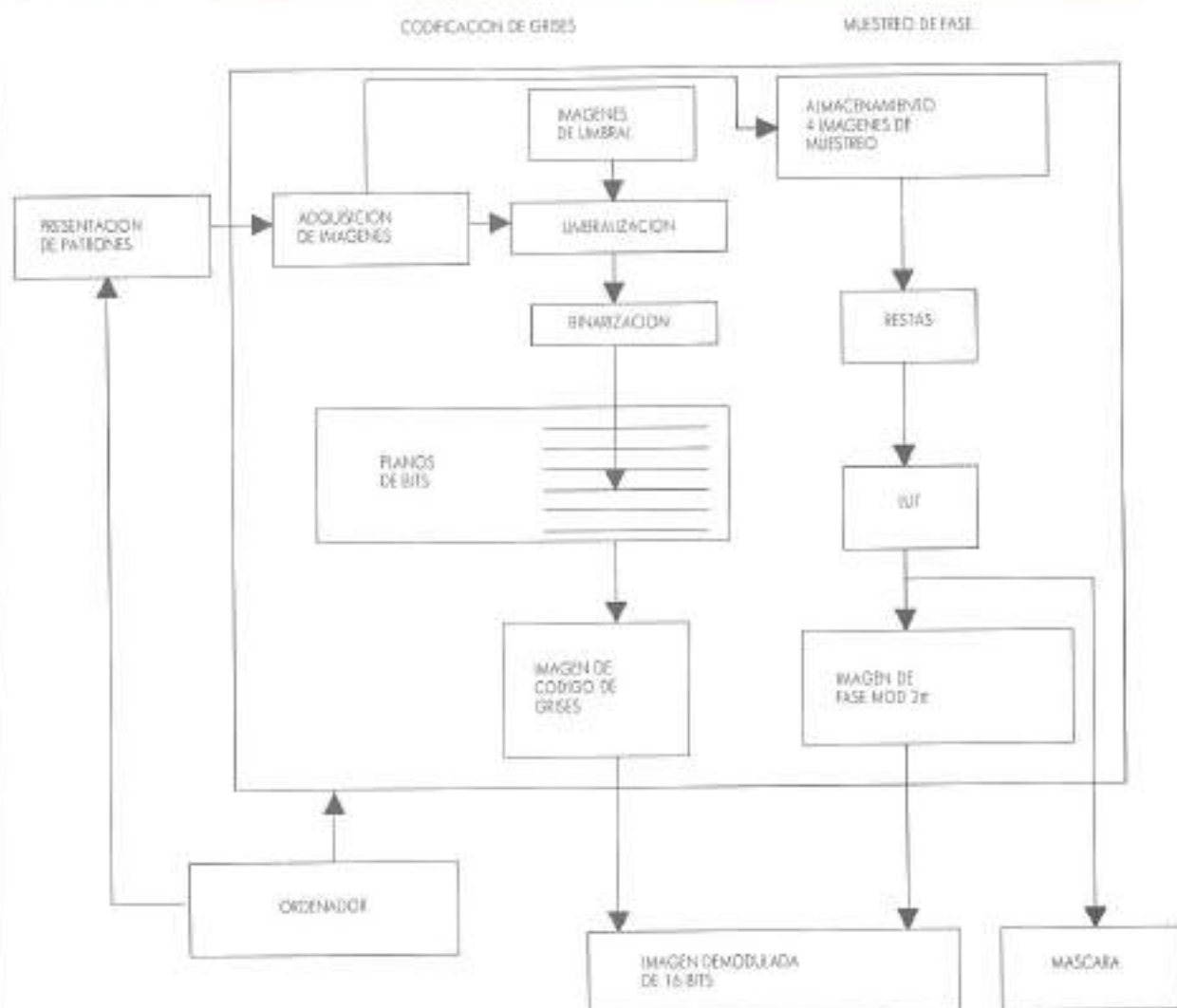


FIGURA 6. Algoritmo. En la caja principal se engloban las operaciones soportadas por la tarjeta de procesamiento de imágenes.

la codificación de grises (imagen de 6 bits) y el resto de la resolución para el muestreo de fase (imagen de 8 bits).

El LCD se controla mediante una tarjeta de tal forma que se puede sincronizar la proyección con la adquisición de imágenes.

