Captura de la información espacio-angular de escenas 3D

Capture of the spatio-angular information of a 3D scene

H. Navarro^(1,S), M. Martínez-Corral^(1,S), A. Dorado⁽¹⁾, G. Saavedra^(1,S),
A. Llavador^(1,*), B. Javidi⁽²⁾

1. Departamento de Óptica, Universidad de Valencia, Dr. Moliner 50, 46100, Burjassot, Spain.

2. Electrical and Computer Engineering Department, University of Connecticut, Storrs, CT 06269-2157, USA.

(*) Email: Anabel.llavador@uv.es

S: miembro de SEDOPTICA / SEDOPTICA member

Recibido / Received: 05/12/2012. Revisado / Revised: 15/03/2013. Aceptado / Accepted: 16/03/2013.

DOI: <http://dx.doi.org/10.7149/OPA.46.2.147>

RESUMEN:

En este trabajo se han estudiado los principios básicos de los sistemas de captura plenópticos, mediante los cuales es posible capturar no sólo la información espacial, sino también la información angular de un objeto o escena 3D. Esta información espacio-angular queda determinada a través del concepto de *lightfield* o función plenóptica. Para la verificación de los conceptos presentados, se ha procedido a la captura de una imagen plenóptica mediante un sistema compuesto por un conjunto de microlentes distribuidas en disposición hexagonal. De esta manera, se ha podido comprobar que el registro de la información espacio-angular permite reconstruir la escena 3D a partir de una única captura. Esta reconstrucción se ha llevado a cabo, por una parte, mediante un reenfoque en profundidad de la escena. Por otro lado, también ha sido posible llevar a cabo una reconstrucción angular, permitiendo al observador observar la escena desde diferentes perspectivas.

Palabras clave: Sistemas Plenópticos, Captura Espacio-Angular, Lightfield, Reconstrucción 3D.

ABSTRACT:

In this paper, we have studied the basic principles of the plenoptic systems for the capture of the spatio-angular information of a 3D scene. This spatio-angular information is determined through the concept of the lightfield or plenoptic function. In order to verify the concepts presented, we have taken a plenoptic image through a system with a microlens array, arranged in a hexagonal grid. In this way we have checked that the spatio-angular information allows us to reconstruct the 3D scene from only one capture. This reconstruction has been made, on the one hand, through a refocus of the scene at different depths. On the other hand, it has also been possible to carry out an angular reconstruction which allows the observer to see different views of the scene.

Key words: Plenoptic Systems, Spatio-Angular Capture, Lightfield, 3D Reconstruction.

REFERENCIAS Y ENLACES / REFERENCES AND LINKS

- [1]. C. Wheatstone, "Contributions to the physiology of vision. Part the First. On some remarkable, and hitherto unobserved, phenomena of binocular vision", *Phil. Trans. Roy. Soc. London* **128**, 371-394 (1838). [DOI](#)
- [2]. F. Ives, US patent 725, 567, 4 (1903).
- [3]. G. Lippmann, "Epreuves reversibles donnant la sensation du relief", *J. Phys.* **7**, 821-825 (1908).
- [4]. M. Cho, B. Javidi, "Computational reconstruction of three-dimensional integral imaging by rearrangement of elemental image pixels", *J. Disp. Technol.* **5**, 61-65 (2009). [DOI](#)
- [5]. S. H. Hong, J. S. Jang, B. Javidi, "Three-dimensional volumetric object reconstruction using computational integral imaging", *Opt. Express* **12**, 483-491 (2004). [DOI](#)

- [6]. R. Martínez-Cuenca, G. Saavedra, M. Martínez-Corral, B. Javidi, "Enhanced depth of field integral imaging with sensor resolution constraints", *Opt. Express* **12**, 5237-5242 (2004). [DOI](#)
- [7]. R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, P. Hanrahan, "Light field photography with a hand-held plenoptic camera", *Tech. Rep. CSTR 2* (2005).
- [8]. R. Ng, *Digital Light Field Photography*, PhD Thesis, Stanford University, USA (2006).
- [9]. M. Martínez Corral, W. Furlan, A. Pons, G. Saavedra, *Instrumentos Ópticos y Optométricos: Teoría y Prácticas*, Universitat de València, Valencia (1998).
- [10]. T. Georgiev, A. Lumsdaine, "The focused plenoptic camera and rendering", *J. Electron. Imaging* **19**, 021106 (2010). [DOI](#)
- [11]. M. Levoy, "Light fields and computational imaging", *IEEE Computer* **39**, 46-55 (2006). [DOI](#)
- [12]. M. Levoy, P. Hanrahan, "Light field rendering", *Proc. ACM SIGGRAPH'96*, pp. 31-42 (1996).
- [13]. A. Lumsdaine, T. Georgiev, "The focused plenoptic camera", *2009 IEEE International Conference on Computational Photography, ICCP 09*, art. no. 5559008 (2009).
- [14]. R. Ng, "Fourier slice photography", *ACM T. Graphic.* **24**, 735-744 (2005). [DOI](#)

