

Tecnologías Emergentes

Viendo lo invisible con óptica de ondas virtuales

Autores: Adrián Jarabo, Julio Marco, Adolfo Muñoz, Diego Gutiérrez, Universidad de Zaragoza, I3A.

Palabras clave: ondas virtuales, femtofotografía, láser, reflectancia luz, visión de escenas ocultas, línea de visión, transporte de luz, fotografía computacional.

Líneas Tecnológicas ETID relacionadas: 2.3.1, 2.3.2, 2.5.1, Línea 2.6.

Introducción

En 1878, el fotógrafo inglés Eadweard Muybridge capturó su famosa serie *Horse in Motion*, la primera secuencia de fotografías que capturaba una escena dinámica usando la técnica fotográfica de *stop-motion*. Su invención no solo revolucionó la fotografía, inspirando el cinematógrafo de los hermanos Lumière; por primera vez se pudo analizar la dinámica del galope de un caballo basado en imágenes, permitiendo una mejor comprensión de la dinámica de la locomoción. Con la aparición de sistemas de captura cada vez más rápidos, dicho análisis se trasladó a sistemas cada vez más complejos. Así, casi 60 años después de la invención de Muybridge, Harold Edgerton demostró capturas de eventos ocurriendo en milésimas de segundo mediante técnicas

estroboscópicas ultrarrápidas, incluyendo estudios fotográficos de explosiones nucleares durante el desarrollo de la bomba atómica, o su icónica fotografía *Bullet through Apple* en la que congelaba en tiempo una bala atravesando una manzana (Figura 1, izquierda).

Inspirados por el trabajo pionero de Muybridge y Edgerton, en 2013 investigadores del MIT (*Massachusetts Institute of Technology*) y de la Universidad de Zaragoza, presentaron un sistema de imagen ultrarrápida bautizado como *femtofotografía* [1]. La femtofotografía tomaba la idea de Edgerton y la explotaba hasta nuevos límites: por primera vez, se podían capturar eventos ocurriendo en billonésimas de segundo, velocidad suficiente como ver la propagación de la luz. A esa resolución temporal (un fotograma cada 2 picosegundos) la luz avanza menos de un milímetro por fotograma (Figura 1, derecha). Así, al igual que los trabajos de Muybridge y Edgerton permitieron estudiar la dinámica del mundo con una resolución temporal sin precedentes, la femtofotografía ha revolucionado los campos de visión por computador y reconocimiento de escenas. Problemas históricos como la detección de profundidad o detección de materiales, fundamentales en campos como la robótica o sistemas autónomos, son fácilmente resolubles con la información contenida en el

transporte de luz a escala de picosegundos.

Al margen de estos problemas clásicos, la enorme cantidad de información codificada en el dominio temporal ha abierto nuevas oportunidades y desafíos, permitiendo **ver lo invisible**: desde escenas con escasa iluminación, de poca visibilidad por la presencia de medios turbios como la niebla, e incluso totalmente ocluidas o fuera de la línea de visión (a través de esquinas) [2]. Todas estas modalidades de visión aumentada pueden tener un elevado número de aplicaciones en seguridad, robótica, vehículos autónomos o defensa.

Desafortunadamente, muchas de esas aplicaciones requieren invertir el transporte de luz, infiriendo numéricamente qué escena ha generado la señal medida por la cámara. Esto resulta extremadamente complejo y costoso, y requiere, en la mayoría de casos, realizar numerosas simplificaciones sobre el proceso físico subyacente. Esto limitaba significativamente la aplicabilidad de esta modalidad de imagen a condiciones de laboratorio muy controladas. Para solventar dichas limitaciones, investigadores de la Universidad de Zaragoza y de la *University of Wisconsin-Madison* demostraron recientemente que la información codificada en el dominio temporal del transporte de luz permitía generar en computación