Programación

1. Codificación de grises. El programa (Fig. 6) tiene que controlar la proyección de los patrones y la adquisición de las imágenes. Cada imagen se somete a la umbralización pixel a pixel mediante la comparación lógica con dos imágenes de referencia, que corresponden, una al objeto iluminado y otra al objeto sin iluminar. Se evita así una codificación incorrecta debida a cambios zonales de la reflectancia del objeto y efectos de la iluminación ambiente. Una vez umbralizada y binarizada

(píxeles con valores 0 o 1), cada imagen sirve para rellenar cada uno de los planos de bits hasta formar la imagen de grises, en este caso de 6 bits.

2. Muestreo de fase. La implementación de la fórmula 1 es simple. Las operaciones de resta y división vienen soportadas por la tarjeta y se realizan inmediatamente. La implementación del arco tangente se realiza mediante la utilización de LUTs (look-up tables). Son tablas en las que se pueden cargar los valores resultado de una operación matemática. Así, en este caso, con una LUT bidimensional se rellena la tabla con todos los resultados del arco tangente de la división para entradas de las dos restas. De esta forma, el cálculo, una vez almacenadas las cuatro muestras de fase, es instantáneo.

Con el fin de encontrar los píxeles de baja modulación



FOTO 3. Imagen de la LUT bidimensional. En este caso se descartan pixels con valor de umbral inferior a 20 pixels.

se implementa la fórmula 2. Descartar estos pixels no es más que «recortar» en la LUT bidimensional entradas con $B(x, y) \leq B_n$ o, lo que es lo mismo, asignar un valor cero a los pixels que cumplan esta condición (9).

En la imagen 3 se muestra la LUT bidimensional con el «corte» de pixels de baja modulación. Esta LUT sirve para crear una máscara (imagen binaria con valor cero en pixels erróneos) de las zonas a descartar.

Una vez obtenida la imagen de fase demodulada del objeto sólo queda restar la imagen de un plano de referencia y aplicar la máscara.

RESULTADOS

El sistema es independiente de lo complicada que sea la forma geométrica del objeto. Según se aumenta la complejidad el único inconveniente es, en principio, el au-

mento de zonas de baja modulación y por lo tanto, el aumento de zonas enmascaradas que no contienen información útil.

Hay que destacar que el sistema realiza una medida relativa, es decir, es necesario contar con un plano de referencia y con un objeto de calibración. La medida obtenida se refiere a ese plano de referencia.

Se podrían obtener medidas absolutas respecto a un punto del sistema si se calculara la triangulación. En este caso, no sería necesario utilizar un plano de referencia; como contrapartida, el ajuste del sistema sería especialmente delicado (7).

Se ha supuesto que en el caso de la aplicación a los modelos del CEPYC siempre se puede contar con un plano de referencia por lo que no se ha efectuado la triangulación.

Como ejemplo se construyó a escala real una reproducción de una maqueta de las utilizadas en los ensayos de dinámica en el CEPYC (Fig. 7). Se eligió una zona especialmente complicada formada por cubos de 3 cm de lado, en distintas orientaciones protegiendo a un dique que sirve de referencia.

Se situó el sistema a 2-2,4 m del objeto con el eje cámara perpendicular al modelo y el proyector a 0,6 m de la cámara (Fig. 8). Se utilizó un objetivo de 35 mm, lo que permitía cubrir un área de unos 0,5 m². El paso de la red más fina proyectada sobre el objeto fue de 5 mm (para el muestro de fase).

La resolución obtenida fue aproximadamente de 1 mm en z y 1 mm en x e y .

En la resolución en z influye además del número de patrones generados, el ángulo entre el sistema de proyección y la cámara, en la resolución en x e y , la óptica utilizada y la resolución del CCD.

El rango de medida viene determinado fundamentalmente por la profundidad de foco del sistema de proyección y la cámara, en este caso el rango está en torno a 100 mm.

