



Universitat d'Alacant
Universidad de Alicante

Estudio e implementación
de técnicas subpíxel para
medidas no invasivas

María Baralida Tomás López



Tesis **Doctorales**

UNIVERSIDAD de ALICANTE

Unitat de Digitalització UA

Unidad de Digitalización UA



Universitat d'Alacant
Universidad de Alicante

Instituto Universitario de Física Aplicada a las Ciencias y las Tecnologías

**ESTUDIO E IMPLEMENTACIÓN DE
TÉCNICAS SUBPÍXEL PARA MEDIDAS NO
INVASIVAS**

MARÍA BARALIDA TOMÁS LÓPEZ

Tesis presentada para aspirar al grado de
DOCTORA POR LA UNIVERSIDAD DE ALICANTE

Programa de Doctorado en Física Aplicada a las Ciencias y las Tecnologías

Dirigida por:

Dr. DAVID MAS CANDELA

Dra. BELÉN FERRER CRESPO

Cofinanciada por:

La Generalitat Valenciana y el Fondo Social Europeo a través de la subvención para la contratación de personal investigador de carácter predoctoral, ACIF/2018/211

*“Una investigación básica de calidad
es fundamental para un posterior desarrollo,
porque de ella saldrán resultados
no previsibles a priori.”*

Margarita Salas Falgueras (1938 – 2019)

Universitat d'Alacant
Universidad de Alicante

AGRADECIMIENTOS

Son muchas las personas e instituciones que me han apoyado durante el proceso de investigación que finaliza con la presentación de la actual tesis doctoral y a las que me gustaría expresar mi más sincero agradecimiento.

En primer lugar, a mi marido Paco por animarme y alentarme desde el primer día durante estos años, tanto en los momentos buenos como en los más delicados.

También a mi familia por estar a mi lado y contribuir a llegar hasta aquí: a mi hermano Francisco Gaspar que siempre me ha incitado a seguir estudiando y me ha enseñado a luchar por mis sueños; a mis padres Gaspar y Baralida que desde pequeña me han inculcado la importancia de la disciplina y la constancia en los estudios; y a mis abuelos Paco y Maruja por creer en mí en todo momento.

Igualmente, quiero dar las gracias a mis directores de tesis David y Belén por ser mis mentores durante estos años, animarme y ayudarme en cada momento, enseñarme a nivel científico y personal y por ser tan cercanos desde el primer instante.

Del mismo modo, quiero agradecer al Grupo de Análisis de Imagen, Sistemas Ópticos y Visión (IMAOS+V) por acogerme y poner a mi disposición los laboratorios y materiales necesarios para realizar los experimentos y trabajos para mi tesis, así como al Instituto Universitario de Física Aplicada a las Ciencias y las Tecnologías por aceptarme en su programa de doctorado.

También, al Departamento de Óptica, Farmacología y Anatomía por permitirme realizar colaboraciones docentes dentro de varias asignaturas de su Grado en Óptica y Optometría. Además, no puedo olvidarme de nombrar a Menchu, Begoña, Joaquín y Chelo por acogerme en las prácticas de sus asignaturas y enseñarme a ser docente.

Finalmente, deseo mostrar mi gratitud a la Generalitat Valenciana y el Fondo Social Europeo por la cofinanciación de este trabajo mediante la subvención para la contratación de personal investigador de carácter predoctoral, ACIF/2018/211.

A todos, MUCHAS GRACIAS.



Universitat d'Alacant
Universidad de Alicante

RESUMEN

Se conocen como técnicas de resolución subpíxel aquellas que son capaces de detectar objetos o movimientos capturados por una cámara digital con una resolución superior a la definida por un único sensor. Dado que estas técnicas son capaces de detectar información por encima de la resolución límite teórica del sistema, se engloban dentro de los llamados métodos superresolventes (“Superresolución,” 2019).

Las técnicas subpíxel para medidas no invasivas se utilizan desde que se empezó a considerar insuficiente la resolución de la imagen digital para ciertas aplicaciones. Esto ocurre porque en una imagen digital, a priori, no es posible distinguir detalles entre dos puntos del objeto de interés si ambos coinciden dentro del mismo píxel en la imagen, ya que cada píxel proporciona una información única y uniforme para toda su superficie.

Originalmente, se comenzaron a utilizar técnicas de correlación para obtener una resolución ampliada (Pearson et al., 1977) y seguidamente, empezaron a aparecer trabajos que trataban de aprovechar las limitadas posibilidades de los sistemas digitales de imagen. Al aumentar el desarrollo tecnológico, empieza a ser más fácil la implementación de métodos de detección de objetos. La mayoría de estudios que utilizan estas técnicas intentan detectar dianas específicas con formas conocidas que se utilizarán como referencia para centrar y guiar sistemas de detección y posicionamiento de objetos con precisión subpíxel.

La tecnología de la imagen digital se ha expandido y mejorado enormemente en las últimas décadas, perfeccionando tanto la capacidad de los sistemas de captación y procesado como la calidad de las imágenes obtenidas. Actualmente, los sensores digitales de imágenes están muy extendidos y sus resoluciones espaciales igualan e incluso mejoran la resolución esperada con las películas analógicas tradicionales, aunque, a pesar de estas mejoras, las

técnicas subpíxel siguen siendo necesarias ya que muchas limitaciones prácticas permanecen.

La localización y detección de movimientos de objetos con precisión subpíxel será de utilidad cuando la amplitud de dichos movimientos en el plano de la imagen sea de una fracción de píxel, por lo que no se podría detectar a priori. No obstante, en cuanto ese movimiento produzca un cambio en la respuesta del sensor de la cámara, esa variación será detectable y el desplazamiento subpíxel quedará registrado. Si este proceso se repite en el mismo u otro píxel, se registrarán movimientos sucesivos y podrá determinarse, mediante el análisis matemático pertinente, la trayectoria del objeto observado.

Bajo esta premisa, se pueden encontrar diferentes métodos para el seguimiento de objetos en una imagen. Si en la escena se incluye una diana que tenga una forma matemáticamente descriptible, se podrían utilizar las restricciones geométricas que definen el objeto para detectar de manera precisa su posición o desplazamiento (Espinosa et al., 2015). Por ejemplo, un segmento recto mantiene todos los puntos alineados, de modo que una ligera desviación en la posición de los píxeles que registran el objeto podría interpretarse como un desplazamiento o un giro, aunque no haya un desplazamiento claro de un píxel a otro. Para efectuar esta detección podrían utilizarse técnicas de reconocimiento de formas o técnicas estadísticas como el cálculo del centro de masas o del centroide.

Desafortunadamente, no siempre es posible reconocer el objeto de interés porque puede no estar bien contrastado o aislado del fondo o bien, porque no tiene una forma matemáticamente descriptible. En estos casos, se puede utilizar un modelo del objeto obtenido, por ejemplo, a partir de una imagen de la secuencia grabada para detectar el movimiento y localizarlo luego mediante la correlación cruzada mejorada con un ajuste del pico para poder describir el movimiento a nivel subpíxel. También pueden ser útiles las técnicas basadas en la detección local de vibraciones, donde no será necesario conocer la forma del objeto, sino detectar zonas o incluso píxeles aislados que muestren una oscilación periódica en la luminancia (Ferrer et al., 2013).

La hipótesis principal de este trabajo consiste en que, mediante la detección de pequeños cambios de intensidad luminosa en el sensor de la cámara, se pueden identificar y seguir los movimientos producidos en el objeto de interés incluso con errores inferiores al tamaño del píxel en la imagen.

La finalidad es estudiar teórica y experimentalmente los métodos de detección y seguimiento de objetos con precisión subpíxel sin identificación previa del objeto. Para ello, los dos objetivos principales son estudiar y optimizar los métodos de seguimiento con precisión subpíxel e implementarlos después experimentalmente.

Para desarrollar estos objetivos, es imprescindible realizar una búsqueda bibliográfica exhaustiva de las técnicas subpíxel desarrolladas y empleadas por otros grupos de investigación. Se ha visto con esta búsqueda que el método principal de medida de desplazamientos en imágenes es la correlación cruzada entre una imagen de referencia y su versión desplazada, aunque para conseguir una resolución subpíxel es necesario introducir un ajuste en el entorno del pico de correlación para recalcular la nueva posición del máximo. Así, se explicará de modo pormenorizado esta técnica resaltando sus ventajas e

inconvenientes, así como las diferentes técnicas que se utilizan para superar dichos problemas (Tomás et al., 2020).

La búsqueda bibliográfica ha revelado que algunos autores aseguran que desenfocar ligeramente el objeto de interés puede mejorar la detección con precisión subpíxel de la correlación cruzada (Michaelis et al., 2016; Overmars et al., 2010; Zhou et al., 2015), aunque no especifican el nivel de desenfoco. Por ello, se ha llevado a cabo un estudio teórico para poder proporcionar unos valores de desenfoco que mejoren dicha detección. Para ello, se han generado 5 tests binarios con un desplazamiento dado y se han evaluado sus errores al realizar la correlación cruzada utilizando tres tipos de interpolaciones del pico de correlación con diferentes tamaños de la matriz alrededor del pico y con 6 niveles de desenfoco distintos (contando también sin desenfocar) introducido mediante un filtro de desenfoco gaussiano. En segundo lugar y para obtener unos resultados directamente comparables con los que se presentan en la bibliografía, se ha estudiado el método utilizando 5 secuencias de imágenes con speckle procedentes de un banco de imágenes público (“Previous DIC Challenge 1.0 Data,” 2020). Estas secuencias suelen ser empleadas para validar los métodos de seguimiento con precisión subpíxel ya que en el banco de imágenes se proporcionan los datos de desplazamiento, ruido y contraste de las imágenes.

El estudio plantea la correlación con un ajuste analítico del área local alrededor del pico de correlación. Se ha analizado tanto el efecto del tamaño de la zona de análisis como el método de ajuste del entorno a una función continua para relocalizar el pico de correlación. Para ello, se han implementado los tres métodos de ajuste más utilizados en la bibliografía (Michaelis et al., 2016; Roesgen, 2003; Nobach et al., 2005) como son el ajuste del pico de correlación por splines, el ajuste a una función polinómica y el ajuste a una función gaussiana.

Como resultado de este estudio (Tomás et al., 2020), se ha obtenido que, según el tipo de ajuste utilizado, puede ser mejor que el objeto de interés esté ligeramente desenfocado o no. Además, para cada ajuste, la matriz de píxeles alrededor del pico de correlación utilizada para el interpolado es óptima con un tamaño distinto. En el caso del ajuste con splines, se han obtenido mejores resultados con vecindarios en torno al pico de correlación de gran tamaño y con un nivel alto de desenfoco. Las características particulares de los splines hacen que este tipo de ajuste se adapte mucho a la forma del pico, por lo que sus resultados son muy variables en función del objeto que se tenga. En cuanto al ajuste polinómico, los mejores resultados ocurren con tamaños de matriz de ajuste pequeños y desenfocos altos. Por último, en el ajuste gaussiano se han obtenido buenos resultados (los mejores de todo el estudio) con tamaños de matriz de ajuste grandes y sin desenfocar, aunque para este ajuste, la selección de los diferentes parámetros es mucho más crítica que en las otras dos interpolaciones, produciendo resultados muy malos si los parámetros no son los óptimos. Se debe tener en cuenta también que los objetos con speckle utilizados en este estudio no son siempre los que se encontrarán en los experimentos reales y que, con otro tipo de objetos, estos resultados podrían verse ligeramente modificados.

Siguiendo con el método de la correlación cruzada junto con un ajuste del pico de correlación, se ha llevado a cabo un experimento práctico para medir la expansión de una roca arenisca al hincharse por hidratación estando parcialmente sumergida en agua (Ferrer

et al., 2021). Este trabajo tiene gran interés en el campo de la restauración de edificios históricos para evaluar el estado de deterioro de las construcciones (Di Benedetto et al., 2015). Además, el método estándar para medir esta expansión por hidratación en muestras de roca consiste en la utilización de un transformador diferencial variable lineal (LVDT por sus siglas en inglés) que, colocado en la parte superior de la muestra, permite medir movimientos verticales, pero para medir también los horizontales se debe realizar un montaje bastante sofisticado. Realizar esta medida mediante imágenes permite reducir costes económicos y complejidad de montaje experimental, además de proporcionar resultados de desplazamientos tanto horizontales como verticales con solo un montaje sencillo.

En este estudio experimental se ha obtenido que los resultados son parecidos a aquellos medidos con el LVDT, por lo que se está introduciendo una mejora en el proceso de medida al ser un experimento más sencillo que proporciona unos resultados similares. Además, si se quisiera realizar este mismo experimento fuera del laboratorio sobre la roca utilizada como material de construcción de una edificación, sería factible realizar esta medida y analizarla posteriormente grabando solo la zona de interés de la roca. Sin embargo, realizar esta medida con el palpador, implicaría un montaje mucho más sofisticado incluyendo la extracción de una muestra, lo que puede ser problemático en un edificio histórico.

Por otra parte, se ha realizado otro estudio experimental para determinar si la textura propia del hormigón puede ser utilizada como diana natural utilizando el método de la correlación cruzada junto con un ajuste del pico de correlación. Para ello, se han evaluado un total de 21 probetas de hormigón generadas con distintas técnicas de encofrado, compactación y consistencia del hormigón y se han grabado distintas secuencias estáticas de cada probeta de hormigón variando la iluminación y el desenfoque en condiciones de laboratorio. El objetivo final de este estudio sería analizar la viabilidad de utilizar la correlación cruzada para analizar los movimientos de un objeto con textura de hormigón en condiciones reales, esto es, al aire libre y a gran distancia del objeto. No obstante, este objetivo final tiene muchas incertidumbres ya que al aire libre no se pueden controlar muchas variables del ensayo como la iluminación, el viento sobre el trípode de la cámara, el ruido o la distorsión atmosférica, entre otras. Por ello, se ha evaluado en primer lugar la viabilidad de la medida con la textura de hormigón en condiciones controladas de laboratorio y, tras obtener un resultado positivo, ya es posible realizar medidas más complejas y con más variables en juego.

Finalmente, se ha encontrado en la bibliografía reciente una nueva técnica para medidas de desplazamientos con precisión subpíxel basada en el análisis de la diferencia en la luminancia entre dos imágenes (Wan et al., 2020). Este análisis de la luminancia se realiza calculando la diferencia de nivel de gris entre las imágenes en horizontal y en vertical para poder determinar el desplazamiento en ambas direcciones. La ventaja principal de este método es su rapidez en el cálculo, aunque presenta los inconvenientes de que es muy sensible al ruido y a cambios de iluminación en la escena y que necesita una calibración previa.