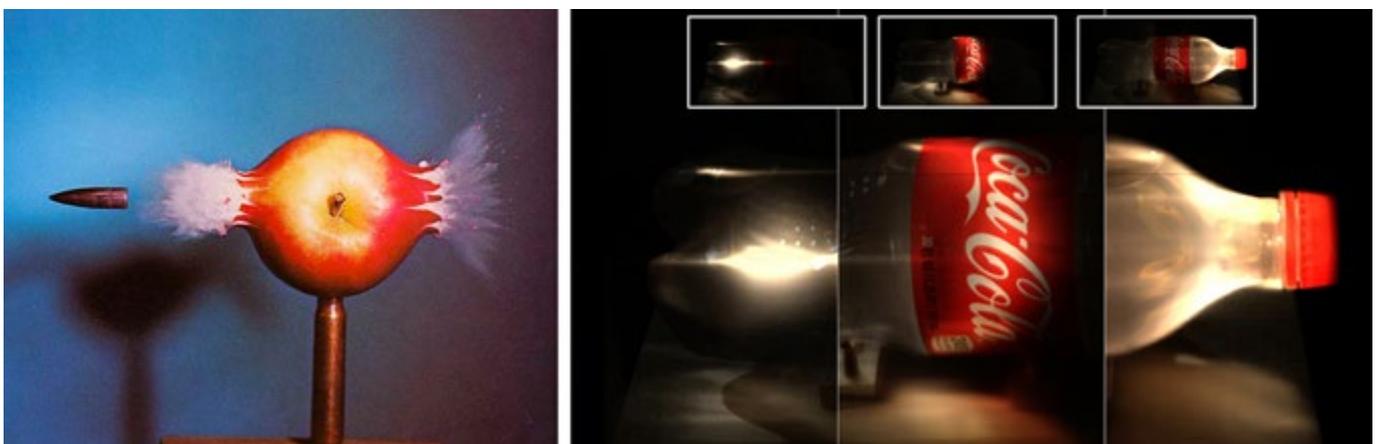


Fig. 1. En 1964, Doc Edgerton capturó su icónica imagen *Bullet through Apple* (© MIT Museum), en la que congelaba una bala atravesando una manzana a alrededor de 850 m/s (izquierda). Casi 50 años después, Velten y colegas introdujeron femtofotografía, una tecnología capaz de capturar la luz en movimiento a casi 300.000 km/s, a un billón de fotogramas por segundo (derecha). La imagen es una composición de los tres fotogramas en el que un pulso láser viaja desde el fondo de la botella hasta el tapón. (Fuente: Jarabo y otros, 2017. *Vis. informatics*. 1).

Tecnologías emergentes

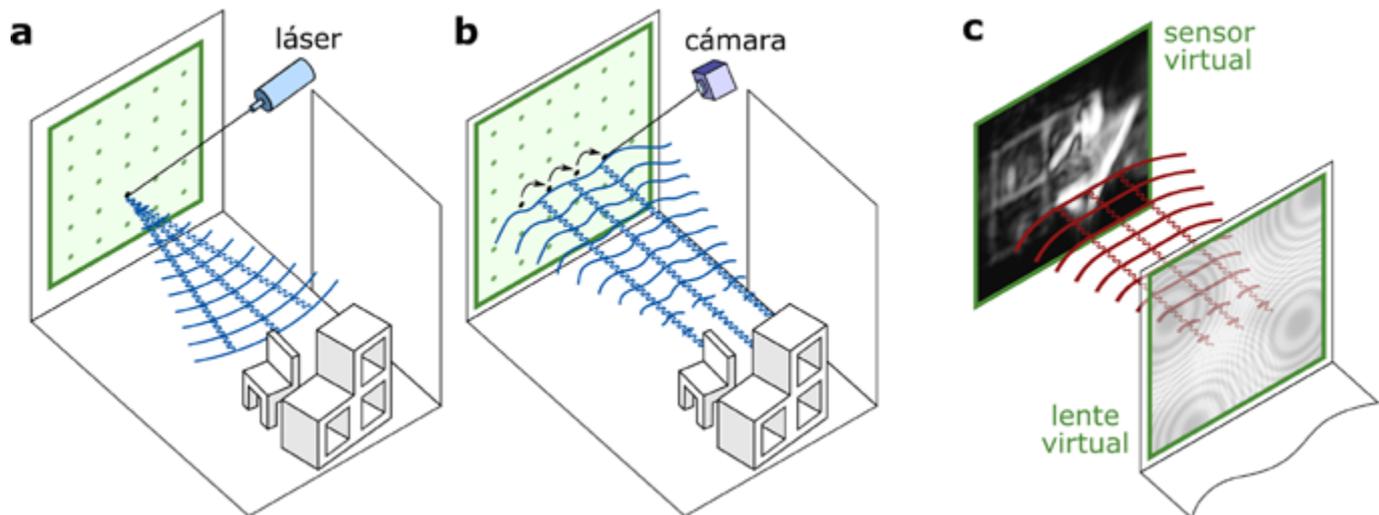


Fig. 2. Reconstrucción de escenas ocluidas como cámara virtual: a) La información de la escena ocluida es capturada iluminando la superficie visible (en verde) con un láser; b) la luz reflejada desde la escena de vuelta a la superficie visible es capturada por la cámara. c) Transformando computacionalmente la luz capturada por la cámara en ondas virtuales (*phasor field*) podemos transformar la superficie visible en una lente virtual, que nos permite enfocar en la escena ocluida de forma análoga a la fotografía convencional. Eso nos permite fotografiar escenas ocluidas. (Fuente: Liu y otros, 2019. *Nature* 572).

