

# Localizando partículas en un fluido a partir de holografías y computación de altas prestaciones

Gloria Ortega<sup>1</sup>, Julia Lobera<sup>2</sup>, Inmaculada García<sup>3</sup>, María del Pilar Arroyo<sup>4</sup> y Ester M. Garzón<sup>1</sup>

**Resumen**— La tomografía se ha introducido recientemente en la velocimetría de fluidos para proporcionar información en tres dimensiones de la localización de partículas en el seno de fluidos. En concreto, algunos trabajos previos han probado el potencial de la Tomografía Óptica Difraccional en este campo. En general, los métodos de reconstrucción en longitudes de ondas visibles tienen que tener en cuenta la difracción. La Primera Aproximación de Born (aproximación lineal) se ha utilizado para la reconstrucción 3D de imágenes, pero un método no lineal de reconstrucción es necesario cuando la dispersión múltiple no es despreciable. Por lo tanto, si nuestro interés se centra en el desarrollo de técnicas alta precisión de velocimetría de fluidos, una estrategia interesante consiste en incluir en los fluidos un conjunto de partículas sembradoras cuya posición se puede determinar mediante TOD pero el proceso de reconstrucción debe considerar los campos dispersados. En este trabajo se describe una optimización iterativa no lineal para localizar las partículas sembradoras y computar después el campo de velocidad del fluido y, por tanto, métodos de reconstrucción no lineales. Este método de inversión requiere la resolución de la ecuación de Helmholtz, con una alta demanda computacional relacionada con el tamaño del problema. Por lo tanto, la computación de altas prestaciones se hace imprescindible para la solución de este tipo de problemas. Este trabajo muestra los resultados de acelerar esta técnica de reconstrucción utilizando la computación con GPUs y un formato adaptado a las características de la matriz involucrada en la ecuación de Helmholtz.

**Palabras clave**—ODT, computación GPU, Computación de altas prestaciones.

## INTRODUCCION

El análisis clásico de datos Velocimetría de Imágenes de Partículas basada en Holografía (HPIV) asume que la interacción entre el haz de luz de iluminación y las partículas no considera la dispersión múltiple [1], [2]. En la práctica, sin embargo, los efectos de la dispersión múltiple aumentan el ruido de fondo, decrementando el número de vectores de velocidad que pueden ser recuperados de un campo de fluido dado [3]. Los métodos tomográficos que han sido utilizados en los últimos años para mitigar este problema, usan varias holografías desde distintas direcciones de observación [4-5]. En la Tomografía basada en la Velocimetría de

Imágenes de Partículas, las partículas del volumen necesitan ser enfocadas con una apertura numérica (NA) apropiada. Como consecuencia, se requiere un láser poderoso y se sufren algunas distorsiones, especialmente en los flujos líquidos. Además, esta técnica asume una difracción despreciable lo que consecuentemente limita su aplicación a problemas de pequeña escala, tales como los flujos biológicos, donde la densidad de las partículas para una buena resolución espacial produce una considerable dispersión múltiple.

La aplicación de la Tomografía Óptica Difraccional (ODT) en HPIV [6], basada en versiones no lineales mejoraría la resolución espacial.

La ODT es una técnica de radiación no invasiva que suministra un mapa 3D del índice de refracción del objeto a partir de la información holográfica de los campos dispersados al iluminar el objeto con diferentes direcciones de iluminación. Si se considera la aproximación lineal, los componentes espectrales del campo dispersado por el objeto están directamente relacionados con los componentes espectrales del campo del índice de refracción del objeto [7]. En esta línea, recientemente se han hecho algunos avances en microscopía coherente, incluso si estas medidas están incompletas y se asume que la débil dispersión es muy restrictiva.

Para la aproximación de ODT no lineal (NLODT), la reconstrucción de la imagen se plantea como la solución que mejor explica el campo dispersado medido a través de los datos holográficos [8]. La reconstrucción de la imagen se plantea entonces como un problema de optimización. Pero este método tiene un coste computacional considerable que obliga a que estos métodos sean necesariamente implementados con tecnologías de computación de alto rendimiento.

Por otra parte, el uso de la información del objeto que ya es conocida a priori puede servir para reducir el tiempo de computación y la inestabilidad de los resultados [6], [9]. En velocimetría de fluidos, normalmente se tiene información adicional acerca del objeto, por ejemplo el diámetro e índice de refracción de las partículas sembradoras, reduciendo el problema de reconstrucción de la imagen a un problema de localización de partículas [6], [10].

El problema de optimización que se plantea puede ser resuelto con una estrategia de mínimos cuadrados. Este planteamiento incluye la solución de la ecuación de Helmholtz para una distribución de un índice de refracción estimado y para un cierto campo de iluminación. Además, para garantizar una solución estable el muestreo espacial mínimo debe ser aproximadamente una décima parte de la longitud de onda. Debido al tamaño del volumen de interés en

<sup>1</sup> Dpto. de Informática, Universidad de Almería (ceia3), e-mail: gloriaortega@ual.es, gmartin@ual.es.

<sup>2</sup> Centro Universitario de la Defensa de Zaragoza, e-mail: jlobera@unizar.es.

<sup>3</sup> Dpto. de Arquitectura, Universidad de Málaga, e-mail: igarcia@ac.uma.es.

<sup>4</sup> Instituto Universitario de Investigación en Ingeniería de Aragón (I3A), Universidad de Zaragoza, e-mail: arroyo@unizar.es.

flujos biológicos y las restricciones del muestreo el coste computacional del método es bastante alto.

En este contexto, la computación de altas prestaciones (HPC) se necesita para implementar y validar el modelo anteriormente mencionado. HPC permite a la comunidad científica extender sus modelos y acelerar sus simulaciones mediante la explotación de una gran variedad de herramientas de cómputo.