Una medida experimental previa para utilizarla como calibración no siempre será posible, ya que las medidas pueden consistir en tests destructivos o procesos irreversibles. Así pues, nos hemos planteado realizar una precalibración digital mediante secuencias de

movimiento subpíxel generadas por ordenador. Para ello, hemos implementado dos métodos diferentes, uno basado en propiedades de la Transformada de Fourier y otro con interpolación de rejilla. Con este análisis, se demuestra que el método de rejilla que utiliza el algoritmo *griddedInterpolant* de Matlab es efectivo para realizar un desplazamiento digital del objeto que sirva como calibración.

De este modo, se ha procedido a comparar de manera numérica y experimental el nuevo método con el de la correlación anteriormente presentado. En la prueba teórica se ha obtenido que, en general, el método basado en la diferencia de imágenes proporciona mejores resultados que el basado en la correlación. No obstante, las diferencias son pequeñas y podrían ser irrelevantes a nivel experimental. Para las pruebas de laboratorio se utilizaron los mismos objetos empleados en la evaluación de la precalibración digital, pero esta vez desplazados con un motor micrométrico y controlando la iluminación para que el contraste entre el objeto y el fondo sea máximo y obtener, de esta forma, unos resultados comparables a los obtenidos utilizando la correlación cruzada.

Los resultados mostrados con este nuevo método son que no presenta el error de atrapamiento de pico y que, en general, es más robusto con imágenes reales respecto del método de la correlación, aunque debe controlarse muy bien el contraste de las imágenes grabadas. Por otra parte, aunque este nuevo método proporcione resultados similares o incluso mejores a los de la correlación cruzada, presenta los inconvenientes que ya se han comentado, como son: insensibilidad a la dirección del movimiento, sensibilidad al ruido y a cambios de iluminación y necesidad de una calibración previa para poder conocer la cantidad de movimiento que se está produciendo en el objeto.

Para finalizar, y tras realizar experimentos con los distintos métodos subpíxel para evaluar su funcionamiento, llegamos a la conclusión de que las técnicas subpíxel estudiadas en este trabajo de tesis son adecuadas para medir movimientos con una resolución por encima de la resolución nominal determinada por el sistema de captación de imágenes utilizado. No obstante, es importante continuar investigando en la aplicación práctica de estas técnicas para poder utilizarlas en entornos menos controlados en los que las condiciones de iluminación o contraste no puedan mantenerse constantes durante todo el experimento.

ÍNDICE

| | |
|--|-----|
| AGRADECIMIENTOS | III |
| RESUMEN | V |
| ÍNDICE | XI |
| CAPÍTULO 1. Introducción..... | 1 |
| 1.1. Justificación..... | 1 |
| 1.2. Objetivo de la tesis | 5 |
| 1.3. Organización de la tesis..... | 6 |
| CAPÍTULO 2: Revisión bibliográfica de los métodos subpíxel | 9 |
| 2.1. Ajuste morfológico | 10 |
| 2.2. Centro de masas y centroide..... | 12 |
| 2.3. Correlación cruzada con ajuste del pico de correlación | 14 |
| 2.4. Análisis de luminancia | 19 |
| CAPÍTULO 3: Evaluación teórica del método subpíxel por ajuste del pico de correlación | 23 |

| | |
|--|------------|
| CAPÍTULO 4: Aplicaciones del método subpíxel por ajuste del pico de correlación | 39 |
| 4.1. Medida de la expansión hídrica de una roca arenisca | 40 |
| 4.2. Evaluación de la textura natural del hormigón como diana en medidas a través de la correlación cruzada | 52 |
| CAPÍTULO 5: Métodos subpíxel por análisis de luminancia | 63 |
| 5.1. Evaluación de métodos de generación de secuencias de calibración digitales | 65 |
| 5.2. Comparación numérica de los métodos VarCGL y correlación cruzada | 72 |
| 5.3. Comprobación del método VarCGL utilizando secuencias reales | 75 |
| CAPÍTULO 6: Conclusiones | 81 |
| CAPÍTULO 7: Futuras líneas de investigación | 83 |
| BIBLIOGRAFÍA | 85 |
| ANEXO I: Cámaras y objetivos | 91 |
| AI.1. Resoluciones espacial y temporal | 91 |
| AI.2. Características de las cámaras | 93 |
| AI.3. Características de los objetivos | 95 |
| ANEXO II: Características y uso de las cámaras | 99 |
| ANEXO III: Sistemas de iluminación y cabinas de iluminación | 103 |
| ANEXO IV: Publicaciones | 107 |
| PUBLICACIÓN 1: Influence of neighborhood size and cross-correlation peak-fitting method on location accuracy | 109 |
| PUBLICACIÓN 2: Use of image correlation to measure macroscopic strains by hygric swelling in sandstone rocks | 129 |

CAPÍTULO 1. Introducción

1.1. Justificación

La necesidad de utilizar técnicas subpíxel para medidas no invasivas no es nueva y surge desde el momento en que la resolución de la imagen digital se considera insuficiente para algunas aplicaciones. En primer lugar, conviene definir qué es un “píxel”, ya que es un concepto clave en esta tesis. Se tiende a pensar que los píxeles son las superficies pequeñas y homogéneas que recogen la información y forman la imagen digital. Sin embargo, los píxeles pueden ser definidos de diferentes formas según el contexto (“Pixel,” 2021). Por ejemplo, pueden ser los píxeles impresos en una página (medidos como puntos por pulgada, ppi), los píxeles transportados por señales eléctricas, los píxeles presentes en una pantalla de visualización de datos o los píxeles como elementos fotosensibles en una cámara digital, entre otros. En el caso de la presente tesis doctoral, nos referiremos a los píxeles de la cámara digital como el elemento fotosensible menor que es capaz de recoger información sobre la intensidad de la luz en el sensor de la misma, de forma que cada píxel tendrá una intensidad de luz constante en toda su superficie.

En una imagen digital, por tanto, no es posible conocer la diferencia de información entre un punto y otro de la imagen si ambos puntos se encuentran situados dentro del mismo píxel. Esto es así porque cada píxel se puede considerar homogéneo en toda su extensión, por lo que contiene una información uniforme para toda su superficie. En la figura 1.1 se puede ver una imagen ampliada en la que se observan unos cuadrados con diferencias de color entre unos y otros. Estos cuadrados representan los píxeles del sensor de la cámara digital con una información constante en toda su superficie (la intensidad de luz en este

caso). Si en una zona del objeto real hay una diferencia de información, como podría ser un borde o un cambio en la textura, en dos posiciones contiguas que en el sensor de la cámara estén representadas dentro del mismo píxel, esa diferencia de información no será captada en la imagen digital, por lo que la distancia entre ambas posiciones no se podrá medir si no es utilizando técnicas subpíxel.

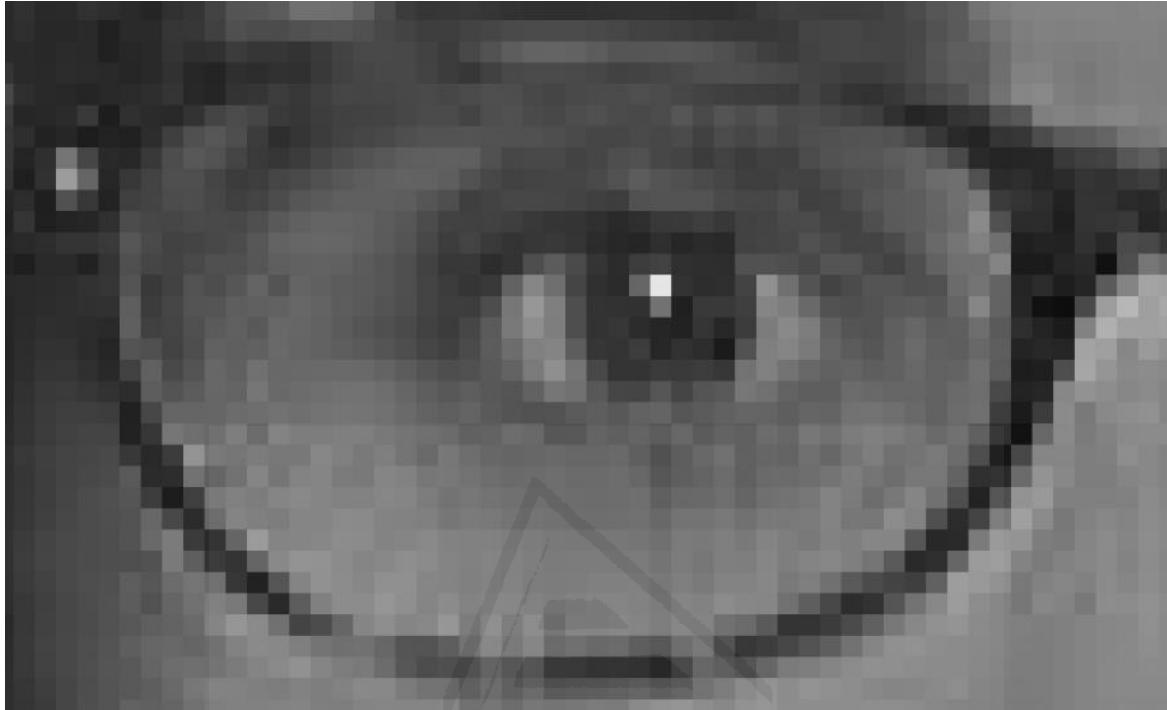


Figura 1.1. Imagen pixelada donde se puede observar que cada píxel del sensor contiene una información diferente (la misma intensidad de luz en toda la superficie de cada píxel).

En el Anexo I.1 se explican más detalladamente las resoluciones espaciales y temporales de las cámaras junto con ejemplos gráficos para su mejor comprensión. No obstante, creemos conveniente realizar aquí una breve introducción sobre estos dos conceptos para la mejor comprensión del resto de los antecedentes.

La resolución espacial de una cámara digital está muy relacionada con el concepto de píxel ya que se refiere a la densidad de píxeles que hay en el sensor de la cámara (Yang and Huang, 2017). De esta manera, una disminución en el tamaño del píxel proporcionará una información más detallada del objeto de interés ya que los píxeles serán menores y más cercanos unos a otros y, en general, los detalles de la imagen estarán mejor muestreados. Una mayor densidad de píxeles mitigará el problema de la resolución, pero nunca podrá ser eliminado completamente ya que existe un límite tecnológico por el cual no se podrá reducir el tamaño de los píxeles infinitamente. Por otro lado, la relación coste-beneficio de utilizar sistemas cercanos al límite tecnológico, en ocasiones no es conveniente y, en esos casos, suelen preferirse sistemas no óptimos más ajustados en coste y utilizar técnicas subpíxel para el análisis posterior de las imágenes, paliando así los defectos en la imagen capturada.

Otro parámetro ligado a la calidad de imagen proporcionada por una cámara es la resolución temporal. Se trata de la velocidad de captura de la cámara, es decir, el número de fotogramas que es capaz de grabar en un determinado espacio de tiempo. Se mide en fotogramas por segundo [fps] y, a mayor velocidad, más información temporal se tendrá del

objeto. En aquellos casos en los que sea muy importante conocer la posición exacta del objeto en movimiento en cada momento porque pueda tener desplazamientos o vibraciones rápidos, será necesario utilizar una resolución temporal alta para tener la mayor información posible de su posicionamiento en cada instante.

Nuestro sistema visual es capaz de combinar imágenes percibiéndolas como un objeto en movimiento, sin percibir imágenes individuales, a una velocidad mínima de 16 fps (Pueo, 2016). Sin embargo, para poder considerar que una cámara es de alta velocidad, la resolución temporal mínima que deberá tener es de 250 fps, habiendo cámaras de alta resolución científicas con resoluciones temporales de hasta 1000 fps o incluso más (“High-speed camera,” 2021). Las cámaras de alta velocidad, generalmente tienen una memoria interna RAM donde almacenan las imágenes antes de pasarlas al ordenador ya que la velocidad de transferencia de datos de la cámara al ordenador suele ser menor a la de captura de las imágenes, por lo que si no existiera esa memoria interna no se conseguiría grabar a tan altas velocidades (Roche, 2013). Algunas de ellas graban la memoria RAM completa y luego hacen toda la transferencia de datos al ordenador, mientras que otras van traspasando datos al ordenador mientras se realiza la medida, de forma que no utilizan la memoria interna completamente, sino que sirve de almacenamiento intermedio para poder capturar más imágenes.

Las resoluciones espacial y temporal están inversamente relacionadas ya que el parámetro crítico de una cámara es la cantidad de información que puede procesar el sistema. La información que capturan los sensores es digitalizada, redirigida a los procesadores y almacenada en la memoria. Estas operaciones tardarán un tiempo en función del hardware y el software de la cámara y/o el ordenador. Así, una imagen de alta densidad generará una gran cantidad de datos, que el sistema deberá almacenar antes de proceder con el siguiente fotograma. Por tanto, si se aumenta la resolución espacial será a costa de reducir la resolución temporal y viceversa. Esto provoca que, en el caso en el que una alta resolución temporal sea prioritaria, la resolución espacial se verá perjudicada y, si fuera necesario aumentarla, deberán utilizarse técnicas subpíxel.

Las técnicas subpíxel fueron introducidas en 1977 (Pearson et al., 1977) a través del análisis de la diferencia de fase entre imágenes mediante la Transformada de Fourier. Durante los siguientes años fueron apareciendo diferentes trabajos tratando de aprovechar las limitadas posibilidades de los sistemas digitales de captación de imagen. En los años 90, en un entorno tecnológico más desarrollado, empiezan a perfilarse métodos estandarizados para la detección de objetos determinados y propuestas de optimización de su seguimiento. En (Shortis et al., 1994, 1995) se puede consultar un resumen de las técnicas más adecuadas de entre las estudiadas en las décadas anteriores, así como la precisión esperada de cada una de ellas. La mayoría de estos estudios tratan sobre la detección de dianas específicas, llamadas objetos fiduciaros, que se utilizarán como referencia para el centrado y guiado de sistemas de detección y posicionamiento de objetos con precisión subpíxel.

Desde los años 90, la tecnología de imagen digital ha tenido una expansión importante, con una mejora evidente tanto en la capacidad de los sistemas de captación y procesado como en su calidad de imagen. En la actualidad, los sensores de imagen están ampliamente extendidos y cuentan con resoluciones espaciales que igualan e incluso superan la calidad

esperada con películas analógicas tradicionales. Sin embargo, pese a esas mejoras, las técnicas subpíxel siguen siendo necesarias por varios motivos. El primero de ellos es la limitación física de la resolución espacial en la formación de imágenes que puede ser debida a la difracción, aberraciones, difusión en los métodos de propagación y transmisión, etc. El segundo motivo es la existencia de un límite tecnológico que afecta a la cantidad de información que puede procesar un sistema. Finalmente, existen limitaciones prácticas, que si bien, son superables con mejores sistemas, no siempre es posible superarlas y en ocasiones es preferible utilizar soluciones no óptimas más ajustadas en coste con el posterior uso de técnicas subpíxel durante el análisis de las imágenes.