un campo de ondas virtuales (*phasor fields*) que se propagaban con los mismos principios que las ondas electromagnéticas. Esta observación dio lugar a una aplicación sorprendente: utilizando principios bien establecidos de óptica clásica se podían generar sistemas ópticos virtuales en computación, que permitían ver a través de esquinas como si se usase un sistema óptico convencional. La gran ventaja de este enfoque es que, al igual que los sistemas ópticos convencionales, se pueden reconstruir escenas de complejidad arbitraria, sin ninguna limitación impuesta por ningún modelo de formación de imagen. Así, por primera vez se podía *fotografiar computacionalmente* escenas de tamaño y complejidad real fuera de la línea de visión.

Viendo escenas fuera de la línea de visión con *phasor fields*

El acceso a la información temporal del transporte de luz mediante imagen ultrarrápida ha dado lugar a un emergente campo en la visión por computador que, en lugar de centrarse en reconocer y reconstruir zonas directamente visibles por parte del sensor, reconstruye partes de la escena fuera de la línea de visión (NLOS, del inglés “*non-line of sight*”). Para ello, se utiliza la iluminación que la escena ocluida refleja de forma indirecta sobre una superficie visible por la cámara. La Figura 2 (a,b) muestra la configuración: se utiliza un sistema de imagen con

iluminación activa controlada, en concreto un pulso láser (la iluminación ambiente se considera ruido en la señal). Dicho láser es emitido hacia la superficie visible, que lo refleja de manera difusa iluminando la escena ocluida. La escena ocluida refleja dicha iluminación indirecta de vuelta a la superficie visible, y ésta es capturada por la cámara.

Debido a que tenemos información del dominio temporal del transporte de luz a escala de picosegundos, podemos calcular el tiempo que le ha costado a un fotón viajar desde la fuente de luz (el láser) hasta la cámara. Este tiempo está directamente relacionado con la distancia recorrida por el fotón desde la superficie visible hacia la escena ocluida, y de vuelta a la superficie visible. Por tanto, el tiempo de vuelo del fotón nos da pistas de dónde ha interactuado el fotón dentro de la escena ocluida, y se puede usar para inferir matemáticamente su camino mediante una inversión del proceso físico del transporte de luz. Sin embargo, aunque este enfoque se ha mostrado moderadamente efectivo en condiciones de laboratorio muy controladas, impone un gran número de limitaciones en las escenas que se pueden reconstruir: objetos aislados de alta reflectancia, en condiciones con poca o nula iluminación ambiental. Esto limita significativamente el rango de aplicabilidad de esta tecnología en situaciones reales.

Este tipo de limitaciones no existen cuando se captura con una cámara una escena directamente visible (LOS, del inglés “*line of sight*”). En ese caso, la cámara es agnóstica a la complejidad del transporte de luz que ocurre en la escena: simplemente toma el campo electromagnético incidente en la apertura, el sistema de lentes lo enfoca en el sensor y el sensor captura el campo enfocado. Motivados por esta observación, Liu y colegas [3] propusieron transformar el problema de reconstrucción de escenas NLOS, en un problema de reconstrucción **LOS virtual**. Para ello, observaron la información en el dominio temporal del transporte de luz en la escena ocluida capturada en la superficie visible, se puede transformar de forma **computacional** un conjunto de ondas virtuales (*phasor fields*) con un comportamiento análogo al de las ondas electromagnéticas en la apertura de una cámara convencional. Estas ondas virtuales transforman la superficie visible en una apertura virtual orientada hacia la escena ocluida, permitiendo modelar todo proceso de emisión, transporte y captura de dichos *phasor fields* utilizando el conocimiento existente en óptica de Fourier. La consecuencia fundamental es que es posible definir computacionalmente emisores y cámaras virtuales sobre las superficies visibles sobre las que capturamos la información de la escena ocluida, y fotografiarla desde nuestra cámara virtual mediante computación