Sin embargo, la selección de las arquitecturas HPC y los interfaces de programación para desarrollar los modelos requiere de un esfuerzo importante.

Hoy en día, las plataformas de HPC están caracterizadas por la heterogeneidad de sus recursos. La mayoría de los supercomputadores modernos consisten en clústeres de nodos multi-core que incluyen elementos aceleradores, como son las GPUs [11].

Algunos trabajos anteriores han mostrado la capacidad de la computación paralela utilizando GPUs a la hora de recrear modelos basados en ODT [12], [13].

En este trabajo nuestro interés se centra en la explotación de la enorme potencia computacional de la GPU. Esta se debe, fundamentalmente, al elevado número de cores que contiene y al gran ancho de banda de su jerarquía de memoria.

Por otra parte, MATLAB es un entorno muy utilizado por la comunidad científica para desarrollar y testear sus modelos [14]. MATLAB puede ser útil para la descripción de modelos porque ofrece un alto nivel de abstracción para operaciones matemáticas fundamentales y proporciona herramientas de visualización que facilitan el ajuste de los modelos en desarrollo. Adicionalmente, otra cualidad muy interesante para el desarrollo de modelos científicos es su posible combinación con interfaces paralelos que permiten explotar distintos recursos de los computadores actuales.

Centrando nuestra atención en las GPUs, está claro que pueden mejorar el rendimiento de MATLAB, si se realiza parte de la computación en la GPU. Esto es debido al hecho de que las rutinas específicas de GPU pueden ser llamadas y ejecutadas directamente desde códigos MATLAB. Para este propósito, aparecen los ficheros MEX-files como la interfaz para la comunicación entre MATLAB y las subrutinas escritas en distintos lenguajes de programación. Aquí, nosotros nos centramos en una implementación de un modelo de ODT no lineal que se ha desarrollado en MATLAB y mediante MEX-files se aprovecha el paralelismo proporcionado por una GPU. Además, este modelo es validado y evaluado computacionalmente.

En este artículo se analiza la mejora en la precisión de la imagen reconstruida con el método ODT no lineal comparado con las estrategias tomográficas lineales. Se analizan experimentos numéricos para volúmenes tridimensionales de  $(10\mu\text{m})^3$ . Está claro que para flujos biológicos de  $(10\text{mm})^3$  se necesita una estrategia multiescalar: una primera imagen menos precisa sería obtenida con el algoritmo de ODT lineal mientras que una imagen con mejor resolución sería obtenida con un algoritmo local de ODT no lineal.

En resumen, las principales contribuciones de este trabajo son: (1) el desarrollo y validación de un modelo de ODT no lineal para la localización de partículas en

situaciones de dispersión múltiple; (2) una explotación eficiente de la capacidad de cómputo de la GPUs para incrementar el rendimiento de la reconstrucción de imágenes con ODT; y (3) comprobación de que el uso de ficheros MEX-files permite a los científicos extender el rango de aplicación de algoritmos de reconstrucción, ya que pueden ser resueltos si se combinan las ventajas de MATLAB para la descripción de modelos junto con las técnicas de HPC.

El resto del trabajo se organiza de este modo. La Sección 2 estudia y analiza el modelo de ODT no lineal propuesto (NLODT-P). La Sección 3 valida el modelo a través de análisis de ejemplos de reconstrucciones TOD. En la Sección 4 se describe la implementación basada en GPU. La Sección 5 ofrece algunos resultados experimentales computacionales al evaluar el modelo utilizando MATLAB y una GPU. Finalmente, la Sección 6 expone las conclusiones y las principales líneas de actuación futuras.

## DESCRIPCIÓN DEL MODELO

En la velocimetría de fluidos, el flujo estudiado es sembrado con un conjunto de pequeñas partículas que tienen un índice de refracción distinto del medio. Una fuente coherente es utilizada para iluminar el flujo, y este haz de iluminación es dispersado por la presencia de las partículas sembradoras, apuntando a su dirección en un instante de tiempo. Para determinar la velocidad de las partículas del fluido, dos grabaciones consecutivas del montaje son requeridas. En particular, en este trabajo nosotros nos hemos centrado en la combinación de varias imágenes tomadas de forma simultánea para recuperar la posición de cada partícula en un volumen de interés en un instante de tiempo. El mismo procedimiento puede ser realizado para cualquier conjunto de grabaciones y, de este modo, se puede rastrear la velocidad de las partículas, consiguiendo determinar la velocidad del flujo.

El método de inversión propuesto consiste en la minimización de la función coste definida por el cuadrado de la diferencia entre el campo medido y el campo dispersado por la presencia de las partículas sembradoras [15]. Esto se puede trasladar en un problema de optimización que puede ser resuelto utilizando un método de optimización modificado basado en el Método del Gradiente Conjugado (CGM) **¡Error! No se encuentra el origen de la referencia.**

La dirección de búsqueda en la primera iteración es el negativo del gradiente de la función coste. Aunque el gradiente se puede expresar como una ecuación analítica, esto aún requiere la resolución del problema *Forward* dos veces en cada iteración. Este problema *Forward* consiste en computar el campo dispersado por el objeto ( $E_{sth}(\mathbf{r})$ ) con un haz de iluminación ( $E_r(\mathbf{r})$ ).