Sea cual sea la causa por la que se utilizan estos sistemas, los métodos de procesado que permiten aumentar la resolución espacial de la imagen capturada siguen siendo de gran interés para la comunidad científica. Además, la disponibilidad de cámaras con velocidad de captura elevada a un coste asequible junto con ordenadores con capacidad suficiente para el manejo de gran cantidad de datos (imágenes en este caso), muestra que la tecnología es suficientemente madura como para pasar a un plano aplicado, lo que hace que los estudios realizados durante el desarrollo de la tesis sean novedosos y necesarios.

Como se ha mencionado antes, si un objeto se mueve de modo que su desplazamiento en el plano imagen es de una fracción de píxel, dicho movimiento no será detectable a priori. Sin embargo, en el momento en que ese cambio produzca una alteración en la respuesta del sensor de la cámara, esa alteración será detectable y, por tanto, el desplazamiento subpíxel quedará registrado. La repetición de este proceso sobre el mismo u otro píxel permitirá registrar movimientos sucesivos y determinar la trayectoria del objeto medido con una precisión subpíxel sin necesidad de que se produzcan cambios en todos los píxeles del sensor simultáneamente.

Existen ciertas formas de las dianas que facilitan el seguimiento de objetos con una resolución por encima de la resolución nominal de la cámara facilitando lo que se conoce como superresolución (O’Gorman et al., 1990). Para poder detectar estos pequeños movimientos de dichos objetos, es necesario, en primer lugar, reconocer el objeto y después, poder localizarlo en un punto del espacio. Para ello, se utilizaron algoritmos de localización estadísticos como el cálculo de la posición del centroide, el reconocimiento y ajuste del contorno del objeto o el ajuste del perfil de iluminación (Shortis et al., 1994). Desafortunadamente, estas técnicas requieren de objetos bien contrastados y aislados del fondo, lo que no siempre será posible.

Si el movimiento esperado del objeto es una vibración, existe otra técnica para la monitorización de movimientos subpíxel que se basa en la detección local de vibraciones (Espinosa et al., 2014; Mas et al., 2016a). En este caso, no es necesario conocer la forma del objeto, sino detectar zonas o píxeles aislados que presenten una oscilación periódica de la luminancia. Esta idea resulta de gran utilidad para la medida no invasiva de vibraciones en estructuras u objetos complejos, pero no resuelve el problema del seguimiento continuado del objeto en su trayectoria, ya que el método sólo registra frecuencias de movimiento, pero no permite el cálculo de amplitudes.

En general, el seguimiento preciso de objetos es de gran interés en una gran variedad de aplicaciones científicas (Mecánica, Óptica, Biología, etc.). En esta tesis nos centraremos en

aplicaciones en el campo de la Mecánica y la Ingeniería Civil, en el que la monitorización y control de vibraciones y desplazamientos de estructuras bajo excitaciones dinámicas o impactos es un tema de gran interés. La medición de la respuesta dinámica de una estructura a una excitación ambiental o puntual permite la obtención de importantes parámetros de movimiento, contribuyendo al conocimiento de sus características y permitiendo la identificación de posibles daños (Ferrer et al., 2016).

Además, los sistemas basados en procesado de imagen para la detección de posicionamientos con precisión subpíxel sin utilizar dianas artificiales, permiten medidas de alta resolución a un coste razonable y sin necesidad de andamiajes o instrumentación invasiva de la estructura o espécimen a medir. Hasta hace algunos años, estos sistemas estaban orientados a medidas de deformaciones estáticas (Hack and Leroy, 2005; Wahbeh et al., 2003), sin embargo, en los últimos años, las cámaras de alta velocidad han permitido medidas dinámicas de estructuras en movimiento con resultados precisos y comparables a los obtenidos por otros medios (Feng and Feng, 2017). No obstante, es cierto que su precisión queda lejos de los métodos estándar propios de la disciplina, lo que justifica la búsqueda de técnicas superresolventes capaces de aumentar la resolución de los métodos basados en procesado de imágenes en varios órdenes de magnitud.

1.2. Objetivo de la tesis

La principal hipótesis de partida es que, mediante la detección de pequeños cambios de intensidad luminosa en el sensor de la cámara, se pueden identificar y seguir los movimientos que se están produciendo en el objeto de interés utilizando dispositivos de vídeo estándar o de bajo coste para monitorizar, de forma no invasiva, los movimientos en condiciones adecuadas de distancia y seguridad.

La finalidad de la tesis es estudiar teóricamente e implementar experimentalmente métodos de detección y seguimiento de objetos con resolución subpíxel y sin necesidad de identificación previa del objeto en cuestión. En cuanto al desarrollo teórico de los métodos, se realiza estudiando los fundamentos de la técnica y tratando de avanzar en ella para mejorarla y, con respecto a la implementación experimental, se realizan ensayos en laboratorio para obtener datos y conclusiones experimentales tratando de exportarlos después a medidas reales de parámetros en fenómenos dinámicos en cualquier ámbito como podría ser el de la Ingeniería o la Física.

Para ello se proponen dos objetivos principales:

1. Estudiar y optimizar los métodos de seguimiento, incluyendo un estudio del estado del arte y un análisis crítico de las propuestas que puedan ser interesantes para el estudio de movimientos a distancia.
2. Implementar experimentalmente los métodos subpíxel. Este objetivo contempla la implementación en laboratorio de las técnicas subpíxel estudiadas, así como su adaptación a aplicaciones de interés.

1.3. Organización de la tesis

La presente tesis doctoral se divide en siete capítulos bien diferenciados.

En primer lugar, en el Capítulo 1, el actual, se realiza una introducción general de la obra. Además, se incluye una justificación del trabajo realizado y los objetivos como punto de partida de la investigación.

A continuación, el Capítulo 2 pretende explicar en profundidad las técnicas subpíxel más utilizadas que se conocen, detallando los antecedentes hasta llegar a los conocimientos que tenemos hoy en día sobre las técnicas subpíxel. Estas técnicas se explican en orden ascendente de complejidad, de modo que las primeras técnicas tienden a ser más sencillas y rápidas de implementar aunque pueden no ser adecuadas para analizar objetos muy complejos o para implementarlas en experimentos reales con condiciones de iluminación y ruido inestables que dificulten su uso, mientras que los últimos métodos subpíxel explicados serán más difíciles de implementar, aunque también tienden a ser más robustos, lo que los hace adecuados para situaciones reales más complejas y cambiantes.

Posteriormente, en los siguientes 3 capítulos, el objetivo es mostrar diferentes trabajos, tanto teóricos como experimentales, que se han desarrollado durante la realización de esta tesis doctoral.

En el Capítulo 3 se explica un experimento teórico llevado a cabo para comparar tres métodos de ajuste del pico de correlación y determinar cuál proporciona mejores resultados. Para ello se han utilizado unas imágenes de un banco de imágenes público específicamente diseñado para experimentar con ellas, ya que proporciona secuencias de imágenes con un desplazamiento conocido junto con sus características de ruido y contraste. En cuanto al método de ajuste del pico de correlación, se ha ido variando el tamaño de la matriz utilizada para realizar el ajuste y determinar posteriormente el mejor tamaño en función del tipo de ajuste del pico de correlación utilizado. Además, como se han encontrado referencias que aseguran que un ligero desenfoque del objeto de interés puede mejorar la precisión del seguimiento utilizando la correlación cruzada, se han introducido también distintos niveles de desenfoque para cada uno de los ajustes del pico de correlación y con todos esos datos se proporciona una tendencia del tipo de ajuste junto con su tamaño de la matriz alrededor del pico y el desenfoque que proporciona los mejores resultados para las imágenes utilizadas.

En el Capítulo 4, continuando con el método subpíxel por ajuste del pico de correlación, se desarrollan dos aplicaciones prácticas del método. La primera de las aplicaciones trata de medir la expansión por hidratación de una roca arenisca al estar parcialmente sumergida en agua, cosa que es útil en el campo de la restauración de edificios históricos para evaluar el estado de deterioro de las construcciones. Este experimento se ha llevado a cabo en condiciones de laboratorio, midiendo cuatro muestras de la misma roca sumergida en agua hasta cubrir 1/3 de su altura, permitiendo que la roca sufra una dilatación, debida a las características expansivas de uno de sus componentes, conforme el agua asciende por capilaridad y va hidratando estos componentes. La evaluación de las expansiones relativas de la roca se ha realizado utilizando la correlación cruzada junto con un ajuste del pico de correlación. La segunda de las aplicaciones prácticas del método subpíxel por ajuste del pico

de correlación consiste en evaluar la textura propia del hormigón como diana natural para el seguimiento de estructuras construidas con este material. Para ello, se han evaluado un total de 21 probetas de hormigón fabricadas con distintos tipos de encofrado, compactación y consistencia del hormigón. Además, para cada probeta se han realizado 6 ensayos cambiando tanto el nivel de iluminación como el desenfoque.

En el Capítulo 5, se comenta un nuevo método que ha llegado a nuestro conocimiento en el último año y que pretende medir movimientos con precisión subpíxel utilizando las diferencias de los niveles de gris acumulados por filas y por columnas en todas las imágenes de la secuencia. El interés de este método reside en su rapidez de uso respecto de otros métodos subpíxel ya que los cálculos que realiza son sencillos (solamente sumas y restas). En este caso, se ha llevado a cabo una parte teórica y otra práctica para evaluar el método. Respecto a la parte teórica, se han generado distintos objetos (sencillos y complejos) con desplazamientos digitales y se han evaluado los resultados proporcionados por este método y comparado con los obtenidos utilizando la correlación cruzada normalizada. En cuanto a la parte experimental, se han impreso los mismos objetos utilizados en la evaluación teórica y se han grabado con un desplazamiento generado utilizando un motor de desplazamiento continuo en condiciones controladas de laboratorio. De nuevo, se han comparado los resultados obtenidos con este método y con la correlación cruzada para validar el método en condiciones experimentales controladas.

Después de estos capítulos que componen el grueso de la tesis, en el Capítulo 6 se muestran las conclusiones de la tesis y en el Capítulo 7 se comentan las líneas futuras por las que podría continuar la investigación en este campo tan en auge.

Finalmente, tras el apartado de Bibliografía, que se ha unificado en una única sección para todas las referencias incluidas en la tesis, se muestran 4 anexos que se han considerado importantes para poder ampliar información que puede ser de interés pero que no se ha incluido en los capítulos anteriores para facilitar su lectura.

En el Anexo I se realiza una reseña sobre las cámaras digitales para conocer sus características técnicas y, así, poder ampliar la información que se menciona en el texto general de la tesis en caso de que interese al lector.

En el Anexo II, se describen en profundidad las características técnicas de las cámaras empleadas en las medidas de los diferentes experimentos descritos en los Capítulos 4 y 5, así como el software que las controla.

Relacionado con las pruebas experimentales, en el Anexo III se detallan los sistemas de iluminación disponibles en nuestro laboratorio, así como una cabina de iluminación empleada para experimentos largos en los que las condiciones de iluminación ambiente puedan cambiar.

Por último, en el Anexo IV se muestran los artículos publicados durante la presente tesis doctoral con los que se puede ampliar información de los Capítulos 3 y 4. Para cada artículo publicado, se ha descrito la revista científica en la que se ha publicado y sus características.

CAPÍTULO 2: Revisión bibliográfica de los métodos subpíxel

Puesto que la resolución de una imagen digital puede no ser suficiente para algunas aplicaciones, como cuando existen objetos de apariencia muy pequeña (bien porque lo son o bien por su lejanía) o cuando existen limitaciones técnicas, es necesario introducir técnicas subpíxel para localizar los objetos de interés con mayor precisión. Mediante estas técnicas es posible detectar un movimiento sin que se haya producido un cambio en todos los píxeles de la imagen simultáneamente, sino que en el momento en el que se produzca una alteración en la respuesta de alguno de los píxeles del sensor, ese cambio será detectado y el desplazamiento subpíxel quedará registrado. A continuación, otros cambios en el mismo u otro píxel permitirán registrar los siguientes movimientos estableciendo así la trayectoria del objeto medido.

Basándose en esta premisa se pueden encontrar diferentes métodos para el seguimiento de objetos en una imagen. En el caso de que el objeto de interés tenga una forma regular con una expresión analítica conocida, la aproximación más sencilla sería reconstruir la forma original de la diana ajustando algunos puntos de la forma detectada para obtener así de manera precisa su posición o desplazamiento (Espinosa et al., 2015). Sin embargo, no siempre es posible introducir una diana con una forma regular en la escena y en ocasiones será necesario utilizar como diana aquellas formas naturalmente presentes en la escena como por ejemplo bordes, líneas contrastadas o formas reconocibles en la imagen. Al utilizar dianas naturales es común necesitar realizar un procesamiento de imágenes anterior para aislar el objeto, facilitando así la identificación y su posterior localización y seguimiento.

En el caso de objetos bien aislados y con contornos complejos es posible detectar su desplazamiento en una escena mediante el cálculo de su centroide o de su centro de gravedad.

Esta técnica resulta muy precisa y útil, pero requiere que la diana sea indeformable, esté perfectamente aislada y tenga una iluminación uniforme a lo largo de toda la secuencia. Por ese motivo, es una técnica que se usa en operaciones de calibración con condiciones experimentales controladas y no tanto en experimentos reales.

En el caso de escenas complejas resulta complicado encontrar una expresión matemática que describa el objeto de interés. En este caso, si el objeto es compacto o si contamos con un modelo del objeto a identificar, se puede localizar dicho modelo en la imagen mediante la operación de correlación. El modelo puede proceder de una imagen aparte o de un fragmento de la propia escena. En cualquier caso, la precisión subpíxel vendría dada por ajustes en el objeto y en la escena a comparar o mediante un ajuste por mínimos cuadrados de la propia función de correlación, como se verá más adelante.

Recientemente, en (Wan et al., 2020), se ha presentado un modelo mixto en el que a partir de una escena se modelizan los cambios en la luminancia producidos por un objeto en movimiento durante un ensayo de calibración. Dicho modelo se aplica luego a un experimento real, de modo que el desplazamiento se deduce a partir del modelo de variación de luminancias.

A continuación, se describe de forma detallada cómo funciona cada uno de estos métodos de detección con precisión subpíxel.

2.1. Ajuste morfológico

El ajuste morfológico es una herramienta común cuando el objeto de interés tiene una forma que facilita su localización con superresolución o se introduce en la escena una diana regular cuya forma es conocida matemáticamente y permite realizar un ajuste del objeto capturado a su forma conocida. De esta manera, sólo con la detección de movimiento en unos pocos puntos de la diana es posible reconstruir la forma original y detectar su nueva localización (Espinosa et al., 2015).