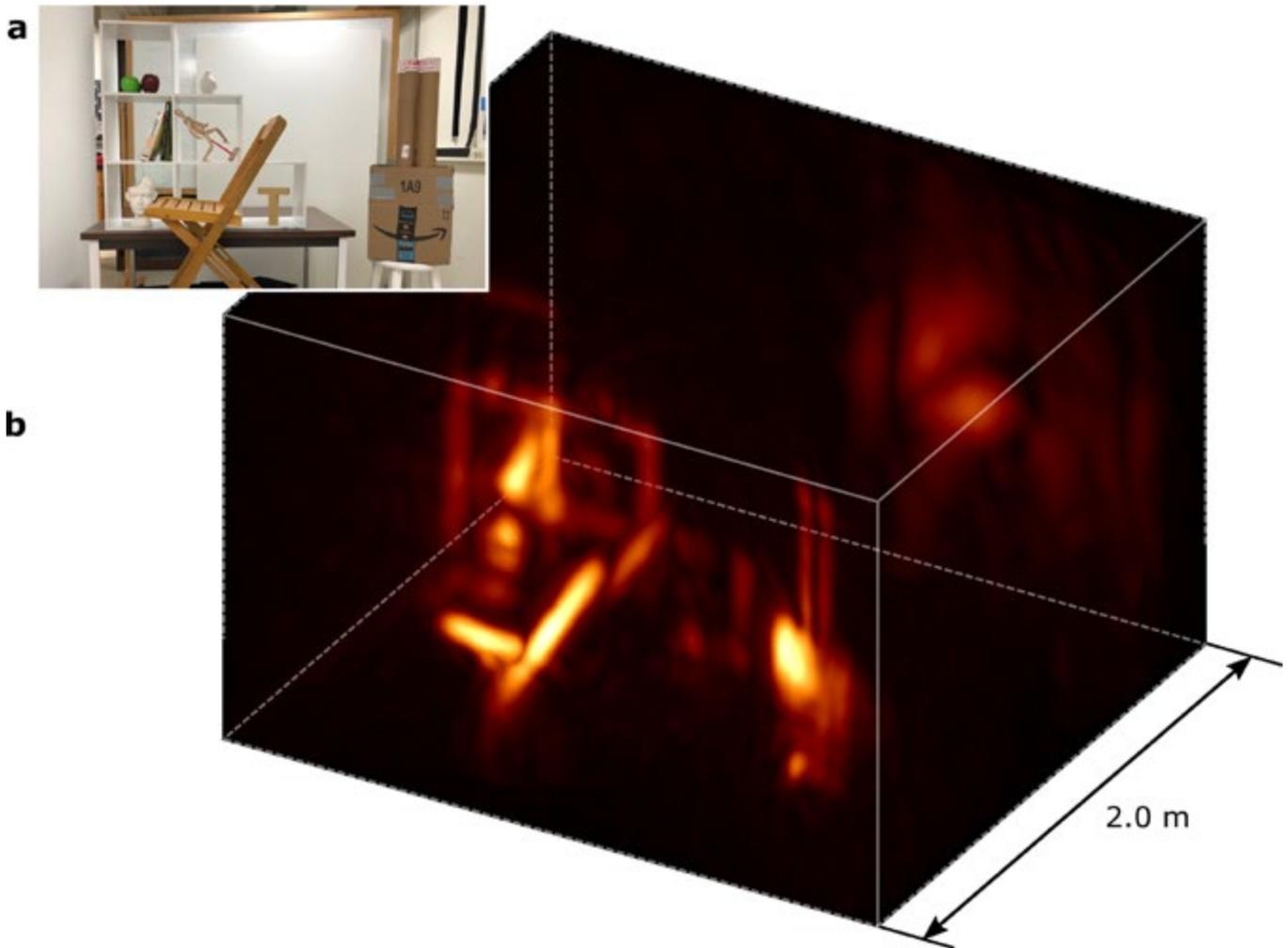


Fig. 3. Reconstrucción de una escena ocluida compleja. a) Fotografía de la escena vista desde la superficie visible por la cámara. La escena contiene oclusiones, múltiples reflectancias, gran tamaño, y ruido en forma de luz ambiental y múltiples interreflexiones. b) Visualización 3D de la escena reconstruida usando *phasor fields* (Fuente: Liu y otros, 2019. *Nature* 572).