Teniendo en cuenta la teoría de difracción escalar, la amplitud de un campo eléctrico monocromático,  $E(\mathbf{r}) = E_{sth}(\mathbf{r}) + E_r(\mathbf{r})$ , propagándose en un medio de índice de refracción (complejo),  $n(\mathbf{r})$ , obedece la ecuación de Helmholtz **¡Error! No se encuentra el origen de la referencia.** y se puede expresar de la siguiente manera:

$$\left(\nabla^2 + k_0^2 n^2(\mathbf{r})\right)E_{sth}(\mathbf{r}) = f(r)E_r(r) \quad (1)$$

donde  $f(r)$  es el potencial dispersado, definido como:

$$f(r) = -k_0^2(n^2(\mathbf{r}) - 1) \quad (2)$$

Nuestro objetivo es encontrar la posición de partículas que minimicen el cuadrado de la diferencia entre el campo medido  $E_m(\mathbf{r})$  y el campo computado que debería haber sido medido de acuerdo a una estimación del índice de refracción,  $E_c(f, r)$ . Para cualquier holograma existe una contribución separada de la función coste que esta denotada por el índice  $i$ .

$$\text{cost}(f) = \sum_i |E_m^i(r) - E_c^i(f, r)|^2 \quad (3)$$

La estrategia del método del gradiente conjugado ha sido utilizada para minimizar esta función coste. Entonces, dado un potencial dispersado, el gradiente de la función coste se puede tomar como una imagen de la diferencia entre el  $n(r)$  real y su estimación.

Esta imagen es normalmente una distribución homogénea en la que se pueden encontrar varios máximos locales, que indican cambios bruscos en el índice de refracción con respecto al esperado, como es nuestro caso de localización de partículas. Por consiguiente, una matriz relativamente pequeña ( $P$ ) almacena la localización de las partículas que podrían describir el índice de refracción 3D del campo.

Teniendo en cuenta las consideraciones previas, el modelo desarrollado, referido como NLOPDT-P, se detalla más abajo. Un diagrama de flujo con los pasos más importantes del modelo se muestra en la Figura 1. Adicionalmente, la notación utilizada en este trabajo se muestra en la Tabla 1.

#### Paso 1. Localización de la 1ª partícula

El modelo está compuesto por una primera iteración, donde se localiza la primera partícula, y un proceso iterativo en el cual se colocan el resto de partículas. La iteración inicial contiene pasos similares al resto de iteraciones, pero no se necesitan grandes requerimientos de memoria para computarlo. Por lo tanto, se considera de forma separada.

En **¡Error! No se encuentra el origen de la referencia.** se ha demostrado que el gradiente de la función coste se puede expresar como la suma de los hologramas simulados de Bragg entre el campo de iluminación y el campo retropropagado medido.

En particular, para la iteración inicial, al campo de iluminación  $E_r^i(r)$  y al campo de retropropagación medido  $E_m^i(r)$  se les considera inalterados. Entonces, el gradiente es fundamentalmente la Primera Aproximación de Born del potencial dispersado:

$$f(r)^* \approx \sum_i E_m^i(r) * E_r^i(r) \quad (4)$$

donde \* representa el conjugado del valor.

La localización de la partícula coincide con la localización del máximo del valor absoluto de  $f(r)$ . Se puede conseguir un mejor rendimiento del modelo si se aplica previamente un proceso de filtro. Este filtro está basado en el espectro de frecuencias del patrón o partícula tipo. De este modo, permite tener un pico donde encuentre ese patrón. Este filtro se va a denominar de aquí en adelante, matched-filter, y puede ser obtenido considerando la imagen ODT (lineal) de una particular aislada.

#### Paso 2. Actualización del índice de refracción del campo, $n(\mathbf{r})$

Del conocimiento a priori del objeto en la velocimetría de fluidos, nosotros asumimos que el campo del índice de refracción solo puede tomar los siguientes valores: (1) el índice de refracción de las partículas sembradoras en una región alrededor de la partícula localizada; y (2) el índice de refracción del fluido en el resto de vóxeles del volumen de interés.

#### Paso 3. Computar la actualización del gradiente, $g(\mathbf{r})$

El gradiente  $g(\mathbf{r})$  de la función coste proporciona la distribución que se debería añadir al potencial dispersado  $f(\mathbf{r})$  para minimizar la función coste de acuerdo al Método de Gradiente Conjugado clásico. Una vez estimado el campo del índice de refracción, una expresión similar a la Ecuación (4) puede ser obtenida del nuevo gradiente  $g(\mathbf{r})$ . Sin embargo, el significado de ambos campos de interferencia y su estrategia computacional han cambiado:

- **Actualización del campo de iluminación:**  $E_{r,n}^i(\mathbf{r})$ . Nosotros necesitamos tener en cuenta la presencia de las partículas que ya han sido localizadas. El procedimiento *Forward* computa el campo dispersado  $E_{sth}^i(r)$  debido al objeto descrito por  $n(\mathbf{r})$  y el campo de iluminación original  $E_r^i(\mathbf{r})$ . El láser de iluminación será la suma de ambos campos  $E_{r,n}^i = E_r^i(\mathbf{r}) + E_{sth}^i(\mathbf{r})$ .
- **Actualización del campo medido:**  $E_{m,n}^i(\mathbf{r})$ . La retropropagación del campo medido es computado en dos pasos: Primero  $E_{sth}^i$  tiene que ser filtrado para tener en cuenta la apertura numérica y la situación de los dispositivos que recogen los campos y entonces poder obtener el campo medido esperado:  $E_c^i(\mathbf{r})$ . En segundo lugar, el campo medido que todavía no ha sido explicado por las partículas ya localizadas ( $E_m^i - E_c^i$ ) es retropropagado teniendo en cuenta la difracción introducida por las partículas localizadas. El solver *Forward* provee el campo dispersado  $E_{m,n}^i$  debido a  $n(\mathbf{r})$  siendo el láser de iluminación  $(E_m^i - E_c^i)^*$ . La suma del campo dispersado  $E_{m,n}^i$  y el láser de iluminación, son tomados como el campo medido actualizado.

Como en la iteración inicial, una contribución separada al gradiente  $g(\mathbf{r})$  es obtenida para cada holograma.