1. Introducción

La obtención de imágenes tridimensionales (3D) de objetos de volumen posee un gran interés en muchas áreas de la ciencia y la tecnología. Los sistemas ópticos tradicionales no resultan eficaces en el registro de este tipo de imágenes ya que en el proceso de captura se produce una fuerte pérdida de información sobre la localización en profundidad de los diferentes elementos en la escena registrada. Para paliar este defecto se han propuesto diferentes técnicas que han supuesto una revolución en el ámbito de la adquisición de imágenes 3D. Aunque en el siglo XIX ya se usaban estereoscopios que proporcionaban imágenes con *sensación* de profundidad, éstos no generaban más que una única perspectiva de la escena 3D, independientemente de la posición del observador [1]. Esta solución, en la que se basan los sistemas actuales de cine y televisión 3D, no proporciona, pues, una verdadera reproducción volumétrica de la escena.

La primera propuesta para la generación de verdaderas imágenes 3D con iluminación natural fue la *Fotografía Integral*, propuesta inicialmente a principios del siglo XX por Ives [2] y desarrollada posteriormente por Lippmann en 1908 [3]. Esta técnica permite obtener imágenes 3D autoestereoscópicas mediante una matriz de microlentes y un proceso simétrico de captura/reproducción. Estas imágenes, que se pueden observar sin ningún tipo de gafas especiales, permiten ver distintas perspectivas del objeto desde distintas posiciones de

observación. A partir de esta idea y del desarrollo de dispositivos digitales de alta resolución, se han implementado recientemente una serie de sistemas de captura y reproducción de escenas 3D tales como los sistemas de Imagen Integral o los sistemas Plenópticos [4-8]. Por otro lado, en los últimos años se han desarrollado otras aplicaciones muy interesantes de dichos sistemas como, por ejemplo, la reconstrucción topográfica/volumétrica de escenas 3D.

Las propuestas presentadas en este trabajo se basan en las mismas ideas que fundamentan la reproducción de la imagen integral o plenóptica para llevar a cabo la reconstrucción de la escena digitalmente. Así, en las secciones 2 y 3 se presentan las bases teóricas de los sistemas de captura plenópticos. La sección 4 está dedicada a presentar los fundamentos del proceso de reconstrucción de la escena 3D, cuyos resultados pueden observarse en la sección 5. Finalmente, en la sección 6 se presentan las conclusiones de este trabajo.

2. La función plenóptica (o *lightfield*)

Un objeto luminoso puede ser considerado como un conjunto de emisores puntuales, emitiendo luz en todas direcciones (Fig. 1). Estas direcciones se representan en el marco de la Óptica Geométrica por medio de los rayos luminosos [9]. Dado un plano arbitrario perpendicular al eje óptico, es posible parametrizar cada uno de los rayos por medio de

sus coordenadas (x, θ) , donde x y θ representan, respectivamente, la posición y la dirección o inclinación del rayo en su intersección con dicho plano [10], tal y como se observa en la Fig. 1.

Con esta parametrización, cada rayo puede representarse como un punto en un diagrama espacio-angular como el que se muestra en la Fig. 2(a). Físicamente, el valor de los distintos puntos representados en dicho diagrama está asociado con la radiancia emitida por cada uno de los puntos del objeto [9]. De esta forma, es posible definir el *lightfield* o función plenóptica como la distribución total de radiancia asociada a un campo luminoso dado [11].

Tal y como se ha definido la función plenóptica, hay que pensar en ella como en una función 5D, 3D para las coordenadas espaciales (x, y, z) y 2D para las coordenadas angulares (θ, ϕ) . Sin embargo, en la mayoría de los casos la escena 3D que se quiere describir es una escena libre de obstáculos (oclusores) y de difusores de la radiación, por lo que la radiancia a lo largo de un rayo permanece constante. Dicho de otro modo, conocida la distribución de radiancia o

Fig. 1. Parametrización de los rayos emitidos por un objeto luminoso mediante las coordenadas x y θ .