(Figura 2, c). Eso permite, por primera vez, fotografiar escenas ocultas independientemente de su tamaño y complejidad, y de forma robusta respecto al ruido en la captura o a la iluminación ambiental en la escena.

Creando cámaras virtuales

Siguiendo la analogía de la cámara virtual, la iluminación que llega a una superficie difusa es equivalente a la recogida por un sensor fotográfico desnudo (sin lente, *bare sensor*), en el que sólo se recoge información de la amplitud del campo incidente (es decir, la intensidad). Esto resulta en una enorme pérdida de información, ya que la información de fase del campo, así como sus altas frecuencias se pierden. En fotografía tradicional, esto se evita posicionando lentes frente al sensor, que enfocan el campo

incidente en determinados puntos del sensor, transformando la fase del mismo. Sin embargo, esto no es posible en superficies arbitrarias.

En este sentido, la mayoría de métodos de reconstrucción de escenas NLOS tratan este problema: tienen que reconstruir la escena a partir de la imagen desenfocada que llega a la pared directamente visible (el sensor virtual). Es la diferencia clave resultante de la introducción de *phasor fields*: debido a que tenemos la información completa de nuestro campo virtual, incluyendo amplitud y fase a su llegada al sensor virtual, podemos introducir tantas lentes como queramos de forma virtual. Estas lentes virtuales operan modificando la fase incidente, lo cual nos permite *enfo-car* una imagen de la escena oculta con nuestra cámara virtual.

Esta flexibilidad para introducir sistemas ópticos arbitrarios de forma computacional tiene un enorme potencial para fotografiar escenas ocultas. Por ejemplo, nos permite definir una cámara que enfoque puntos individuales en la escena oculta, por lo que es posible reconstruir representaciones tridimensionales de la escena, como se observa en la Figura 3. También nos permite generar de forma virtual patrones de iluminación arbitrarios en la escena oculta, lo que nos permite trasladar el vasto conocimiento en técnicas de iluminación estructurada al dominio de las escenas ocluidas. Estas técnicas de iluminación estructurada son comunes en sensores tridimensionales utilizados en robótica, y tienen el potencial de permitir capturas de muy alta resolución de la escena oculta.

Tecnologías emergentes

Aplicaciones y trabajo futuro

La tecnología de imagen ultra-rápida, y en concreto su aplicación a reconstrucción de escenas fuera de la línea de visión, tiene un enorme potencial en campos como la robótica, los vehículos autónomos, situaciones de rescate, seguridad o defensa. La capacidad de reconocer y reconstruir escenas ocluidas permitiría planificar movimientos y detectar potenciales amenazas de forma segura. Una muestra de este potencial es la inversión del Departamento de Defensa de EE.UU. a través de su Agencia de Investigación Avanzada (DARPA) en explorar los límites de esta tecnología a través del programa REVEAL [4]. El salto tecnológico propiciado por *phasor fields*, permitiendo reconstruir escenas complejas, de geometría y materiales arbitrarios, y con iluminación ambiente acerca el objetivo final del uso de esta tecnología en condiciones reales, fuera de laboratorio.

Esta tecnología también tiene un gran potencial en aplicaciones de exploración remota (*remote sensing*) y mapeo en superficies complejas como cuevas, cavidades y simas submarinas, o glaciares. En este contexto, el programa PERISCOPE de la NASA (*National Aeronautics and Space Administration*) plantea el uso de tecnología de reconstrucción de geometría no visible para cartografiar de forma remota cuevas lunares [5], con unos primeros resultados (en la Tierra) muy prometedores.

Por supuesto, para la adopción de la imagen ultrarrápida en entornos realistas de robótica, vehículos autónomos o seguridad aún se requiere de hardware más robusto, portable y barato. Los sensores actuales requieren múltiples medidas y tiempos de exposición de varios segundos, lo cual limita la reconstrucción en tiempo real. El desarrollo de sensores capaces de detectar fotones individuales (*single photon detectors*, SPADs) de gran resolución permitirá hacer un gran número de capturas en paralelo, potencialmente permitiendo tiempos de exposición de decenas de microsegundos, suficientes para tiempo real.