El caso más sencillo es el de una línea recta. Su ecuación es conocida por lo que es fácil obtener información de posiciones intermedias de un detalle realizando un ajuste para conocer su posición y/o desplazamiento (Espinosa et al., 2015). Tal y como se muestra en la figura 2.1, si se produce un pequeño desplazamiento en el objeto, parte de él pasará a ser detectado por algún píxel contiguo, aunque el resto de partes de la diana continúen siendo detectadas por el mismo elemento sensor. Sabiendo la forma de la diana es fácil obtener, mediante ajustes estadísticos, el desplazamiento producido.

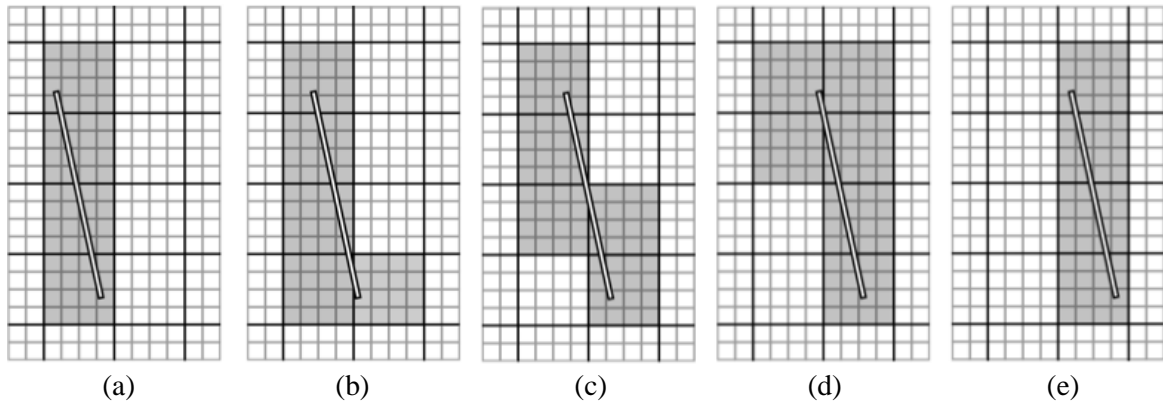


Figura 2.1. Simulación de un movimiento subpíxel en pasos de 0.25 px de una línea recta con una cierta inclinación (Espinosa et al., 2015).

El uso de objetos rectilíneos, aunque sencillo, no es aconsejable. En primer lugar, sólo tienen dos parámetros de ajuste (pendiente y ordenada en el origen), lo que los hace poco versátiles. En segundo lugar, el pixelado introduce escalones y altera su forma, lo que introduce errores en función de su alineamiento con la matriz de sensores. Además, si el objeto está perfectamente alineado con el sensor, esta detección con precisión subpíxel no será posible.

Para paliar este defecto, se pueden utilizar dianas más complejas pero que continúen siendo reconocibles geoméricamente como las elipses que presentan un perfil más suave. Dado que cuentan con más variables de ajuste, los valores obtenidos de los parámetros geoméricos (centro y semiejes) serán más robustos que en el caso de las líneas rectas, llegando a una resolución en la localización de su posición de centésimas de píxel (Ferrer et al., 2016; Mas et al., 2012a, 2012b). Puesto que con la ecuación general de la elipse se pueden calcular sus centros, los ejes largo y corto y su orientación, siempre se podrá conocer la posición y orientación de la elipse, obteniendo así una posible deformación del objeto. Igual que sucedía anteriormente, al producirse un desplazamiento en la diana, algunos puntos del objeto pasarán a ser detectados por el píxel contiguo, mientras que otros puntos continuarán siendo captados por el mismo píxel (figura 2.2). De esta manera, se producirá una deformación en el objeto y se podrá detectar el movimiento con precisión subpíxel siempre y cuando algún píxel comience o deje de detectar el objeto.

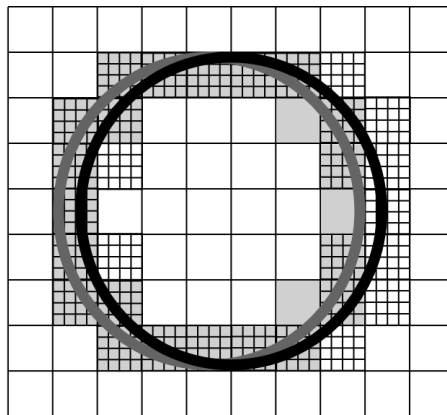


Figura 2.2. Circunferencia utilizada como diana con un movimiento inferior al tamaño de un píxel pero detectable ya que algunos puntos de su borde han cambiado de píxel aunque otros todavía no lo hayan hecho (Mas et al., 2012b).

Aunque la utilización de objetos geoméricamente conocidos con formas regulares y sencillas presenta muchas ventajas, no siempre es posible utilizarlos. En primer lugar, puede no ser posible pegar una diana con la forma óptima en la escena por la dificultad de acceso o por el pequeño tamaño del objeto a medir. En segundo lugar, si no es posible pegar la diana y se necesita utilizar una forma reconocible presente en la escena, puede que ésta no se encuentre en la zona de interés, esté mal iluminada o resulte ocluida parcial o totalmente durante el movimiento.

2.2. Centro de masas y centroide

El centro de masas es un concepto habitualmente utilizado en Mecánica. Físicamente, se trata del punto en el que se puede aplicar la resultante de las fuerzas gravitatorias que experimentan cada una de las partículas del cuerpo, de forma que no cambien los efectos sobre el cuerpo globalmente considerado. Matemáticamente, se define como la posición promedio de todas las partes de un sistema pesados de acuerdo con su masa. En el caso de un cuerpo rígido con densidad uniforme, todas las partes del cuerpo tendrán la misma masa, con lo que el centro de masas coincidirá con el centroide, que es un concepto puramente geométrico.

En procesado de imagen, se ha adaptado el concepto de centro de masas cambiando la masa por la luminancia. Así, este concepto se define como la posición promedio de un objeto o conjunto de objetos en una escena pesado por la intensidad de cada píxel (Goicochea, 2014). Las coordenadas \bar{X} e \bar{Y} del centro de masas de una imagen se definen como:

$$\bar{X} = \frac{\sum X_i \cdot I_i}{\sum I_i} \quad (2.1)$$

$$\bar{Y} = \frac{\sum Y_i \cdot I_i}{\sum I_i} \quad (2.2)$$

donde X_i e Y_i son las coordenadas del píxel considerado e I_i su intensidad.

Se debe destacar que el término de centro de masas se confunde, generalmente, con el de centroide. El centroide se define como la posición promedio del objeto, es decir, es una característica geométrica independiente de la intensidad I_i (en el caso de una imagen). Su expresión se obtiene simplemente cambiando intensidades (I_i) por áreas en las ecuaciones (2.1) y (2.2), considerando que en imágenes el área de un píxel se considera la unidad. No obstante, las referencias que se hacen al centro de masa en diferentes programas de procesado de imagen pueden variar y llevar a confusión. Por ejemplo, en ImageJ (Schneider et al., 2012) se menciona el centro de masas de una imagen como la posición obtenida a través de (2.1) y (2.2), mientras que en Matlab (MATLAB, R2020a) se menciona el mismo

concepto como centroide pesado (weighted centroid). Cabe destacar, además, que en el caso de un objeto binario, el centro de masas corresponderá con el centroide (“Centroide,” 2019), lo que posiblemente sea el origen de la confusión de los términos.

Debe notarse que el concepto de centro de masas sólo tiene sentido en el caso de que el objeto o grupo de objetos que queramos localizar con este parámetro sean claramente distinguibles del fondo y puedan segmentarse sin ambigüedad. La presencia de un fondo de imagen no nulo distorsionará la posición del objeto introduciendo un sesgo importante.

Asimismo, también se debe resaltar que el centro de masas proporciona la posición promedio de la diana, de modo que, si a lo largo de una secuencia la diana se deforma, la coordenada obtenida no coincidirá necesariamente con su trayectoria.

De un modo similar, un cambio de luminancia no homogéneo durante la secuencia debido a sombras o cambio de reflectividad de la diana también afectará a la precisión de la localización del objeto, distorsionando el cálculo de su trayectoria.

Pese a todas estas limitaciones, el centro de masas es una herramienta muy sensible para detectar pequeñas variaciones en la estructura de la diana y medir desplazamientos subpíxel, ya que un pequeño cambio en la luminancia o posición de la diana afectará a la posición promedio. No obstante, todos los condicionantes expuestos hacen que su uso esté restringido a experimentos muy controlados con dianas artificiales o para tareas de calibrado.

Como ejemplo de uso de esta técnica, en (Mas et al., 2016b) los autores usaron una diana gaussiana para, mediante el centro de masas, determinar de modo experimental el mínimo desplazamiento detectable para una resolución y un rango dinámico determinados, obteniéndose una resolución de 0.006 px.



Figura 2.3. Imagen de la diana utilizada para determinar la precisión experimental de los métodos subpíxel. En el recuadro rojo se muestra una ampliación de la diana en negativo (Mas et al., 2016b).

Otro ejemplo de esa técnica es el mostrado en (Ferrer et al., 2016) donde se ha utilizado una diana circular pegada en el centro de un puente peatonal para medir la vibración que se produce en él. La medida se ha realizado con dos cámaras distintas y se ha comparado con los resultados proporcionados por un radar interferométrico que se suele utilizar para medir vibraciones y proporciona los valores de referencia. El error relativo obtenido con las cámaras es menor de un 5%, lo que quiere decir que para el peor caso (para una distancia de 200 m entre el puente y la cámara) se obtiene una desviación estándar del error absoluto de 0.12 mm teniendo un movimiento de amplitud 2.77 mm.



Figura 2.4. Grabación experimental de una diana circular sobre un puente para medir su vibración (Ferrer et al., 2016). A la derecha se muestra la diana ampliada.

2.3. Correlación cruzada con ajuste del pico de correlación

En ocasiones es necesario realizar una aproximación diferente porque se necesita un método más estable frente a cambios de luminancia, ruido o deformaciones de la diana, porque la diana que se quiere identificar es compleja, porque no existe un objeto identificable previamente en la escena o porque no es posible acceder al punto de medida y colocar una diana artificial con las características adecuadas, entre otros. En estos casos, una de las herramientas más utilizadas es la correlación cruzada. Esta técnica se utiliza en aplicaciones variadas y aparentemente dispares como la óptica oftálmica (Ryle et al., 2015), la vibrometría (Ferrer et al., 2011; Mas et al., 2016a), la mecánica de fluidos (Nobach and Honkanen, 2005), la microscopía (Rieger et al., 2015) o la fisiología (Ozana et al., 2015), entre otras disciplinas.

La operación de correlación cruzada surge de la necesidad de obtener una estimación de la similitud entre dos señales o imágenes discretas. Dada una señal de referencia y una señal problema cuya presencia y posición se quiere encontrar dentro de la referencia, desde un punto de vista matemático, la similitud entre ellas se cuantifica mediante el error cuadrático medio o *MSE*. Dado que las señales a comparar pueden estar desplazadas una respecto a la otra, el *MSE* se obtiene para todas las posiciones posibles:

$$MSE_j = \frac{1}{N} \sum_{i=1}^N |f[j] - g[i+j]|^2 \quad (2.3)$$

donde f es la señal de referencia, g la señal problema, N el número total de muestras de la señal problema e i y j son diferentes posiciones en las señales analizadas. Así, la posición más probable de la señal problema en la escena será aquella en la que el MSE sea mínimo, o de forma equivalente, aquella que haga que la correlación cruzada sea máxima:

$$C[j] = \sum_{i=1}^N f^*[j] \cdot g[i+j] \quad (2.4)$$

donde f^* es el complejo conjugado de f .

En el caso de imágenes, tanto el MSE como la correlación pueden extenderse a dos dimensiones de manera inmediata:

$$C[m,n] = \sum_{i,j=1}^N f^*[m,n] \cdot g[m+i,n+j] \quad (2.5)$$

donde m y n son las coordenadas en la imagen resultado de la correlación.

En este caso, a la señal de referencia se le denomina escena o imagen y a la señal problema, patrón o template.

La correlación cruzada se utiliza para determinar si dos imágenes tienen una coincidencia alta y en qué punto ocurre dicha similitud (Lava, 2015). Si se utiliza una secuencia de imágenes, una porción de la primera de ellas es cogida como referencia. Así, la localización en fotogramas posteriores del objeto seleccionado proporcionará su trayectoria y permitirá realizar un seguimiento del objeto. Además, es una herramienta robusta para comparar imágenes alteradas por ruido gaussiano (Pearson et al., 1977) que es bastante habitual en la mayoría de circunstancias durante el procesado de imágenes, ya que este tipo de ruido aparece durante la captura de imágenes por un dispositivo en malas condiciones, una iluminación pobre o altas temperaturas (Pérez Benito, 2015).

En la figura 2.6 se muestra el patrón de correlación obtenido al realizar la correlación cruzada normalizada de las dos imágenes de la figura 2.5, que en realidad contienen el mismo objeto pero desplazado. Se toma la primera de ellas como escena (referencia) y la segunda como patrón a localizar y se consigue una coincidencia perfecta (un valor de 1 en un punto de la imagen en la figura 2.6) ya que el objeto es el mismo en ambas imágenes generadas digitalmente y, por tanto, sin ruido ni deformaciones. No obstante, en un experimento real en el que el objeto de interés puede estar deformándose ligeramente mientras ocurre su movimiento o vibración y las imágenes presentan algo de ruido al grabarlas con la cámara, puede que la coincidencia entre la escena y cada patrón no sea perfecta y el pico máximo de correlación sea un valor cercano a 1 sin llegar a la unidad, que sería la coincidencia perfecta.

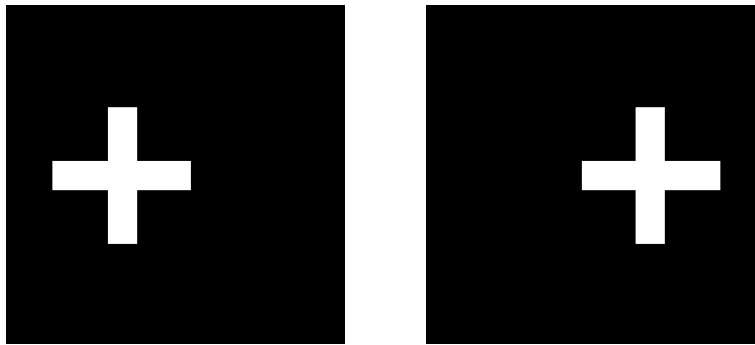


Figura 2.5. Dos imágenes del mismo objeto de interés desplazado al que se le ha realizado la correlación cruzada normalizada.