#### Paso 4. Localización de la siguiente partícula y salida

Como para la iteración inicial, la localización más probable de la partícula es el máximo absoluto del matched-filtered del gradiente. La posición es almacenada en la variable de salida  $P$ , y el valor pico, *value*, es utilizado para decidir el final del modelo. La experiencia ha demostrado que el valor decrecerá en cada iteración y un umbral conveniente asegura la parada del proceso. Sin embargo, el criterio para seleccionar un valor apropiado del umbral no está detallado en este trabajo. Por lo tanto, el máximo número de iteraciones ha sido el criterio impuesto.

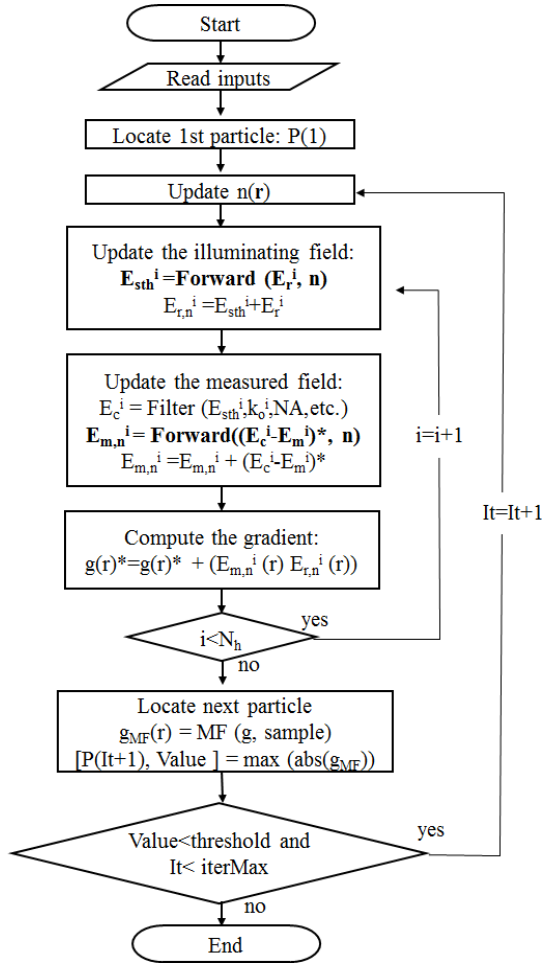


Fig. 1. Diagrama de flujo que presenta los pasos más importantes involucrados en el modelo NLODT-P. Los pasos que usaron la computación basada en GPUs están marcados en negrita.

En cada iteración del problema de optimización, la ecuación de Helmholtz (procedimiento *Forward*) tiene que ser resuelta. La ecuación de Helmholtz es un ejemplo de una Ecuación Diferencial Parcial lineal elíptica (PDE), que ha sido extensamente estudiada. **¡Error! No se encuentra el origen de la referencia..** Esta ecuación puede ser numéricamente resuelta utilizando una transformación basada en las funciones de Green y una discretización espacial [16], por ejemplo, el Método de los Elementos Finitos (FEM). **¡Error! No se encuentra el origen de la referencia..** FEM discretiza la región de interés en pequeños elementos, asumiendo que la función  $E(\mathbf{r})$  puede ser aproximada a un valor constante en cada uno de esos elementos. Una malla regular de elementos es normalmente considerada cuando la forma del objeto es la variable desconocida (problema inverso). Por lo tanto, las derivadas espaciales del operador de Laplace pueden ser discretizadas con un estencil de 7 puntos en 3D.

Para una malla tridimensional, el sistema de ecuaciones lineal resultante de la Ecuación (1) se describe por una matriz con solo 7 diagonales no nulas. Por lo tanto, FEM transforma las ecuaciones 3D de Helmholtz en un sistema lineal ( $Mx = b$ ), donde  $b$  es el termino independiente,  $x$  está relacionado con el campo dispersado y  $M$  está relacionado con índice de

refracción,  $M$  es dispersa y exhibe un patrón muy regular que depende del número de puntos de discretización o vóxeles del volumen ( $Vol$ ) **¡Error! No se encuentra el origen de la referencia..**

Es bien conocido que el “efecto de polución” limita la fiabilidad de la solución FEM de la ecuación de Helmholtz cuando se tiene un alto número de ondas **¡Error! No se encuentra el origen de la referencia..**

TABLA 1

NOTACIÓN UTILIZADA EN ESTE TRABAJO

Abreviación	Descripción
$E_r^i$	Campo de iluminación del holograma $i$ -th
$E_{sth}^i$	Campo dispersado del holograma $i$ -th
$n$	Índice de refracción estimado del objeto (teniendo en cuenta las partículas que ya han sido localizadas previamente)
$E_m^i$	Campo medido del holograma $i$ -th
$E_{m,n}^i$	Campo medido del holograma $i$ -th, dado un $n(r)$
$g(r)$	Gradiente de la función coste
<i>sample</i>	Imagen de una partícula aislada
$Vol$	Número de vóxeles utilizado para la discretización del volumen a reconstruir
$N_h$	Número de hologramas
<i>iterMax</i>	Número máximo de iteraciones del modelo
$P$	Localizaciones de las partículas
<i>Value</i>	Valor máximo del módulo de $g(r)$

Para evitar este efecto de inestabilidad, el volumen de interés debe ser muestreado. Las muestras del volumen de interés (vóxeles) deben estar localizadas de tal forma que el vecino más cercano no esté más alejado de una décima de la longitud de onda. Por tanto, el valor de  $Vol$  es muy alto. Como consecuencia, es necesario resolver un gran sistema de ecuaciones lineales de aritmética compleja en el procedimiento *Forward*. De este modo, dicho procedimiento es el más costoso computacionalmente en el modelo NLODT-P.