Por otra parte, los entornos reales tienen una enorme complejidad, tanto geoméricamente como en términos de materiales y condiciones de visibilidad. Encontrar superficies adecuadas para su uso como cámaras virtuales

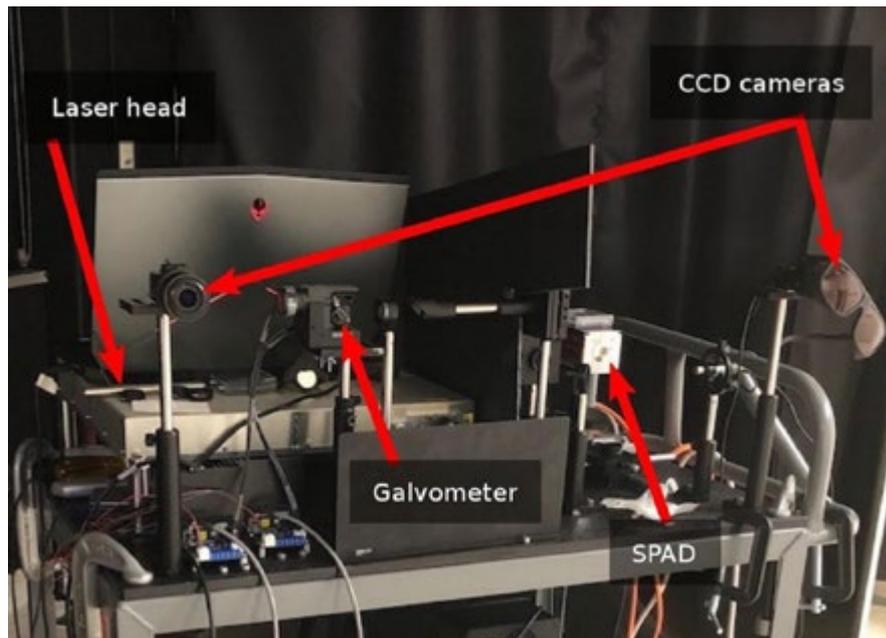


Fig. 4. Sistema de imagen utilizado para la visión de escenas ocluidas. Utilizando un sensor ultrarrápido con sensibilidad para fotones individuales (SPAD, *single photon avalanche diode*) y un láser pulsado, es posible capturar la respuesta del transporte de luz a escala de picosegundo. Esta respuesta temporal permite implementar en computación cámaras virtuales capaces de fotografiar escenas ocultas. Este tipo de sensores se está empezando a producir en masa para su uso en sistemas LiDAR, e incluso en los recientes iPhone 12. (Nótese que el par de cámaras convencionales se usan por motivos de calibración en laboratorio, y no son necesarias. Fuente: [3]).

no es siempre posible, lo cual requerirá de nuevas técnicas de calibración para poder utilizar dichas superficies para observar a través de esquinas. Más complejo aún son las condiciones de visibilidad reducidas, debido a la presencia de medios como p.ej. niebla, humo o agua en entornos subacuáticos. Estos casos son particularmente complejos incluso en escenas en la línea de visión, debido a la degradación de la señal debido a la interacción de la luz con los medios. En estos casos, el potencial de los *phasor fields* es enorme: debido a que podemos controlar computacionalmente cómo emitimos y capturamos mediante de sistemas de lentes virtuales, podemos adaptar los *phasor fields* de modo que pase inalterado por el medio, de modo que éste no afecte a la visibilidad en la escena. Experimentos iniciales a este respecto son enormemente prometedores, y podrían dar lugar a sistemas ópticos robustos a la presencia de medios turbios como niebla en vehículos autónomos, agua en sistemas de exploración submarina, o incluso piel en tecnología óptica de imagen médica.

Finalmente, el campo de imagen NLOS, y en concreto la tecnología

basada en *phasor fields*, ha demostrado que podemos ver objetos ocultos a través de esquinas. Pero también que la señal de rebotes de alto orden (múltiples rebotes de iluminación) en la escena oculta es reconstruible. Esto abre la puerta a definir cámaras virtuales en la escena oculta, y potencialmente ver a través de dos o más esquinas en un futuro cercano.

Referencias

- [1] Velten y otros, 2013. *ACM Trans. Graph.* 32, 44
- [2] Jarabo y otros, 2017. *Vis. Informatics.* 1
- [3] Liu y otros, 2019. *Nature* 572
- [4] REVEAL, *Defense Advanced Research Programs Agency (DARPA)*: <https://www.darpa.mil/program/revolutionary-enhancement-of-visibility-by-exploiting-active-light-fields>
- [5] PERISCOPE, *NASA Innovative Advanced Concepts (NIAC)*: <https://www.nasa.gov/content/periscope-periapsis-subsurface-cave-optical-explorer/>