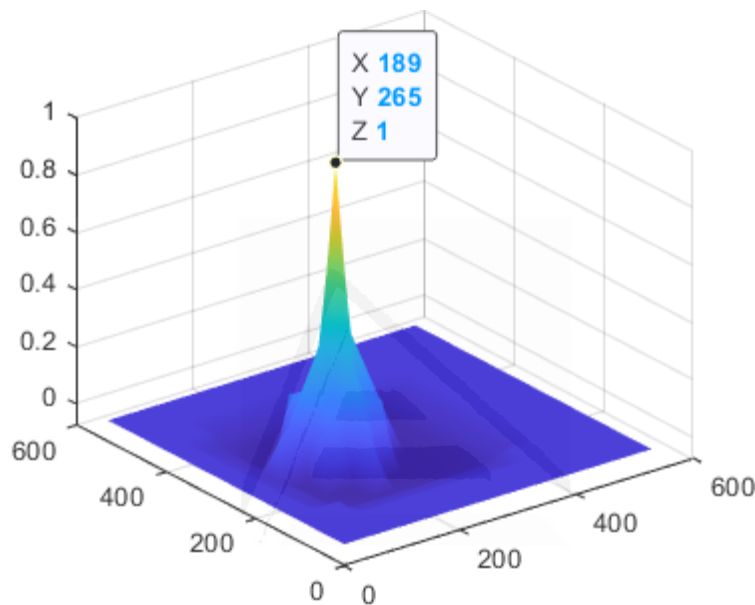


Figura 2.6. Matriz de valores obtenida tras realizar el cálculo de la correlación cruzada normalizada. El pico con un valor de 1 indica la coincidencia perfecta entre la escena y el patrón y la posición en la que ocurre dicha coincidencia dentro de la matriz. Los valores cercanos a 0 indican una pésima coincidencia entre ambas imágenes.

A pesar de las ventajas mencionadas anteriormente, la utilización de la correlación cruzada presenta dos inconvenientes principales. El primero es la dependencia del resultado de la correlación con la luminancia de la escena y el patrón (ya que es un producto de luminancias), lo que puede producir picos de similitud cuando un patrón oscuro se compara con una escena brillante o viceversa, a pesar de que no exista correspondencia y, sin embargo, estos picos pueden ser mayores que el debido a la comparación de un patrón y una escena oscuras, a pesar de tener éstos mayor similitud. El segundo inconveniente es que, por definición, la resolución está limitada a una muestra o un píxel en el caso de imágenes, lo que es a priori una gran desventaja para conseguir detectar movimientos con elevada precisión que es el resultado que aquí se busca.

El primer inconveniente se resuelve utilizando la correlación cruzada normalizada cuyo uso se ha extendido en el estudio de imágenes y cuya ecuación se describe en (2.6). Este cálculo proporciona una matriz de valores adimensionales entre -1 y 1, siendo 1 la coincidencia perfecta y -1 la anticorrelación perfecta, es decir, la coincidencia exacta entre la inversa de la escena y el patrón o la escena y la inversa del patrón. Cuando se obtienen

valores cercanos a 0, indica, por tanto, que la escena y el patrón no se parecen en absoluto (Hayes, 2021).

$$\gamma(m,n) = \frac{\sum_{i,j} [I(i,j) - \bar{I}_{m,n}] [T(i-m, j-n) - \bar{T}]}{\sqrt{\sum_{i,j} [I(i,j) - \bar{I}_{m,n}]^2 \sum_{i,j} [T(i-m, j-n) - \bar{T}]^2}} \quad (2.6)$$

donde I es la imagen de referencia y T el patrón o template, \bar{T} es el valor medio del patrón e \bar{I} el de $I(i,j)$ en la región cubierta por el patrón.

Con respecto al segundo problema, se puede aumentar la resolución obtenida consiguiendo así una resolución subpíxel, mediante un ajuste sobre la escena y el patrón antes de calcular su correlación cruzada (Sutton et al., 2009) o sobre la propia función de correlación para incrementar la precisión en la localización de su máximo (Lei et al., 2015). La primera aproximación tiene el inconveniente de que se produce una manipulación de la información original de las imágenes, de forma que puede distorsionar los sutiles cambios de intensidad que, convenientemente tratados, podrían aportar información sobre los desplazamientos a nivel subpíxel. La segunda aproximación, es decir, el ajuste de la posición del pico de correlación, es un proceso inocuo con la información original de la imagen, ya que únicamente se manipula matemáticamente el resultado de la correlación con el fin de obtener una precisión de la posición del pico inferior al píxel. Por ello, creemos que tiene aplicaciones más interesantes que la interpolación de las intensidades en la imagen y ha sido la técnica elegida para los trabajos que se realizan en esta tesis. Aparte de esto, la interpolación de la escena y el patrón suelen ser procesos computacionalmente intensivos y el aumento de resolución obtenido es muy limitado. En cambio, con el ajuste del pico de correlación, sólo se maneja una pequeña región de la función de correlación alrededor del pico, lo que resulta mucho más rápido de calcular.

Brevemente, la técnica consiste en ajustar la función de correlación en un área local alrededor del pico máximo utilizando una función analítica (Sutton et al., 2009) y reobtener el máximo de la función de correlación a partir de la función ajustada. Este procedimiento puede incrementar la precisión de la localización del pico en casi dos órdenes de magnitud (Ferrer and Mas, 2018; Ferrer et al., 2019), aunque precisa de objetos bien contrastados en las imágenes de referencia y patrón que proporcionen un pico suficientemente agudo como para que el ajuste al entorno local sea lo más preciso posible. Continuando con el ejemplo representado en las dos figuras anteriores, en la figura 2.7 se muestra el ajuste del pico de correlación obtenido y la diferencia entre el pico en valores discretos, es decir, la posición del píxel en el que se encuentra el pico, y el ajustado, que da la localización dentro del píxel en el que se encuentra el pico. Esta diferencia representa el ajuste subpíxel realizado.

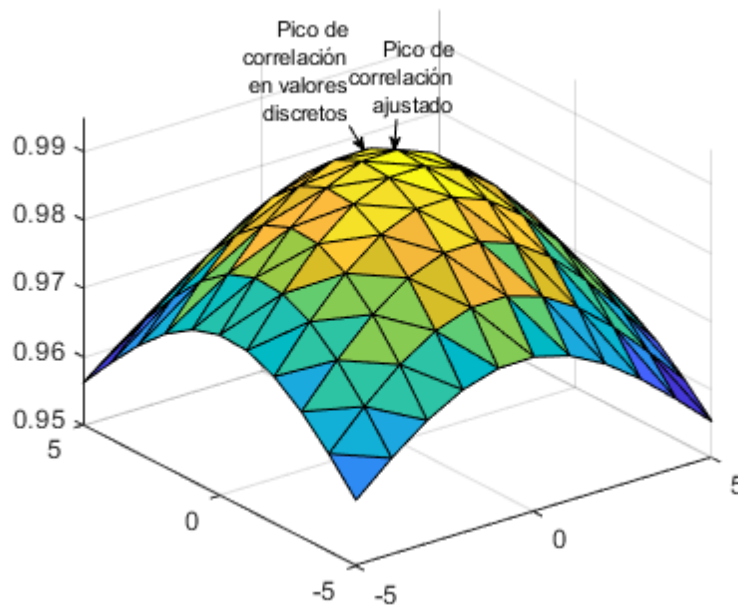


Figura 2.7. Muestra de las posiciones del pico de correlación obtenido en valores discretos y el ajustado. La diferencia entre ambos representa el ajuste subpíxel.

A pesar de la evidente mejora, el procedimiento también introduce un sesgo conocido como “peak-locking”, “pixel-locking” o “atrapamiento de pico” que limita su utilización. Este error implica que la localización del pico obtenida a través del ajuste local está siempre desplazada hacia el entero más cercano (Nogueira et al., 2011).

El origen del atrapamiento de pico se suele atribuir al efecto de aliasing debido a una mala textura de la imagen (Sjödahl and Benckert, 1994) junto con una elección inadecuada de la función de ajuste (Murray et al., 2017). En general, de acuerdo con el límite de Nyquist, una escena con pseudospeckle estará bien muestreada cuando el punto unidad de la textura de speckle sea mayor de 2 px (Stanier et al., 2016).

Sin embargo, se puede dar una explicación alternativa en el dominio espacial. Consideremos una escena cuyo objeto tenga poco ruido y el patrón sea una versión desplazada del objeto original como la que se ha mostrado en la figura 2.5. Cuanto mayor sea el detalle del objeto, más estrecho será el pico de correlación, ya que un desplazamiento pequeño degradará la correspondencia entre el patrón y la escena (Stanier et al., 2016). Si consideramos un dominio local estrecho alrededor del máximo de la función de correlación, habrá un reducido número de muestras para las que tengamos valores alrededor del máximo. De esta manera, el máximo podría tener un excesivo peso en el ajuste posterior del pico y podría empujar al máximo recalculado hacia su posición original. Consecuentemente, se introduce un área de error alrededor del entero más cercano que se corresponde con la posición del máximo original. Asimismo, cuando el máximo está exactamente equidistante entre 2 píxeles vecinos, la distribución del peso está equilibrada y el error es 0. En la figura 2.8 se puede observar la curva en forma de “S” característica del error de atrapamiento de pico.

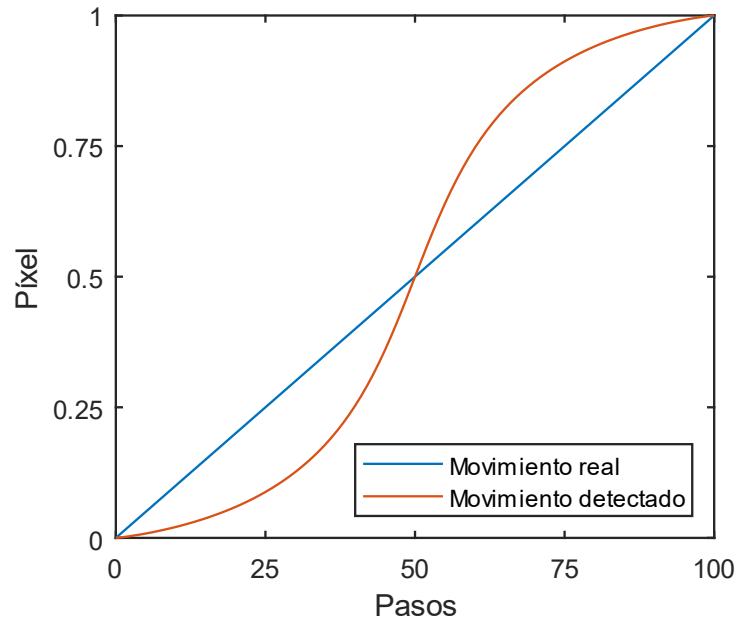


Figura 2.8. Atrapamiento de pico producido sobre un movimiento lineal entre dos píxeles en el que en vez de detectarse una línea recta entre ambos (línea azul) se observa una curva más cercana al primer píxel hasta la mitad del movimiento y más cercana al segundo a partir de la mitad (línea roja).

Se podría tomar un vecindario mayor, aumentando así el número de muestras para ajustar y compensando, de este modo, el excesivo peso del máximo de la correlación, pero esto incluiría información no relacionada con las posiciones, distorsionando, de esta forma, el resultado final. Además, cualquier pico secundario en el entorno seleccionado desequilibraría también el ajuste y empujaría al máximo ajustado hacia él.

De acuerdo con esto, las funciones de ajuste que consideran solamente un pequeño entorno alrededor del pico pueden mostrar un buen ajuste con un vecindario pequeño, pero se verán afectadas por el efecto de atrapamiento de pico con entornos del pico de correlación mayores. Por otra parte, funciones amplias que consideren el pico y la región alrededor de él compensarían este efecto pero requerirían de más cálculos y podrían ser menos precisas. Por lo tanto, una selección adecuada tanto de la función de ajuste como del entorno de aplicación es crítica para disminuir el error de atrapamiento de pico como se verá en el Capítulo 3 en el que se evalúa teóricamente este método en cuanto a los mejores ajustes a utilizar para minimizar los errores propios del método.

2.4. Análisis de luminancia

Uno de los inconvenientes de usar el método de correlación para grandes secuencias numéricas es su carga computacional. La correlación cruzada ya es de por sí intensa numéricamente y depende del número de píxeles que tenga la imagen ($N \times N$). Además, al añadir un ajuste del pico de correlación, se añade tiempo de cálculo. En el dominio temporal,

el método de la correlación cruzada requiere $O(N^4)$ operaciones de multiplicación mientras que en el dominio de la frecuencia, la correlación necesitará $O(N^2 \log_2 N)$ operaciones de multiplicación (Sundararajan, 2017). Sin embargo, en el año 2020, (Wan et al., 2020) propusieron un método alternativo de seguimiento de objetos basado en diferencias en la luminancia. Este método que se describirá seguidamente sólo requiere el orden de $3N^2 + 2N$ operaciones de suma por lo que este método resulta más eficiente que la correlación cruzada.

Aparte de la velocidad, este trabajo resulta interesante porque es capaz de localizar los objetos con precisión sin necesidad de realizar un ajuste, lo que elimina los posibles sesgos de posición que aparecen en el método de la correlación con ajuste del pico.

El método se basa en que la resta de las intensidades de dos imágenes cualesquiera de una misma secuencia indica cambios locales en las imágenes con el paso del tiempo. La suma absoluta de estos cambios en los niveles de gris acumulados proporciona un parámetro que se utiliza para cuantificar el movimiento. Esta aproximación solamente incluye sumas y restas por lo que, aunque se reduce la complejidad del cálculo, se requiere de un conocimiento previo del objeto de interés, debiéndose realizar un proceso de calibración anterior para asociar el valor del parámetro con el desplazamiento (Wan et al., 2020).

Para implementar este método, en primer lugar se calcula la matriz de la diferencia (D) de los niveles de gris entre la imagen original (I_0) y la misma imagen trasladada una cantidad δ_x (I_{δ_x}) menor de 1 px, tal y como se muestra en la ecuación 2.7 y la figura 2.9:

$$D = I_0 - I_{\delta_x} \quad (2.7)$$

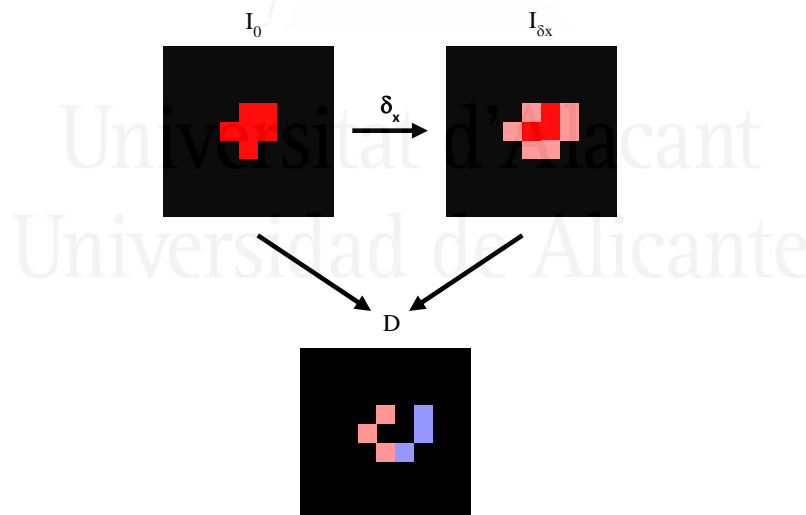


Figura 2.9. Ejemplo gráfico de la imagen de un objeto, la imagen del objeto trasladado y la diferencia entre ambas situaciones (Wan et al., 2020).