En la foto 4 se observa tal como se capta directamente con la cámara. En las fotos 5 y 6 las imágenes correspondientes a la imagen de fase (byte menos significativo) y la imagen de código de grises. Esta última

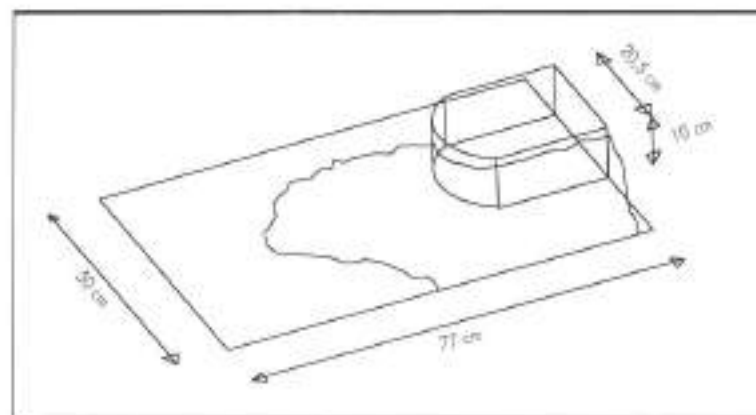


FIGURA 7. Modelo utilizado para la demostración práctica del método.

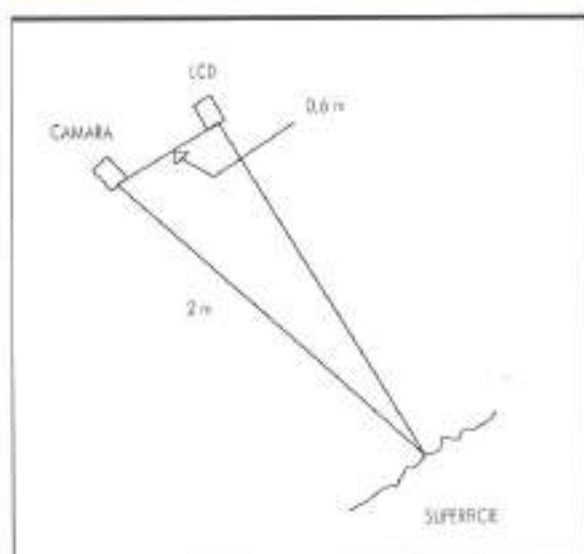


FIGURA 8. Configuración para la demostración.

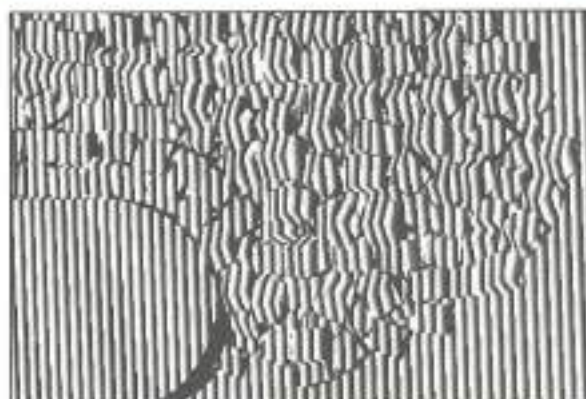
imagen es muy oscura debido a que el valor máximo de un pixel aquí es 64, valor máximo alcanzado con 6 bits. En la foto 7 se presenta la máscara donde las zonas en negro identifican los pixels de baja modulación; la foto 8 corresponde al resultado después de restar el plano de referencia, cada altura del objeto corresponde a un nivel de gris. La imagen ha sido transformada de 16 a 8 bits ya que el sistema de presentación sólo puede mostrar 256 niveles de gris.

La foto 9 es otro ejemplo de la codificación en alturas con tonos de gris, y la 10 otro tipo de representación tridimensional de la misma imagen.

Es interesante destacar la posibilidad de seguir cambios en la superficie, así la foto 11 corresponde a la resta, en lugar de un plano de referencia de otra imagen de la misma zona en la que se ha quitado dos cubos.



FOTO 4. Imagen del objeto tal como es visto por la cámara antes del proceso.



FOTOS 5 y 6. Imágenes de código de grises y fase para el objeto de la figura 7.



CONCLUSIONES

Como se ha expuesto, este método permite una determinación rápida y robusta de superficies. Se pueden destacar los siguientes aspectos:



FOTO 7. Máscara.