Fig. 2. (a) Representación de la radiancia en un diagrama espacio-angular asociado a un rayo luminoso, determinado por sus coordenadas $x - \theta$. (b) Representación del *lightfield* o distribución de radiancia para un conjunto discreto de rayos.

lightfield en un plano, es posible describir la radiancia en toda la escena. Así, esta redundancia permite reducir la dimensionalidad de esta función y considerarla como una función 4D, al eliminar la componente espacial asociada a la dirección axial [11,12]. De esta forma, el *lightfield* 4D vendrá dado por la función $L(x, y, \theta, \phi)$, aunque por simplicidad en las representaciones gráficas, en lo sucesivo se considerará, en este trabajo, el *lightfield* como una función 2D dada por $L(x, \theta)$.

3. Sistemas de captura plenópticos

Los sistemas fotográficos convencionales pueden considerarse, de una forma simplificada, como un sistema formador de imágenes como el de la Fig. 3. De acuerdo con la descripción realizada en la sección anterior, es posible obtener la representación del *lightfield* en el plano imagen. En el caso de objetos con extensión axial, se define un plano central del mismo al que se denomina Plano de Referencia Objeto (PRO). Al conjugado del PRO a través del objetivo fotográfico se le denomina Plano de Referencia Imagen (PRI).

Puesto que los rayos convergen en el plano imagen, el diagrama espacio-angular en dicho plano corresponderá a un conjunto de puntos, de manera que a cada valor de la coordenada espacial le corresponden un conjunto de valores

Fig. 3. Representación simplificada de un sistema fotográfico convencional. Los píxeles del sensor colocado en el PRI capturan la intensidad luminosa acumulada de los rayos que inciden sobre ellos.

Fig. 4. Configuración básica de una cámara plenóptica. El sensor se encuentra separado del plano de las microlentes una distancia f_m , siendo f_m la distancia focal de éstas.

de la coordenada angular (Fig. 2(b)). Sin embargo, los píxeles del sensor colocado en el PRI registran la intensidad luminosa total o irradiancia que incide sobre cada uno de ellos [8], de acuerdo con la ecuación:

$$I(x) = \int L(x, \theta) d\theta. \quad (1)$$

Debido a la integración angular de la Ec. (1), los sistemas fotográficos convencionales sufren una pérdida de la información angular o direccional de los rayos que llegan al PRI. Por tanto, para que el sistema conserve la información angular del *lightfield*, es necesario redirigir los rayos de manera que incidan sobre posiciones (píxeles) diferentes del sensor, contrariamente a como ocurre en una cámara convencional.

Una posible solución a este problema son los llamados sistemas plenópticos. En la Fig. 4 puede observarse el esquema básico de una cámara plenóptica [7,13]. La principal diferencia con respecto a la cámara convencional de la Fig. 3 es la inserción, en el PRI, de una matriz de microlentes de distancia focal f_m .

Por otra parte, tal y como puede observarse en la Fig. 4, el sensor se encuentra desplazado

una cierta distancia con respecto al plano imagen. Por tanto, los rayos que convergen en el plano de las microlentes son separados por cada una de éstas en función de su dirección, impactando sobre píxeles diferentes del sensor, tal y como ocurre en el sistema propuesto por Lippmann [3]. Las imágenes formadas por cada una de las microlentes reciben el nombre de microimágenes. De esta forma, la imagen registrada por el sensor, llamada imagen plenóptica, estará compuesta por un número de microimágenes igual al número de microlentes existentes en el sistema (ver Fig. 10).

Hay que destacar que en esta configuración, la separación entre el plano de las microlentes y el plano del sensor es justamente la distancia focal f_m . Sin embargo, puesto que f_m es usualmente muy pequeña en comparación con la distancia entre las microlentes y el objetivo, se puede considerar, en buena aproximación, que cada una de las microlentes proporciona sobre el sensor la imagen de la apertura del objetivo. Puesto que en la cámara plenóptica el plano objeto (o PRO) está conjugado con el plano de las microlentes, la resolución espacial del sistema estará determinada por el número de microlentes existentes mientras que el número de píxeles del sensor comprendido bajo cada una de las microlentes (o píxeles por microimagen) puede considerarse como una medida de la resolución angular.

4. Reconstrucción de la escena 3D

4.1. La función plenóptica en el plano de las microlentes

Dado un sistema de captura plenóptico como el de la Fig. 4, resulta interesante estudiar cómo es la distribución de radiancia o *lightfield* en el plano de las microlentes en base al diagrama espacio-angular presentado en la sección 2. Para ello, únicamente se considerarán los rayos que atraviesan cada una de las microlentes por su centro, e impactan en los centros de los píxeles [3]. La Fig. 5(a) muestra cómo cada uno de los rayos que atraviesan las microlentes son redirigidos en función de su inclinación, impactando sobre píxeles diferentes del sensor.