#### EXPERIMENTOS NUMÉRICOS

La concentración de partículas determina la resolución espacial de la velocimetría de fluidos. La NLODT permite abordar el problema de la dispersión múltiple, pudiendo incrementar el rendimiento de la velocimetría de partículas clásica. La dispersión múltiple puede ser debida a: (1) el tamaño finito de la partícula y (2) la presencia cercana de otras partículas que modifican tanto el campo de iluminación como el campo dispersado por cada partícula. NLODT-P pretende resolver el problema de la localización de partículas gracias a la estrategia de *matched-filtering* y un procedimiento iterativo.

La relevancia del ruido de la imagen depende también de la configuración óptica. Hemos considerado dos estudios numéricos con diferentes configuraciones: con acceso óptico completo (0) y con acceso óptico limitado (0). Para validar el modelo, hemos escogido una distribución aparentemente sencilla de las partículas que consiste en cuatro partículas de  $2\mu\text{m}$ . La dirección de iluminación para cada holograma es escogida entre los ejes  $x$ ,  $y$  y  $z$ , respectivamente (ilustrado en la 0a por la flecha de color rojo). Las partículas están localizadas de forma que cada partícula está obstaculizada por las otras.

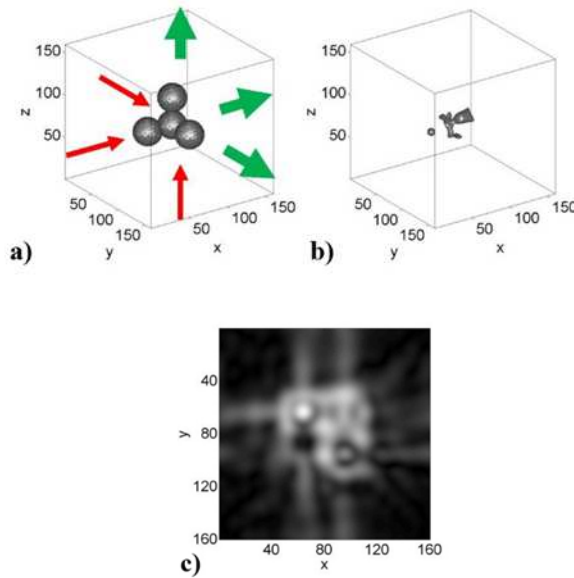


Fig. 2. Vista 3D del problema de distribución de partículas con una configuración de acceso óptico completo (a). Imagen LODT: Vista 3D de los voxeles más brillantes (b); y vista 2D del plano  $z = 64$  pixeles (c).

Hemos considerado una iluminación típica coherente usando un láser He-Ne, con  $\lambda = 0.633\mu\text{m}$  y los hologramas han sido registrados con  $\text{NA} = 0.55$ . El índice de refracción de las partículas es  $n = 1.33$  (agua) y el medio en el que se inserta es aire ( $n = 1.00$ ). El volumen de interés ha sido dividido en  $160 \times 160 \times 160$  vóxeles de una décima de longitud de onda. Primero, por simplicidad, la dirección de observación ha sido elegida para que coincida con la dirección del ángulo de iluminación de cada holograma (marcada con flechas de color verde). Esta configuración maximiza la extensión del espectro medido dispersado por los 3 hologramas, pero requiere de un acceso óptico completo.

La Figura 2b muestra la forma 3D de la superficie dibujada por los 1% vóxeles más brillantes de la imagen de LODT del problema de distribución de partículas. Para ilustrar mejor la imagen borrosa, el módulo del gradiente en el plano  $z = 64$  pixeles es también mostrado en la 0c. Está claro que la posición de las partículas no se puede recuperar de la imagen con LODT. La relevancia del tamaño finito de la partícula en la dispersión múltiple es considerable cuando se considera una sola partícula aislada. La imagen con LODT de una partícula recuperada de un montaje similar con 3 hologramas ha sido computada de forma separada: la vista 3D se muestra en la 0a, y el módulo del plano central en la 0b.

La imagen LODT de una particular aislada está bastante alejada de una esfera e incluso de la distribución Airy-like. Esta amplitud compleja es referida como *sample* en la sección previa ya que patrones similares se podrían esperar alrededor de cada posición de partícula de la imagen LODT para cualquier otra distribución de  $2\mu\text{m}$ .

De este modo, *sample* se puede utilizar para computar un matched-filter sobre la imagen LODT de la 0. Tras el uso del matched-filter sobre la imagen de LODT, las

posiciones de los 4 picos se pueden identificar claramente (0b).

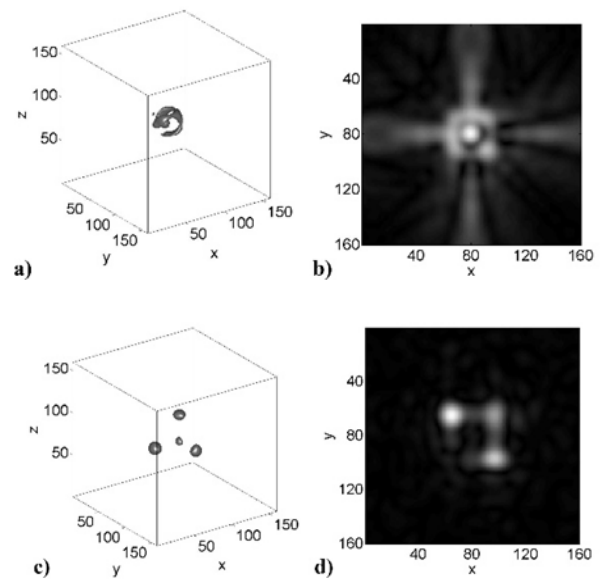


Fig. 3. Imagen de LODT de una partícula aislada (*sample*): (a) Vista 3D y (b) vista 2D del plano central (en  $z = 81$  pixeles). Imagen filtrada de LODT del problema de la distribución de partículas: (c) Vista 3D y (d) Vista 2D del plano  $z = 64$  pixeles.