Con esta matriz se calcula a continuación la suma de valores en las columnas (C_i) y en las filas (R_j) como se muestra en la ecuación 2.8, donde $1 \leq (i, j) \leq N$, siendo N el número de columnas y filas. Después se calculan las sumas absolutas de C_i (T_C) y R_j (T_R) como se muestra en la ecuación 2.9 y la figura 2.10:

$$C_i = \sum_{j=1}^N D_{ij} \quad R_j = \sum_{i=1}^N D_{ij} \quad (2.8)$$

$$T_C = \sum_i |C_i| \quad T_R = \sum_j |R_j| \quad (2.9)$$

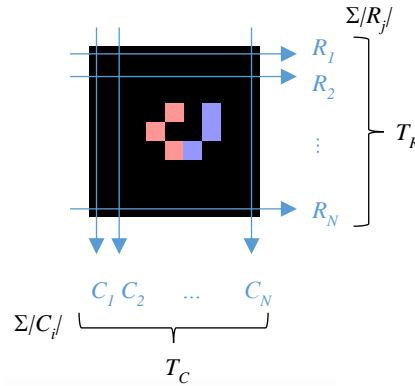


Figura 2.10. Sumas de los niveles de gris por columnas y por filas (Wan et al., 2020).

Nótese que, aunque por simplicidad se asume que la matriz es cuadrada, el método es generalizable a matrices rectangulares. Con este método se utiliza el parámetro T_C como indicador del movimiento en la dirección x y T_R como indicador del movimiento en la dirección y . En el caso de que en alguna de las dos direcciones no haya movimiento, ese valor de la suma absoluta proporcionará información sobre el ruido, y su tamaño dependerá de la forma del objeto y del tamaño del movimiento. Obsérvese que dichos parámetros no proporcionan el valor del movimiento por sí mismos y que son muy dependientes de la imagen. Por ello, para la aplicación de estos métodos se debe calibrar el método en primer lugar, estableciendo la correspondencia entre los parámetros y el desplazamiento real y, una vez calibrado, proceder a los experimentos.

De manera sencilla y visual se puede observar el funcionamiento de este método en la figura 2.11 en la que se emplea un objeto blanco (nivel de gris 255) sobre fondo negro (nivel de gris 0) que se desplaza 1 px horizontalmente. Se realiza, como se ha comentado anteriormente, la diferencia entre ambas imágenes y luego se suman por filas y por columnas los valores obtenidos. Finalmente, se calcula el sumatorio en valor absoluto de todas las filas y todas las columnas por separado para obtener T_R y T_C respectivamente. Como el movimiento del objeto ha sido solo en horizontal, T_C proporciona un valor y en T_R el resultado es 0. Si se tratara de imágenes reales, seguramente T_R no sería 0, sino un valor mucho más bajo al obtenido en T_C , indicando el error de base del método.

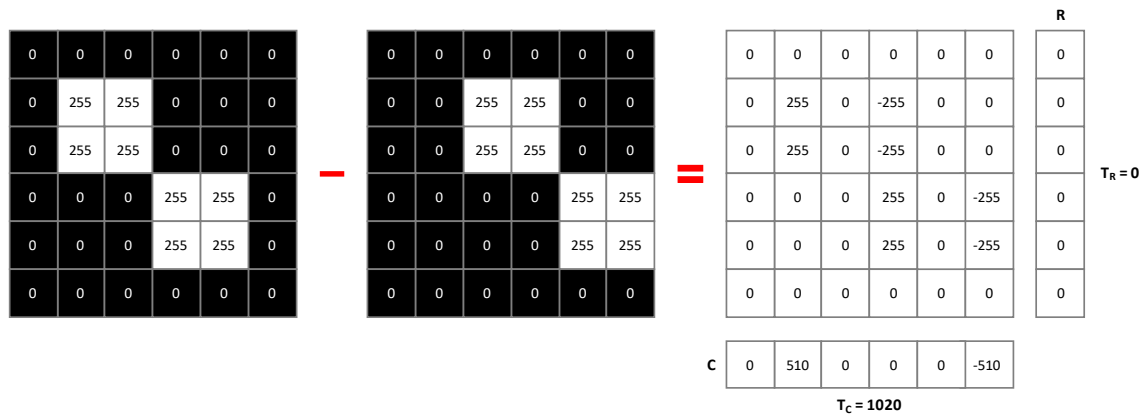


Figura 2.11. Funcionamiento del método de análisis de la luminancia utilizando una imagen digital sencilla (Grehan, 2020).

Para poder apreciar el funcionamiento de este método con objetos más complejos, tanto utilizando cálculos teóricos como en experimentos reales, en el Capítulo 5 se analiza esta propuesta comparada con la correlación cruzada normalizada para poder obtener resultados comparables con el método comúnmente utilizado para medir desplazamientos con una precisión subpíxel.



Universitat d'Alacant
Universidad de Alicante

CAPÍTULO 3: Evaluación teórica del método subpíxel por ajuste del pico de correlación

Como se ha comentado en el Capítulo 2.3, cuando es necesario utilizar dianas presentes de forma natural en la escena de interés, una de las herramientas más utilizadas para seguir el movimiento del objeto es la correlación cruzada normalizada junto con un ajuste y posterior relocalización del pico de correlación. Esta técnica se utiliza para determinar si dos imágenes tienen una coincidencia alta y en qué punto ocurre dicha similitud (Lava, 2015). Por ello, en este capítulo se pretende evaluar de forma teórica qué ajustes son adecuados para proporcionar un buen seguimiento del movimiento del objeto de interés disminuyendo el error de atrapamiento de pico que introduce el propio ajuste.

Antes de analizar teóricamente el método subpíxel por ajuste del pico de correlación, es necesario mencionar que se ha encontrado en la bibliografía que desenfoque ligeramente el objeto de interés puede reducir el error de atrapamiento de pico (Michaelis et al., 2016; Overmars et al., 2010; Zhou et al., 2015). Esto sucede porque una correlación entre objetos desenfocados tenderá a producir un pico ligeramente más suave y redondeado y con apariencia gaussiana, lo que facilitará el ajuste. Lamentablemente, los trabajos encontrados no proporcionan un valor o rango óptimo de desenfoque ni resultados cuantitativos, ya que simplemente introducen el desenfoque de modo experimental y se observa una mejora en los resultados. Por este motivo, se ha decidido llevar a cabo un estudio teórico empleando imágenes digitales de un banco de imágenes público a las que les introducimos después distintos niveles de desenfoque digital para conocer aquellos valores de desenfoque que reducen el error del atrapamiento de pico en mayor medida.

Previamente a la descripción detallada de cómo se ha realizado el estudio teórico, es importante mencionar cómo introducir un desenfoque en una secuencia. Para realizar la

correlación cruzada normalizada (*) entre la escena (I) y el patrón o template (T) sin introducir ningún desenfoque (bien porque no es necesario o bien porque la secuencia ya se ha grabado con el desenfoque deseado), se emplea la ecuación 2.6 y que recordamos aquí (3.1):

$$\gamma(m,n) = \frac{\sum_{i,j} [I(i,j) - \overline{I_{m,n}}] [T(i-m, j-n) - \overline{T}]}{\sqrt{\sum_{i,j} [I(i,j) - \overline{I_{m,n}}]^2 \sum_{i,j} [T(i-m, j-n) - \overline{T}]^2}} \quad (3.1)$$

No obstante, cuando es necesario introducir un desenfoque digital en la secuencia, se debe introducir un filtro de desenfoque gaussiano (G) a través de un kernel de suavizado gaussiano con una desviación estándar especificada por el sigma introducido (r_b). Este filtro se describe matemáticamente según la ecuación 3.2.

$$G_{r_b} = \exp\left(-\frac{x^2}{2 \cdot r_b^2}\right) \quad (3.2)$$

El filtro de desenfoque gaussiano se convoluciona (\otimes) con la escena (I) y con el patrón (T) como se muestra en la ecuación 3.3 donde el subíndice d se refiere a la función desenfocada.

$$C_d = I_d * T_d \equiv (I \otimes G_{r_b}) * (T \otimes G_{r_b}) \quad (3.3)$$

Se pueden aplicar las propiedades de la convolución y la correlación para tratar de simplificar esta expresión:

$$C_d = I_d * T_d = I_d * (T \otimes G_{r_b}) = (I_d * T) \otimes G_{r_b} = [(I \otimes G_{r_b}) * T] * G_{r_b} \quad (3.4)$$

Dado que la gaussiana es una función simétrica, la correlación y la convolución con esta función son intercambiables, de modo que:

$$C_d = [(I * T) \otimes G_{r_b}] \otimes G_{r_b} \equiv (I * T) \otimes (G_{r_b} \otimes G_{r_b}) \quad (3.5)$$

Es decir, la correlación de dos señales desenfocadas es equivalente a realizar un doble desenfoque de la función de correlación. Si además tenemos en cuenta que la autoconvolución de una gaussiana es una nueva gaussiana (3.6) con un sigma ligeramente mayor (3.7) (Bromiley, 2013), concluimos que correlacionar dos señales desenfocadas es equivalente a correlacionar dos señales enfocadas y desenfocar a posteriori el pico de correlación, tal y como se muestra en (3.8).

$$G_{r_b} \otimes G_{r_b} = G'_{r_b'} = \exp\left(-\frac{x^2}{2 \cdot (r_b')^2}\right) \quad (3.6)$$

$$r_b' = r_b \cdot \sqrt{2} \quad (3.7)$$

$$C_d = I_d * T_d = (I * T) \otimes G'_{r_b} \equiv C \otimes G'_{r_b} \quad (3.8)$$

Dado que el perfil de una función desenfocada es más suave que el perfil de la función original, la función de correlación entre dos funciones desenfoçadas dará como resultado una función más suave que la obtenida cuando correlacionamos las funciones enfocadas. Los perfiles suaves resultan más sencillos de interpolar que los abruptos lo que explicaría la mejoría de los errores y la disminución del efecto de atrapamiento de pico.

Para estudiar la tolerancia al desenfoque de modo numérico, se ha generado con Matlab (*MATLAB*, R2020a) una secuencia de un círculo desplazado con la propiedad de traslación de la Transformada de Fourier (conocida en inglés como Shifting Property) que establece que la Transformada de Fourier de un objeto desplazado es igual a la Transformada de Fourier del objeto original multiplicada por una fase (Gaskill, 1978). Así, generando una fase conveniente, se puede calcular la función desplazada incluso una fracción de píxel.

El círculo generado tiene 100 px de radio y en la secuencia realiza un desplazamiento horizontal total de 1 px en pasos de 0.01 px. Una vez generada la secuencia del círculo en movimiento, se ha calculado la correlación cruzada con diferentes niveles de desenfoque.

Este filtro de desenfoque gaussiano se ha introducido mediante el algoritmo *imgaussfilt* de Matlab, que filtra la imagen con un kernel de suavizado gaussiano 2D con la desviación estándar especificada por el sigma (r_b) introducido. Los valores de sigma utilizados han sido: entre 0 (sin desenfoque) y 2 en pasos de 0.1; a partir de 2 hasta 10 en pasos de 0.2; de 10 a 30 en pasos de 1; de 30 a 40 en pasos de 2; y de 40 a 60 en pasos de 5. Se ha realizado de esta forma para poder tener un mapa detallado de lo que ocurre con distintos desenfoques. Además, se ha seleccionado un mayor número de desenfoques pequeños ya que se ha supuesto que la mejora sucederá para desenfoques leves o moderados pero que con desenfoques fuertes los resultados empeorarán de nuevo ya que el objeto de interés dejaría de ser reconocible.

Por otra parte, en cuanto al ajuste del entorno del pico de correlación y posterior localización del pico, se han utilizado el ajuste mediante splines que realiza Matlab (“Thin-plate smoothing spline - MATLAB tpaps - MathWorks España,” 2006), un ajuste polinómico de segundo grado utilizando la función representada en la ecuación (3.9) y un ajuste gaussiano representada en la ecuación (3.10) por ser los tres métodos de ajuste más utilizados en la bibliografía (Michaelis et al., 2016; Roesgen, 2003; Nobach et al., 2005).

$$f(x, y) = p00 + p01 * y + p02 * y^2 + p10 * x + p11 * x * y + p20 * x^2 \quad (3.9)$$

$$f(x, y) = a1 + a2 * \exp\left(-\left(\frac{x-a3}{a4}\right)^2\right) * \exp\left(-\left(\frac{y-a5}{a6}\right)^2\right) \quad (3.10)$$

donde los valores asociados a p y a a son los coeficientes de ajuste.

Los tres ajustes se han realizado empleando un entorno alrededor del pico de correlación igual, de tamaño 7×7 px y situando el pico de correlación en el píxel central. Previamente a la realización de este análisis, se ha evaluado mediante un breve estudio si mantener el pico de correlación centrado o descentrarlo respecto a la matriz utilizada para realizar el ajuste puede mejorarlo, obteniéndose valores similares en todos los casos, por lo que todos los estudios mencionados a continuación se han realizado con el pico de correlación centrado en la matriz utilizada para realizar el ajuste, sea del tipo que sea el ajuste utilizado.

Para comparar los resultados se ha calculado el error máximo de posicionamiento del pico de correlación (3.11) y su desviación estándar (3.12) para todos los niveles de desenfoque gaussiano introducido y con los tres ajustes del pico de correlación realizados. Tanto el error máximo de posicionamiento como la desviación estándar son unos estimadores de la bondad del método y pueden calcularse debido a que la posición real del objeto es conocida al haber diseñado nosotros el movimiento.

$$MaxErr = \max(|x - x'|) \quad (3.11)$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^N ((x_i - x_i') - \mu)^2}{N - 1}} \quad (3.12)$$

donde x es el valor obtenido, x' el valor real, N el número total de valores y μ la media.

Estos errores junto con la desviación estándar obtenidos para todos los desenfoques introducidos y los tres ajustes de la función de correlación empleados se representan en la figura 3.1, en la que se observan tendencias generales similares para los tres ajustes. Dado que en la zona para pequeños desenfoques (hasta 20 px) las líneas se superponen, se ha representado este intervalo ampliado en la figura 3.2.