De esta forma, es posible identificar los diferentes píxeles en la imagen plenóptica con valores concretos del *lightfield* en el plano de las microlentes (definido como plano x_0), de acuerdo con la Fig. 5(b). Así, en cada una de las microimágenes se registra el *lightfield* para la posición central de cada microlente, (posición en el plano de referencia x_0), y para diferentes valores del ángulo de incidencia θ . Por tanto, se puede obtener una representación del *lightfield* en el plano de las microlentes reorganizando los píxeles de la imagen plenóptica, tal como se muestra en la Fig. 5(b). En esta disposición, cada línea vertical en el diagrama espacio-angular está asociada a una microlente diferente, mientras que cada línea horizontal está asociada a una diferente inclinación de los rayos, o punto de vista de la escena 3D.

4.2. Reconstrucción 3D

Una vez capturada la información 3D a través de la imagen plenóptica, la reconstrucción de la escena se llevará a cabo mediante una reordenación de los datos obtenidos. La

reconstrucción de la escena tridimensional ha de plantearse, por una parte, como un reenfoque de diferentes planos de la escena atendiendo a su profundidad. Por otro lado, la reconstrucción también puede entenderse como un proceso que permitirá generar diferentes perspectivas de la escena cambiando el punto de vista.

4.2.1. Reconstrucción en profundidad (reenfoque)

De acuerdo con la Ec. (1), la integración angular del *lightfield* en un determinado plano proporcionará la imagen de la escena enfocada en dicho plano. Por tanto, dado el *lightfield* en el plano de las microlentes (Fig. 5) la imagen enfocada correspondiente a dicho plano (y por tanto, correspondiente a su conjugado en el espacio objeto de la cámara plenóptica) se obtiene mediante la proyección y suma de los píxeles de cada microimagen en intervalos o "cajas" de tamaño igual al diámetro de una microlente, tal como muestra la Fig. 6. Esta operación equivale a sumar todos los píxeles contenidos en cada microimagen [13].

Fig. 5. (a) Representación 1D de la imagen plenóptica. Cada uno de los píxeles de la imagen se corresponde con valores de la radiancia en el plano de las microlentes x_0 ; (b) reordenación de los píxeles de la imagen plenóptica 1D para obtener la representación de la radiancia 2D.

Fig. 6: Representación del *lightfield* en el plano de las microlentes y proyección de los píxeles para obtener la imagen enfocada en dicho plano.

Sin embargo, uno de los principales puntos de interés de este tipo de sistemas de captura es la reconstrucción, no sólo en el plano de las microlentes, sino también en cualquier otro plano paralelo al mismo. Para ello, se considera un plano cualquiera separado del plano de las microlentes una cierta distancia t . Para analizar cómo es el *lightfield* en este nuevo plano, será necesario realizar una traslación de los rayos luminosos desde el plano inicial hasta éste último. De acuerdo con la Fig. 7, puede demostrarse que en el plano x' , las nuevas coordenadas (x', θ') del rayo se relacionan con las coordenadas (x, θ) mediante la Ec. (2) [13].

$$x' = x + t\theta, \quad \theta' = \theta. \quad (2)$$

De acuerdo con estas ecuaciones, la propagación de un rayo mantiene fija su inclinación con respecto al eje óptico pero varía la posición de impacto con un plano perpendicular al mismo (Fig. 7).

Desde el punto de vista del diagrama espacio-angular, esta transformación puede verse como una cizalladura del *lightfield* en la coordenada espacial, tal y como puede apreciarse en la Fig. 8. Al igual que antes, el proceso de reconstrucción se basa en la proyección y suma de los píxeles,

teniendo en cuenta que ahora la representación del *lightfield* en el diagrama espacio-angular es una versión cizallada del *lightfield* en el plano de las microlentes (Fig. 8). Dependiendo de la separación t entre el plano de reconstrucción y el plano de las microlentes, la deformación en el diagrama espacio-angular será distinta por lo que la reorganización de los píxeles también será diferente. Variando progresivamente el valor del parámetro t , se puede reconstruir la escena, plano a plano.

Fig. 7. Transformación de las coordenadas de un rayo como consecuencia de la traslación, una distancia t , a lo largo del eje óptico.

Fig. 8. Cizalladura del lightfield y proyección de los píxeles para la obtención de una imagen enfocada en un plano distinto al plano de las microlentes.

Fig. 9. Selección de los píxeles de cada una de las microimágenes en función del ángulo de incidencia de los rayos luminosos.