Como era de esperar, NLODT-P también puede resolver este problema de localización de partículas. En cada iteración, *sample* también se usa para computar el gradiente filtrado de la función coste. El pico más brillante del gradiente filtrado determina la localización de la partícula en cada iteración. La Fig. 4 muestra de izquierda a derecha las imágenes del gradiente filtrado de la primera iteración (que coincide con la imagen filtrada de LODT) y las siguientes iteraciones. Para este experimento, la LODT no puede resolver la distribución de las partículas; mientras que la LODT con el filtro y la NLODT si pueden resolver el problema, obteniendo errores similares, alrededor de  $2.1$  pixeles ( $0.13\mu\text{m}$ ). Por tanto, el uso de un matched-filter considerando la información conocida a priori de las partículas, ha sido suficiente para obtener la localización de las partículas.

Un segundo experimento ha sido considerado utilizando un montaje más realista, con un acceso óptico limitado. Las direcciones de iluminación también son 3 direcciones ortogonales, como antes, pero todos los hologramas se han registrado desde la misma dirección de observación. La dirección de observación se ha escogido como se muestra en la 0a, por lo tanto, se forma un ángulo de  $45^\circ$  con cualquiera de las direcciones de iluminación. Está claro que la resolución espacial se reduce y las imágenes de las partículas están localizadas a lo largo de la dirección de observación. La LODT no es capaz de resolver este problema de localización de partículas con esta configuración de la observación. Además, después de computar con el correspondiente *sample* y las imágenes filtradas obtenidas con LODT (0b and c), sólo se han podido localizar tres partículas.



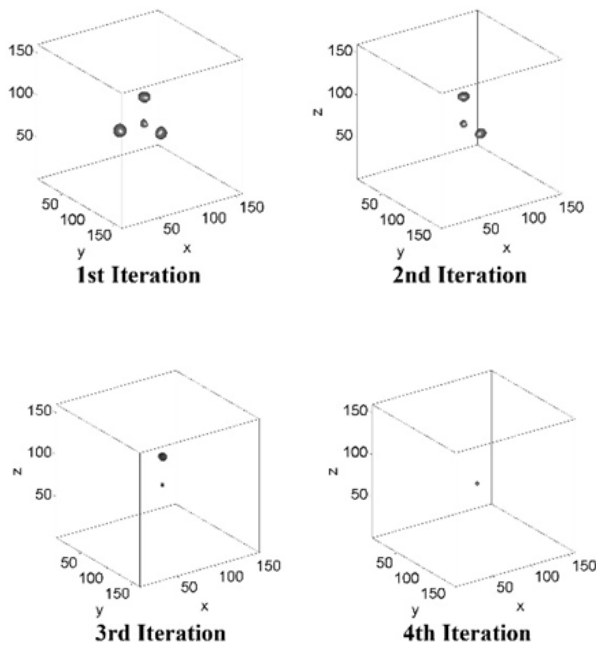


Fig. 4. Gradientes filtrados de la NLODT (imágenes de las partículas que tienen que ser localizadas) para una configuración con acceso óptico completo.

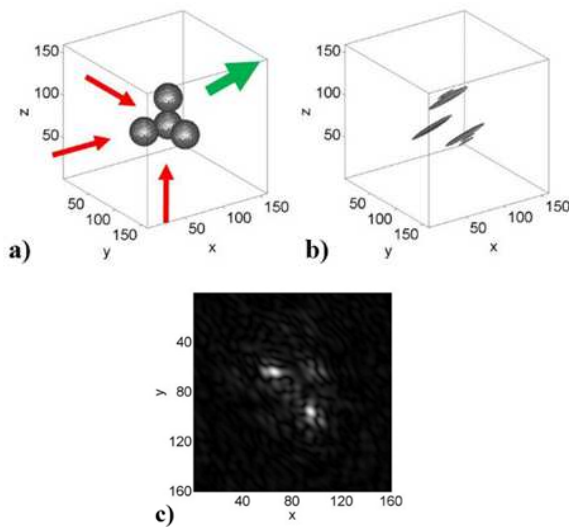


Fig. 5. Vista 3D de una configuración con acceso óptico limitado (a). La imagen filtrada de LODT: vista 3D de los vóxeles más brillantes (b) y vista 2D en el plano  $z = 64$  píxeles (c).

El gradiente filtrado obtenido en cada iteración del modelo NLODT-P se muestra en la 0. De la figura, se puede apreciar que la partícula situada en la esquina inferior es más evidente en cada iteración consecutiva.

El error de la posición final es  $0.3\mu\text{m}$ , aunque el rendimiento de NLODT-P es claramente más ventajoso en comparación con la estrategia lineal, lo cual se puede observar en estos experimentos de pequeño tamaño. De forma adicional, la combinación de MATLAB y la computación con GPUs para resolver el procedimiento *Forward*, discutido en la Sección 4, hace posible el estudio de este tipo de problemas y, por tanto, el uso de NLODT-P para aplicaciones de velocimetría de fluidos.

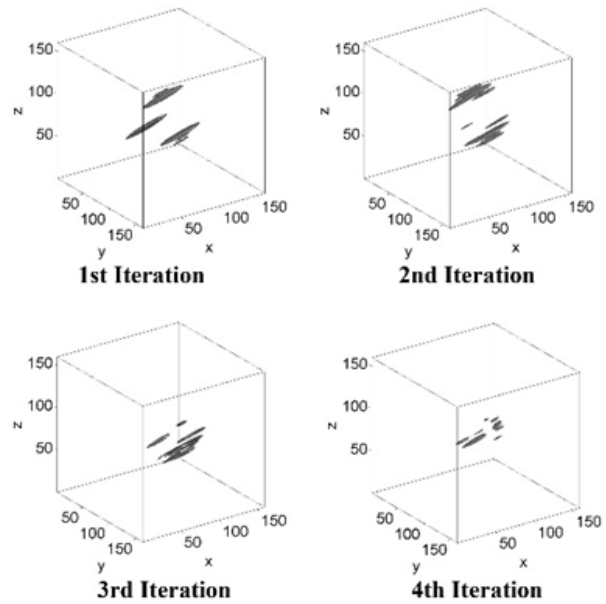


Fig. 6. Gradiente filtrado de la NLODT para una configuración con acceso óptico limitado.