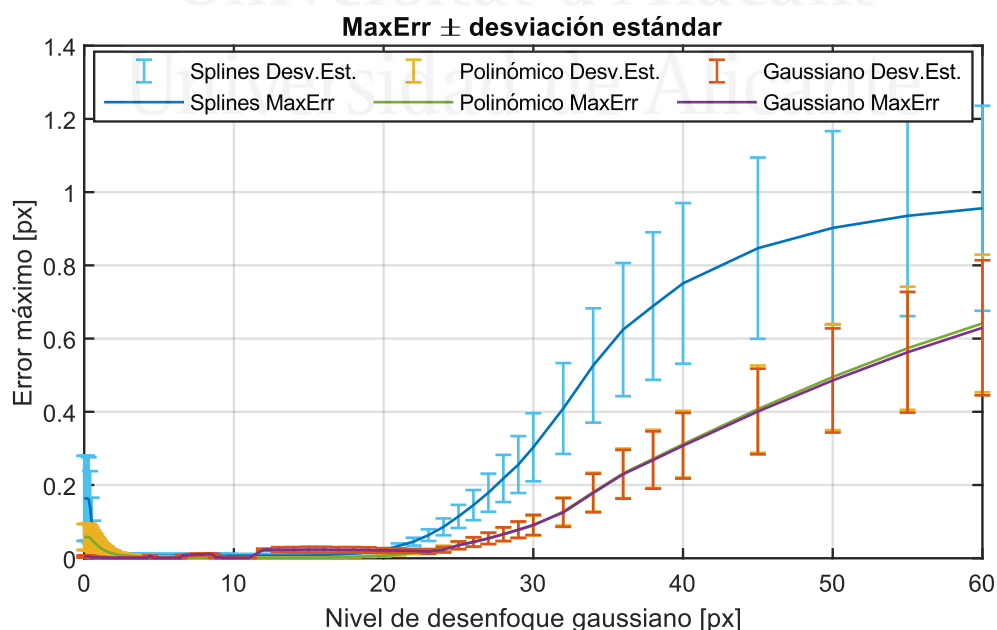


Figura 3.1. Error máximo junto con la desviación estándar para los distintos niveles de desenfoque con los tres ajustes del pico de correlación.

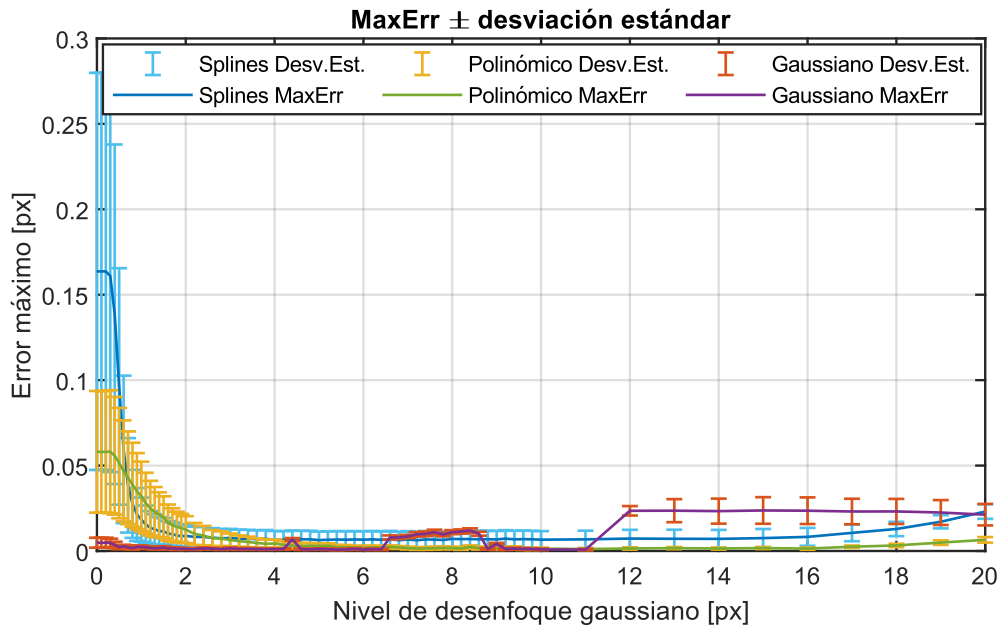


Figura 3.2. Error máximo junto con la desviación estándar para sigmas de desenfoque gaussiano hasta 20 px con los tres ajustes del pico de correlación.

En el caso de los ajustes mediante splines y polinómico se consigue una mejora rápida con un pequeño desenfoque (menor de 1 px) pero para el ajuste gaussiano ya se obtienen buenos resultados incluso sin desenfocar. Además, estos errores se mantienen pequeños y estables al aumentar el desenfoque hasta que, a partir de sigmas de desenfoque mayores de 20 px, empieza a crecer de nuevo el error máximo. Por otra parte, en el ajuste splines aumenta mucho más rápido el error máximo al aumentar el desenfoque que en los otros dos ajustes analizados.

Sabiendo que aumentar ligeramente el desenfoque ayuda a mejorar el ajuste realizado, se ha completado el estudio utilizando las 5 figuras geométricas mostradas en la figura 3.3. En este caso, se ha variado no solo el desenfoque gaussiano (r_b) desde 0 hasta 5 en pasos de 1 px, sino también el entorno del pico de correlación (Nbd) con tamaños de matrices de ajuste desde 3×3 hasta 11×11 px solo con números impares de píxeles para mantener en todo momento el pico de correlación centrado en la matriz de ajuste. Igual que en el círculo generado anteriormente, se ha realizado un movimiento total de 1 px en pasos de 0.01 px.

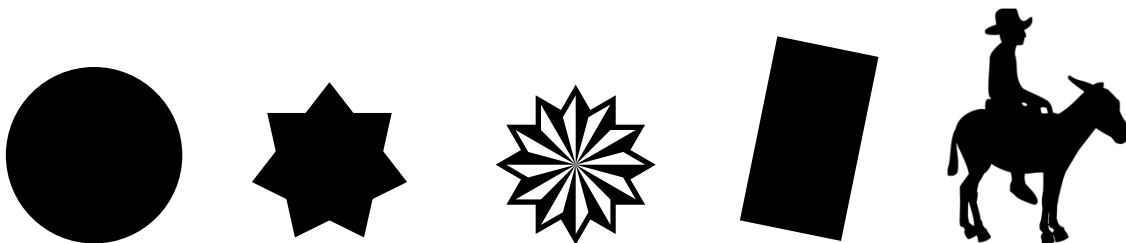


Figura 3.3. Figuras utilizadas para realizar un desplazamiento teórico y evaluar su error máximo con diferentes desenfoques y entornos alrededor del pico de correlación. De izquierda a derecha, las nombramos como *Círculo*, *Estrella*, *Flor*, *Rectángulo* y *Jinete*.

A estas figuras binarias se les ha realizado la correlación cruzada normalizada utilizando el algoritmo *normxcorr2* de Matlab, según la expresión (3.1), empleando la primera imagen de cada secuencia como escena y generando el desplazamiento a partir de esta primera imagen reducida en 8 px por cada lado, evitando así que las imágenes desplazadas se muevan por fuera de los bordes de la imagen de referencia (Sjödahl, 1997).

En la figura 3.4 se presenta un diagrama de flujo del programa utilizado en el que se calcula al final el error máximo junto con la desviación estándar producidos con cada combinación de tipo de ajuste, tamaño de la matriz de ajuste y desenfoque introducido para las 5 figuras analizadas.

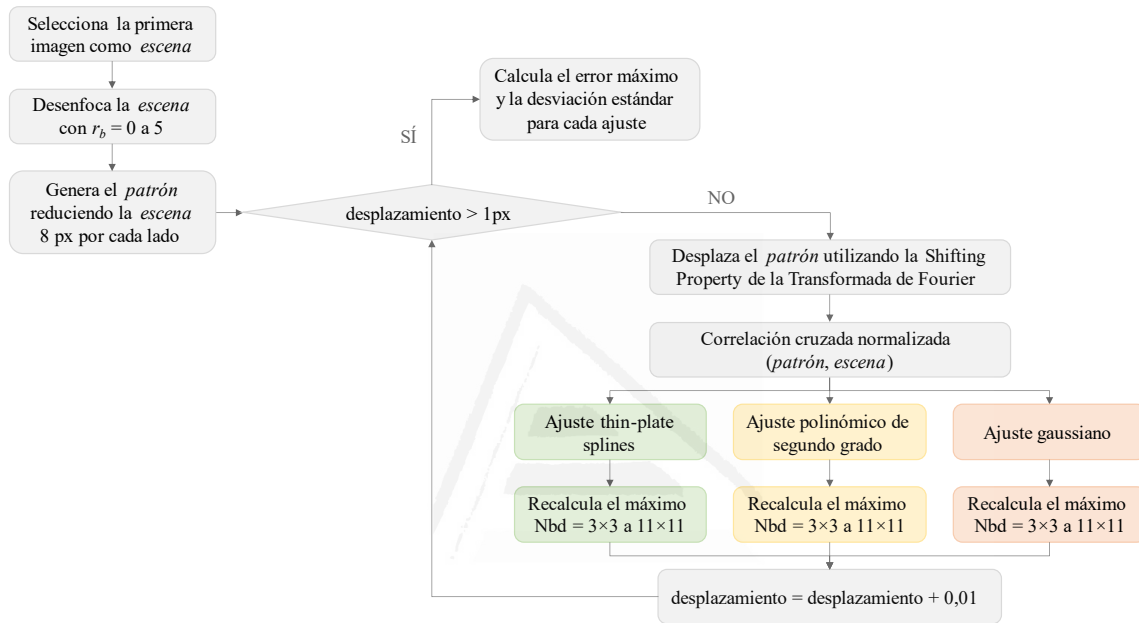


Figura 3.4. Diagrama de flujo del programa realizado con Matlab para calcular la correlación cruzada variando el ajuste utilizado, su tamaño y el desenfoque para las 5 figuras representadas en la figura 3.3.

En la tabla 3.1 se muestran los mejores y peores resultados obtenidos para cada figura y método de ajuste, lo que corresponde con el mínimo y máximo error máximo (MaxErr) respectivamente.

Tabla 3.1. Mínimo y máximo error máximo (mejor y peor resultado respectivamente) utilizando los tres ajustes para cada figura (Fig.) junto con los valores de píxeles alrededor del pico de correlación (Nbd) y desenfoque gaussiano (r_b) que lo producen, así como la media (μ) y desviación estándar (σ) de los resultados de la secuencia.

| Fig. | | Ajuste splines | | Ajuste polinómico | | Ajuste gaussiano | |
|------------|------------------|--|-----------------------------|--|------------------------------|---|---------------|
| | | Mínimo | Máximo | Mínimo | Máximo | Mínimo | Máximo |
| Círculo | $\mu \pm \sigma$ | $6 \times 10^{-5} \pm 2 \times 10^{-4}$ | $-7 \times 10^{-4} \pm 0.1$ | $3 \times 10^{-5} \pm 9 \times 10^{-4}$ | 0.001 ± 0.03 | $4 \cdot 10^{-5} \pm 3 \cdot 10^{-5}$ | 0.1 ± 0.5 |
| | MaxErr | 6×10^{-4} | 0.1760 | 0.0012 | 0.1048 | $1 \cdot 10^{-4}$ | 1.8678 |
| | Nbd | 9x9 | 3x3 | 5x5 | 3x3 | 5x5 | 3x3 |
| | r_b | 5 | 0 | 5 | 0 | 5 | 2 |
| | | | | | | | |
| Estrella | $\mu \pm \sigma$ | $-1 \times 10^{-5} \pm 2 \times 10^{-4}$ | $-7 \times 10^{-4} \pm 0.1$ | $-1 \times 10^{-5} \pm 4 \times 10^{-4}$ | 0.001 ± 0.04 | $4 \cdot 10^{-5} \pm 4 \cdot 10^{-5}$ | 0.2 ± 0.5 |
| | MaxErr | 4×10^{-4} | 0.1766 | 0.0012 | 0.1057 | $2 \cdot 10^{-4}$ | 1.8875 |
| | Nbd | 11x11 | 3x3 | 3x3 | 3x3 | 9x9 | 3x3 |
| | r_b | 5 | 0 | 5 | 0 | 5 | 2 |
| | | | | | | | |
| Flor | $\mu \pm \sigma$ | $2 \times 10^{-4} \pm 2 \times 10^{-4}$ | $7 \times 10^{-4} \pm 0.1$ | $1 \times 10^{-4} \pm 6 \times 10^{-4}$ | -0.001 ± 0.04 | $2 \times 10^{-5} \pm 5 \times 10^{-5}$ | 0.3 ± 0.6 |
| | MaxErr | 6×10^{-4} | 0.1777 | 0.0020 | 0.1041 | 1×10^{-4} | 1.7600 |
| | Nbd | 9x9 | 3x3 | 3x3 | 3x3 | 7x7 | 3x3 |
| | r_b | 5 | 0 | 5 | 0 | 1 | 4 |
| | | | | | | | |
| Rectángulo | $\mu \pm \sigma$ | $-1 \times 10^{-4} \pm 6 \times 10^{-4}$ | $6 \times 10^{-4} \pm 0.1$ | $9 \times 10^{-6} \pm 3 \times 10^{-4}$ | 0.001 ± 0.08 | $3 \times 10^{-4} \pm 0.02$ | 0.5 ± 0.5 |
| | MaxErr | 0.0012 | 0.1682 | 5×10^{-4} | 0.1241 | 0.0308 | 1.8806 |
| | Nbd | 9x9 | 3x3 | 3x3 | 5x5 | 5x5 | 3x3 |
| | r_b | 4 | 0 | 5 | 0 | 5 | 0 |
| | | | | | | | |
| Jinete | $\mu \pm \sigma$ | $-2 \times 10^{-4} \pm 4 \times 10^{-4}$ | $-8 \times 10^{-4} \pm 0.1$ | $-6 \times 10^{-5} \pm 3 \times 10^{-4}$ | $-5 \times 10^{-4} \pm 0.05$ | $5 \times 10^{-4} \pm 0.006$ | 0.6 ± 0.7 |
| | MaxErr | 7×10^{-4} | 0.1672 | 8×10^{-4} | 0.0744 | 0.0083 | 1.9900 |
| | Nbd | 11x11 | 3x3 | 3x3 | 5x5 | 7x7 | 3x3 |
| | r_b | 5 | 0 | 5 | 0 | 1 | 0 |
| | | | | | | | |

Con el fin de analizar los resultados de la tabla 3.1 de una manera más visual, se han representado en la figura 3.5 los valores del error máximo para las 5 figuras utilizadas y los tres ajustes empleados, solo para la dirección horizontal. Asimismo, se han fijado los ejes de las gráficas por tipo de ajuste utilizado para poder comparar los resultados. Sin embargo, no ha sido posible fijar los ejes de todos los ajustes por igual ya que el caso gaussiano proporciona unos errores mucho más pequeños que los otros dos ajustes. Además, también ha sido necesario suprimir los resultados con entornos alrededor del pico de correlación pequeños (Nbd = 3x3 y 5x5) para el ajuste gaussiano ya que estos entornos proporcionan unos errores mucho mayores que los otros tres entornos analizados y su representación enmascara los mejores resultados para este ajuste que son los que realmente interesa conocer.