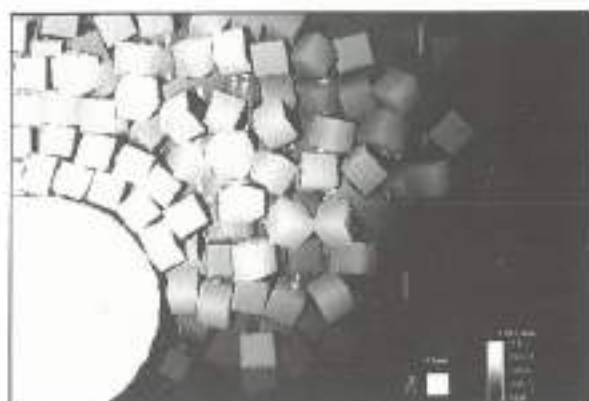
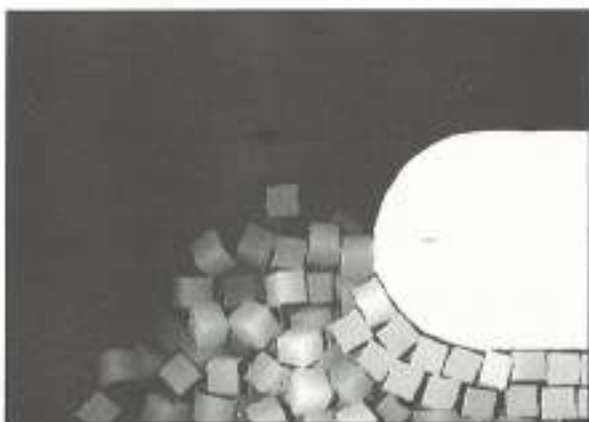


FOTO 8. Resultado después de restar el plano de referencia. La imagen se le ajustó a 8 bits para su presentación, cada ahora corresponde a un tono de gris.

1. Es insensible a las zonas de baja modulación, sombras zonas de ocultación, etc., que simplemente quedan aisladas mediante la máscara.



FOTOS 9 y 10. Imagen del mismo objeto junto con otro tipo de representación tridimensional.



FOTO 11. Resultado de restar, en lugar del plano de referencia, otra imagen con un cambio de estructura. Se observan claramente sólo los cambios.

2. Es en gran parte independiente de la iluminación ambiente gracias a la umbralización puntual.
3. Proporciona una gran resolución tanto en la medida de la coordenada de profundidad (2 bytes) como en x y y (512×512 valores).
4. Es enormemente rápido ya que los cálculos se realizan sobre imágenes completas mediante operaciones implementadas en hardware de forma prácticamente inmediata. El único límite en velocidad es el tiempo entre sucesivas presentaciones de los patrones por el LCD.

Las imágenes obtenidas son de gran calidad sin apenas ruido, directamente utilizables por cualquier otro sistema de representación tridimensional.

Como problemas abiertos quedan la conexión entre diferentes imágenes con el fin de poder determinar superficies mayores que la cubierta por una sola imagen. Sería interesante estudiar la conexión con programas de tratamiento de gráficos 3D y el ajuste entre imágenes tomadas del mismo objeto desde distintas perspectivas con el fin de minimizar los efectos de las zonas de sombra.

BIBLIOGRAFIA

- (1) SIROHI, R. S., y KOTHIYAL M. P. (1991). «Optical Components, Systems and Measurement Techniques». (Optical Engineering núm. 28). Marcel Dekker.
- (2) BALLARD, D., y BROWN, Ch. M. (1982). «Computer Visions». Prentice Hall.
- (3) HUSSAIN, Z. (1991). «Digital Image Processing. Practical Applications of Parallel Processing Techniques». Ellis Horwood.
- (4) MARR, D., y POGIO, T. (1979). «A Computational Theory of Human Stereo Vision». Proc. R. Soc. Lond. B, 204.
- (5) WAHL, F. M. (1986). «A Coded Light Approach for Depth Map Acquisition». G. Hartmann (ed.): Mustererkennung 1986. Springer.

- (6) SCHWIDER, J. (1990). «Advanced Evaluation Techniques in Interferometry». *Progress in Optics XXVIII*.
- (7) STAHS, T. G., y WAHL, F. M. (1990). «Fast and Robust Range Data Acquisition in a Low-Cost Environment». *SPIE*, vol. 1, 395 *Close-Range Photogrammetry Meets Machine Vision*.
- (8) ROSVOLD, O. (1990). «Fast Measurement of Phase using a PC-based Frame Grabber and Phase Stepping Technique». *Applied Optics*, vol. 29, núm. 2.
- (9) VROOMAN, H. A., y MASS, A. M. (1991). «Image Processing Algorithms for the Phase Shifted Speckle Interference Patterns». *Applied Optics*, vol. 30, núm. 13.
- (10) GHIGLIA, D. C.; MASTIN, G. A., y ROMERO, L. A. (1987). «Cellular-automata Method for Phase Unwrapping». *J. Opt. Soc. Am. A*, vol. 4, núm. 1, January.