#### IMPLEMENTACIÓN DE NLODT-P BASADA EN UNA GPU

El modelo NLODT-P tiene unos altos requerimientos de memoria y consume un gran tiempo de ejecución, ya que cada iteración requiere la resolución de varias ecuaciones de Helmholtz. Por lo tanto, para la implementación de NLODT-P es esencial aplicar aproximaciones para reducir los requerimientos de memoria así como usar la computación de altas prestaciones (HPC) para reducir el tiempo de cómputo. Nosotros hemos desarrollado una estrategia para acelerar la resolución de la ecuación de Helmholtz que está basada en la explotación de las regularidades de la matriz dispersa de coeficientes (anteriormente denotada como  $M$ ) y el uso de la computación GPU [19].

El procedimiento con mayor demanda computacional en NLODT-P es el *Forward* que está dedicado a la resolución de la ecuación diferencial de Helmholtz. Hemos considerado el Método del Gradiente Biconjugado (BCG) para resolver el gran sistema de ecuaciones lineal que se obtiene de la discretización de dicha ecuación [19]. Este solver computa de forma iterativa dos productos matriz dispersa vector (SpMVs) que consumen la mayor parte del tiempo de ejecución del BCG [20].

Teniendo en cuenta que la matriz de coeficientes,  $M$ , exhibe varias regularidades cuando está relacionada con la discretización de la ecuación 3D de Helmholtz, se ha considerado un formato específico en este trabajo que almacena la información mínima que define a la matriz dispersa. De este modo, los requerimientos de memoria y el tiempo de ejecución del BCG son reducidos de forma considerable ya que este nuevo formato requiere menos accesos a memoria para leer los elementos de la matriz dispersa. Este formato se ha llamado *Regular Format* (RF). Los requerimientos son 10 veces menores cuando el formato RF se utiliza. Más detalles sobre RF se pueden encontrar en [¡Error! No se encuentra el origen de la referencia.] y en [21].

Adicionalmente, un BCG específico basado en RF se ha acelerado utilizando la computación con GPUs. CUDA es un interfaz paralelo que permite incrementar el rendimiento en la computación con GPUs [22].

Versiones preliminares de NLODT-P han sido descritas mediante MATLAB. Sin embargo, es necesario integrar ambos lenguajes (MATLAB y CUDA) para lograr la facilidad de uso de MATLAB y la aceleración de la computación con GPUs. Rutinas MEX-files específicas se han desarrollado para acelerar el procedimiento con mayor coste computacional del modelo (procedimiento *Forward*). Los MEX-files son subrutinas escritas en C, C++ o CUDA que pueden ser dinámicamente enlazadas con MATLAB de forma transparente al usuario.

Para analizar el rendimiento computacional del modelo NLODT-P, diversos experimentos con varios tamaños del volumen se han llevado a cabo. Estos experimentos ejecutan NLODT-P para la localización de 4 partículas de  $0.5\mu$  de diámetro con volúmenes que van desde  $200^3$  a  $280^3$  vóxeles y tres hologramas de entrada. Para la evaluación, una CPU (2x4 Intel Xeon E5620 cores, 48 GB RAM, 2.4 GHz velocidad de reloj y bajo Linux) y una GPU NVIDIA Tesla M2090 han sido consideradas. La GPU NVIDIA ha sido utilizada para acelerar el procedimiento *Forward*.

La Fig. 7 muestra el tiempo de ejecución (en segundos) de la ejecución del modelo NLODT-P usando dos alternativas: solo MATLAB y una versión CUDA-MATLAB. La implementación con GPU de NLODT-P ha demostrado reducir el tiempo de ejecución en un factor de alrededor de  $37\times$  con respecto a la estrategia que usa solamente MATLAB. Por lo tanto, el uso de MATLAB y la computación hace posible la aplicabilidad del modelo.

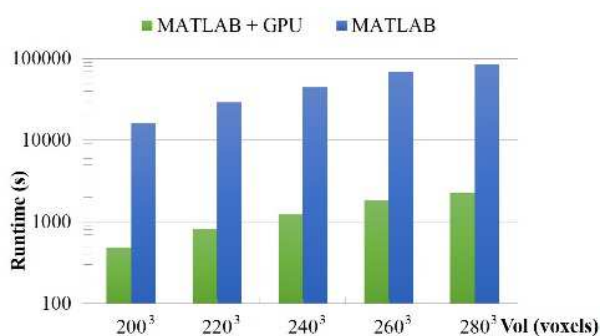


Fig. 1. Evolución del tiempo de ejecución del modelo NLODT-P para distintos tamaños de Vol.

#### CONCLUSIONES Y TRABAJOS FUTUROS

En este trabajo se ha descrito un modelo 3D no lineal de ODT para localizar partículas, como parte de una técnica de velocimetría de fluidos. Algunos experimentos numéricos con dos configuraciones han sido considerados para demostrar las ventajas del modelo NLODT-P frente a la LODT. Se ha mostrado que la dispersión múltiple debida al tamaño finito de las partículas puede manejarse cuando se utiliza un matched-filter adecuado, pero la dispersión múltiple

entre partículas tiene que ser manejado con un método de optimización iterativo, como el NLODT-P propuesto.