Hay que destacar que la descripción que se ha realizado aquí para el proceso de reconstrucción en profundidad de una escena tridimensional no es única, siendo posible encontrar en la literatura diferentes planteamientos [4,7,8,13,14].

4.2.2. Generación de perspectivas

Además del reenfoque selectivo de la escena 3D, es posible también reordenar los píxeles de la imagen plenóptica para la visualización de la escena desde diferentes perspectivas, permitiendo al observador un mejor conocimiento angular de la escena. Puesto que en una imagen plenóptica la resolución angular está relacionada con la resolución de cada una de las microimágenes, es posible obtener un mayor número de vistas o perspectivas cuanto mayor sea el número de píxeles contenido en cada una de estas microimágenes.

La base de la generación de vistas o perspectivas puede observarse en la Fig. 9. Cada uno de los píxeles que componen una microimagen recoge los rayos luminosos que provienen con una determinada inclinación. De esta forma, considerando el mismo píxel de cada una de las microimágenes, es posible componer una imagen de la escena vista desde ese ángulo [7,8,10]. Las imágenes formadas de esta manera reciben el nombre de *subimágenes*.

5. Resultados experimentales

Para la verificación de los conceptos presentados en las secciones anteriores, se ha considerado una imagen plenóptica capturada mediante un sistema compuesto por 325×375 microlentes, colocadas en disposición hexagonal. De esta forma, la imagen plenóptica presentada en la Fig. 10 consta de 325 microimágenes en la dirección

Fig. 10. Imagen plenóptica capturada mediante un sistema de 325×375 microlentes. Cada una de las microimágenes consta de 10×10 píxeles. Las microlentes se encuentran colocadas en disposición hexagonal, tal y como se muestra en las vistas aumentadas.

Fig. 11. Reconstrucción en profundidad de dos planos diferentes de la escena 3D.

Fig. 12. Escena vista desde dos ángulos o perspectivas diferentes. Puede comprobarse el paralaje existente entre ambas imágenes.

horizontal, y 375 en la dirección vertical. Por otra parte, cada una de estas microimágenes tiene un tamaño de 10×10 píxeles.

La razón por la que se emplean aquí microlentes en disposición hexagonal no es otra que la de conseguir un mayor aprovechamiento de los píxeles del sensor, reduciendo de esta forma las regiones vacías o huecos entre las diferentes microimágenes [8]. La implementación de los diferentes algoritmos de reconstrucción habrá de tener en cuenta, por tanto, la nueva posición de los centros de cada una de las microimágenes capturadas.

En la Fig. 11 puede observarse la reconstrucción de la escena en dos planos diferentes de la misma, de acuerdo con lo expuesto en la sección anterior. La Fig. 11(a) muestra la reconstrucción en el plano de las microlentes mientras que en la Fig. 11(b) la reconstrucción se ha llevado a cabo en un plano diferente. La comparación de ambas imágenes permite constatar cómo a partir de una única captura dada por la imagen plenóptica mostrada en la Fig. 10 es posible producir un reenfoque de diferentes planos de una escena 3D.

Por último, la Fig. 12 muestra dos vistas diferentes de la escena. Ambas imágenes han sido obtenidas mediante la extracción de un único píxel por microimagen, de acuerdo con la

sección 4.2.2. Puede comprobarse que ambas imágenes no son iguales, ya que existe cierto paralaje entre ellas, tanto en la dirección horizontal como en la vertical.

6. Conclusiones

En este trabajo se han estudiado los principios básicos de los sistemas de captura plenópticos, que permiten al captura y procesado del *lightfield* generado por una escena 3D. Para demostrar la utilidad de estos sistemas, se ha procedido a la captura de una imagen plenóptica mediante un sistema compuesto por un conjunto de microlentes distribuidas en disposición hexagonal. Esta imagen ha permitido constatar que, efectivamente, los sistemas de captura plenópticos permiten capturar tanto la información espacial como la angular de una escena 3D. La implementación digital de diversos algoritmos ha permitido realizar la reconstrucción de la escena. La reconstrucción en profundidad ha hecho posible el reenfoque a posteriori de diferentes planos, mientras que la generación de diferentes perspectivas de la escena ha hecho que el observador pueda obtener un mejor conocimiento angular de la misma.

Agradecimientos

Este trabajo ha sido financiado, en parte, por el Proyecto FIS2009-9135 del Ministerio de Ciencia e Innovación, y también por la Generalitat Valenciana (Proyecto PROMETEO2009-077). H. Navarro disfruta de un contrato predoctoral financiado por la Generalitat Valenciana (Programa VALi+d).