Nosotros hemos implementado un modelo NLODT-P utilizando la combinación de MATLAB y de una GPU a través de los ficheros MEX-files, que permiten que rutinas en CUDA puedan ser llamadas desde la interfaz MATLAB. Los factores de aceleración conseguidos llegan hasta  $37\times$  con respecto a la versión que solo utiliza MATLAB. Adicionalmente, el uso de un formato específico para almacenar la gran matriz dispersa ( $M$ ) involucrada en el sistema, ha reducido los requerimientos de memoria en un factor de 10 con respecto al formato tradicional que utiliza MATLAB para compactar este tipo de matrices.

Como resultado de este trabajo, se puede afirmar que el modelo NLODT-P es una buena aproximación para mejorar la precisión los métodos de ODT lineales a la hora de localizar partículas en las aplicaciones de velocimetría de fluidos. La computación de altas prestaciones ha sido esencial para desarrollar y aplicar esta técnica. Las líneas de trabajo futuras están relacionadas con la necesidad de un modelo multiescala para incluir este tipo de propuestas en entornos de producción.

#### AGRADECIMIENTOS

G. Ortega es una becaria del programa FPU. Este trabajo ha sido financiado por el Ministerio de Ciencia (TIN2008-01117, TIN2012-37483) y la Junta de Andalucía (P10-TIC-6002, P11-TIC7176, P12-TIC-301), parcialmente financiados por el Fondo Europeo de Desarrollo Regional (FEDER).

#### REFERENCIAS

- [1] M.P. Arroyo and K.D. Hinsch. "Recent developments of PIV towards 3D measurements. Particle Image Velocimetry: New Developments and Recent Applications", Ed. A Schroder, CE Willert, (Springer, New York, 2008), pp. 127-54.
- [2] J.M. Coupland and N.A. Halliwell. "Particle image velocimetry: three-dimensional fluid velocity measurements using holographic recording and optical correlation" *Appl. Opt.* 31 1004-8 (1992).
- [3] W.D. Koek, N. Bhattacharya, J.J.M. Braat, T.A. Ooms and J. Westerweel, . Ooms and J. Westerweel, "Influence of virtual images on the signal-to-noise ratio in digital in-line particle holography" *Opt. Express* 13, 2578-2589. (2005)
- [4] G. E. Elsinga, F. Scarano, B. Wienieke and B.W. van Oudheusden, "Tomographic particle image velocimetry" *Exp. Fluids.*, 41:933-947 (2006).
- [5] Sheng J., Malkiel E. and Katz J., "Single beam two-views holographic particle image velocimetry" *App. Opt.* 42(2), 235-250 (2003).
- [6] J. Lobera and J. M. Coupland. "Optical diffraction tomography in fluid velocimetry: the use of a priori information" *Meas. Sci. and Technol.* 19(7):074013 (2008).
- [7] Wolf E., "Three-dimensional structure determination of semi-transparent objects from holographic data" *Optics Communications*, 1(4), 153-156 (1969).
- [8] K. Belkebir and A. Sentenac. "High resolution optical diffraction microscopy" *J. Opt. Soc. Am. A* 20(7), 1223-1229 (2003).
- [9] P. C. Chaumet, A. Sentenac, K. Belkebir, G. Maire, and H. Giovannini. "Improving the resolution of grating-assisted optical diffraction tomography using a priori information in the reconstruction procedure". *Journal of Modern Optics*, 57(9):798-808 (2010).
- [10] F. Soulez, L. Denis, C. Fournier, É. Thiébaud and C. Goepfert "Inverse-problem approach for particle digital holography: accurate location based on local optimization"- *JOSA A* 24(4),1164-1171 (2007)

- [11] China's TianhE-2 Supercomputer Maintains Top Spot on 42nd TOP500 List. November 2013. <http://s.top500.org/static/lists/2013/11/PressRelease201311.pdf>.
- [12] G. Ortega, J. Lobera, M. P. Arroyo, I. García, and E. M. Garzón. "High performance computing for optical diffraction tomography". In HPCS, pages 195–201. IEEE, ISBN: 978-1-4673-2361-1. (2012).
- [13] G. Ortega, J. Lobera, I. García, M. P. Arroyo and E. M. Garzón. "Parallel resolution of the 3D Helmholtz equation based on multi-GPU clusters". *Concurrency-Pract Ex*. doi:10.1002/cpe.3212. (2014)
- [14] Mathworks. MATLAB with Mex Files. <http://www.mathworks.es/es/help/distcomp/run-mex-functions-containing-cuda-code.html>
- [15] J. M. Coupland and J. Lobera "Holography, tomography and 3D microscopy as linear filtering. *Measurement Science and Technology*", **19**(7):074012, (2008).
- [16] M. N. O. Sadiku. "Numerical Techniques in Electromagnetics", 2nd Edition, (CRC Press, 2001).
- [17] F. Ihlenburg and I. Babuska, "Finite element solution of the Helmholtz equation with high wave number Part I: The h-version of the FEM". *Computers & Mathematics with Applications* 30(9):9 – 37 (1995)
- [18] I.M. Babuska and S.A. Sauter. "Is the pollution effect of the FEM Avoidable for the Helmholtz Equation considering high wave numbers?" *SIAM Rev.* **42**(3):451–484 (2000)
- [19] Y. Saad. "Iterative Methods for Sparse Linear Systems", 2nd Edition. SIAM (2003).
- [20] G. Ortega, E. M. Garzón, F. Vázquez, and I. García. "The BiConjugate gradient method on GPUs". *The Journal of Supercomputing*, 64:49–58, (2013).
- [21] G. Ortega, E. M. Garzón, F. Vázquez, and I. García. "Exploiting the regularity of differential operators to accelerate solutions of PDEs on GPUs". In *Proceedings of the 2011 International CMMSE*, 908–917 (2011).
- [22] NVIDIA. CUDA Toolkit. <https://developer.nvidia.com/cuda-toolkit>