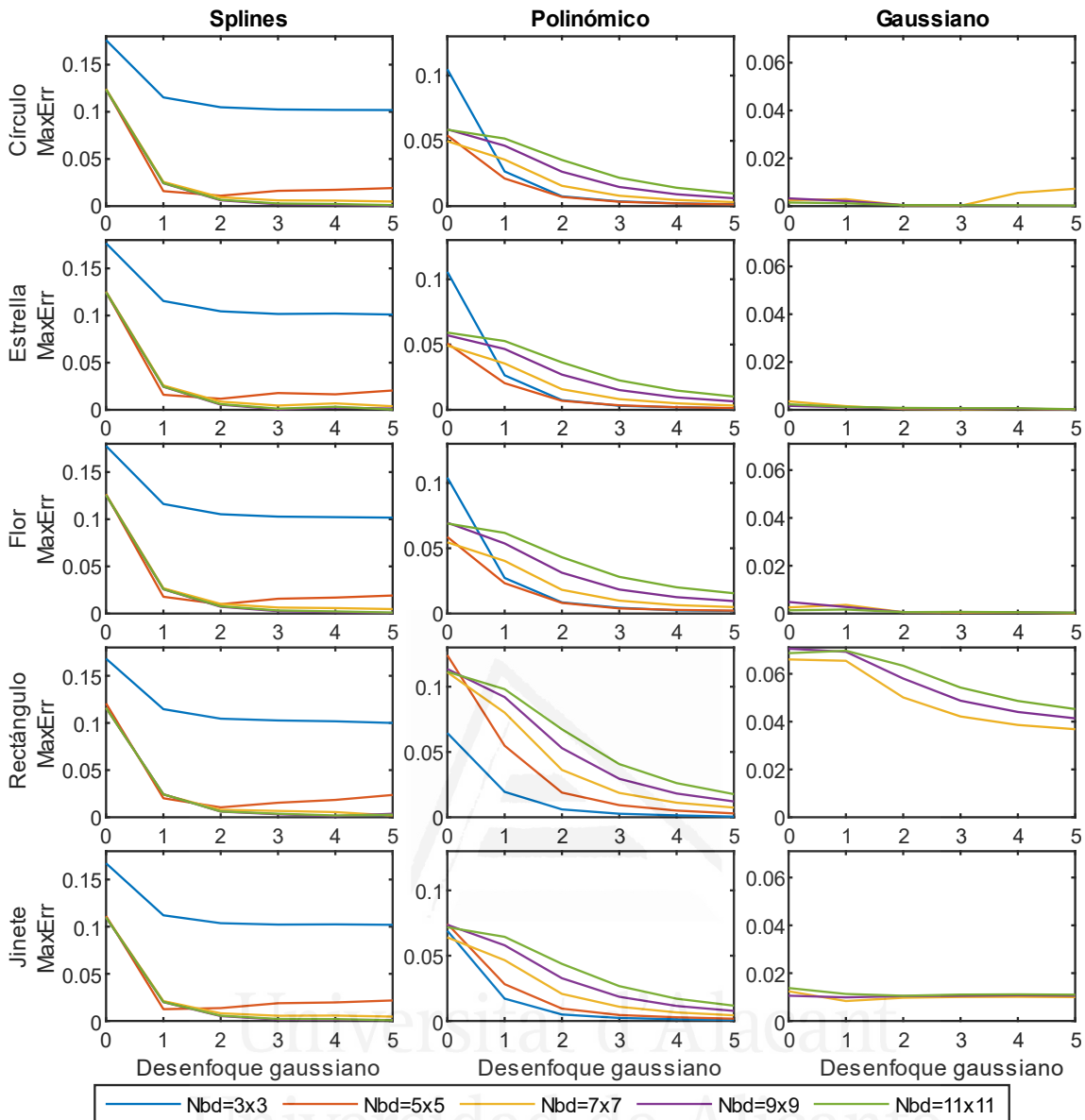


Figura 3.5. Error máximo para los distintos niveles de desenfoque gaussiano (r_b) cambiando el entorno del pico de correlación analizado (Nbd) para las 5 figuras representadas en la figura 3.3 y con los tres ajustes estudiados. Se han suprimido los entornos alrededor del pico de correlación pequeños (Nbd = 3×3 y 5×5) para el ajuste gaussiano para mejorar su visualización por proporcionar unos errores mucho mayores que los otros tres entornos.

A partir de esta gráfica, se puede observar que los entornos alrededor del pico de correlación muy pequeños (Nbd = 3×3) suelen proporcionar errores más grandes que el resto de entornos para los tres tipos de ajuste. Además, en el ajuste gaussiano también se produce un error muy alto con un entorno algo mayor (Nbd = 5×5) que, como se ha comentado anteriormente, ha sido necesario suprimir de la gráfica para lograr una buena visualización de los resultados. Por otra parte, independientemente del objeto utilizado, los errores más pequeños los proporciona el ajuste gaussiano, siempre que se seleccione un entorno mediano o grande, aunque a partir de sigmas de desenfoque de 3 px, los errores producidos con los tres ajustes son parecidos. Nótese por último que el ajuste gaussiano empeora ligeramente sus resultados al utilizar objetos sin simetría radial como son el rectángulo y el jinete.

Con estos análisis se puede concluir que con un pequeño desenfoque gaussiano (el cual se puede introducir antes o después de realizar la correlación cruzada porque produce el mismo efecto) ya se nota una gran mejoría, aunque, dependiendo del objeto evaluado, puede ser mejor o peor un tipo u otro de ajuste del pico de correlación.

Para continuar con las aplicaciones prácticas de la técnica de correlación cruzada, una de sus aplicaciones consiste en el seguimiento de deformaciones en materiales, ya que, si la secuencia de imágenes se captura con una resolución espacial suficientemente alta, la deformación superficial del material puede modelizarse a nivel local como un desplazamiento. Por tanto, sería posible obtener un mapa de desplazamiento mediante la correlación cruzada. Para ello se necesita una referencia que esté homogéneamente distribuida por el material, como, por ejemplo, una textura granulada o de speckle.

El término speckle se refiere a un patrón característico de puntos aleatorios que se produce como resultado de la interferencia de muchas ondas con la misma frecuencia pero diferentes fases y amplitudes, proporcionando una onda resultante cuya amplitud e intensidad varía aleatoriamente. Esto se puede conseguir reflejando un láser en una superficie rugosa (“Encyclopedia of Laser Physics and Technology - Speckle,” 2007; “Speckle pattern,” 2020). Este tipo particular de textura se suele utilizar en imágenes porque se trata de una imagen compleja fácilmente reproducible que se puede generar numéricamente y luego imprimir y pegar sobre una superficie o se puede simular mediante pulverización de pintura sobre un fondo uniforme en la superficie para hacer experimentos reales. En la figura 3.6 se muestra un ejemplo de imagen con speckle obtenida del banco de imágenes público de la Society for Experimental Mechanics y la International Digital Image Correlation Society (“Previous DIC Challenge 1.0 Data,” 2020).

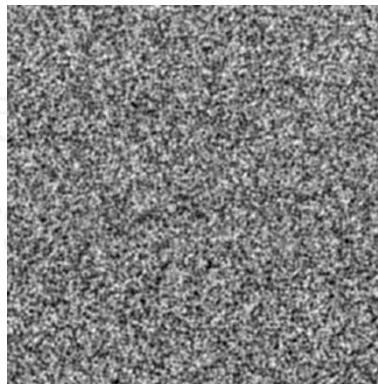
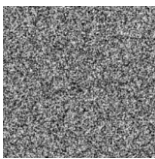
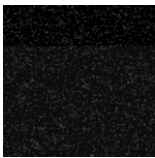
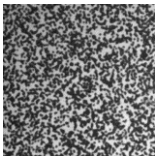

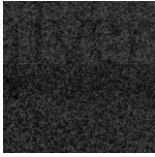


Figura 3.6. Muestra de ejemplo de una imagen con speckle.

Con el fin de evaluar la influencia de las funciones de ajuste del pico y el efecto del desenfoque, se ha extendido este estudio evaluando imágenes de pseudospeckle. Para ello, se han utilizado varias secuencias del banco de imágenes anteriormente mencionado (“Previous DIC Challenge 1.0 Data,” 2020). Las secuencias se componen de imágenes de speckle generadas sintéticamente y cuentan con una descripción completa del contraste, el ruido y el movimiento subpíxel que realizan. En total, en este banco de imágenes hay 19 secuencias diferentes, aunque para el estudio realizado (Publicación 1 del Anexo IV, (Tomás et al., 2020)) solo se han utilizado las 5 primeras que tienen un desplazamiento igual en horizontal y en vertical y el ruido es constante en todas las imágenes de la secuencia. Para más información, en (Reu et al., 2018) se explican con detalle las características de todas las

secuencias presentes en este banco de imágenes. En la tabla 3.2 se muestran, a modo de resumen, las características de las secuencias utilizadas en este estudio junto con la primera imagen de la secuencia. Todas las secuencias empleadas realizan un desplazamiento horizontal y vertical de la primera imagen de acuerdo con los pasos especificados en la columna “Desplazamiento”. Además, el número de imágenes de cada secuencia es la cantidad necesaria para completar 1 píxel de desplazamiento acumulado.

Tabla 3.2. Imágenes de muestra de las secuencias utilizadas junto con sus características (contraste y ruido) y el desplazamiento entre imágenes de la secuencia.

| Secuencia | Muestra | Contraste | Ruido [niveles de gris] | Desplazamiento [píxeles] |
|-----------|---|-----------|-------------------------|------------------------------|
| 1 |  | Variable | 1.5 | $\delta_x = \delta_y = 0.05$ |
| 2 |  | 0 a 50 | 8 | $\delta_x = \delta_y = 0.05$ |
| 3 |  | 0 a 200 | 1.5 | $\delta_x = \delta_y = 0.1$ |
| 4 |  | 0 a 50 | 8 | $\delta_x = \delta_y = 0.1$ |
| 5 |  | Variable | 1.5 | $\delta_x = \delta_y = 0.1$ |

Para este estudio, igual que en el estudio anterior, se ha realizado la correlación cruzada normalizada mediante el algoritmo *normxcorr2* de Matlab (ecuación 3.1) tomando como escena la primera imagen de cada secuencia y como patrón todas las demás imágenes reducidas en 8 px por cada lado para evitar que las imágenes desplazadas se muevan por fuera de los bordes de la imagen de referencia (Sjödahl, 1997) y realizando los mismos tres ajustes del pico de correlación que antes variando los parámetros de N_{bd} y r_b . De nuevo, en la figura 3.7 se muestra un diagrama de flujo de esta segunda parte ya que, al tener toda la secuencia ya generada y no tener que desplazar la primera imagen para generarla, su funcionamiento es ligeramente distinto al del experimento anterior.

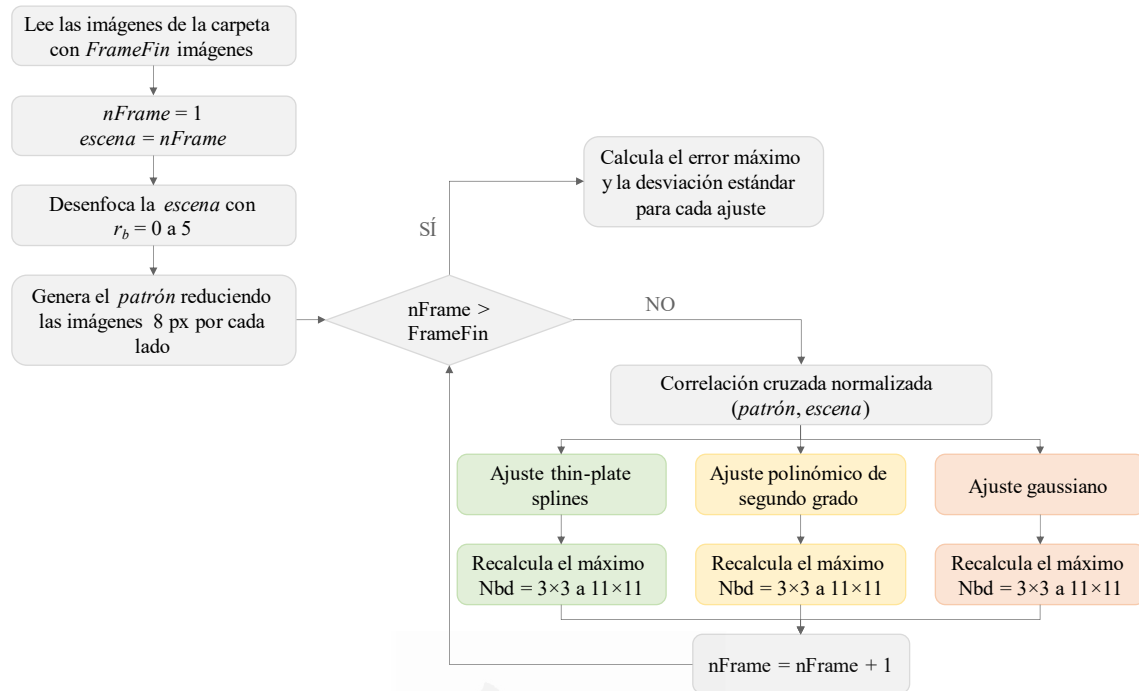


Figura 3.7. Diagrama de flujo del programa realizado con Matlab para analizar las imágenes con speckle del banco de imágenes público (“Previous DIC Challenge 1.0 Data,” 2020).

Igual que para los objetos digitales mostrados anteriormente, se han evaluado tamaños de matrices de ajuste alrededor del pico de correlación desde 3×3 hasta 11×11 px utilizando solo los tamaños impares para mantener el pico de correlación centrado y se ha ido cambiando el radio del desenfoco gaussiano desde 0 (sin desenfocar) hasta 5 px utilizando solamente los números enteros. Todas estas combinaciones de Nbd, r_b y método de ajuste supone un total de 90 pruebas por cada secuencia de imágenes analizada, por lo que se hace difícil su representación y análisis por la gran cantidad de datos obtenidos, así que en la tabla 3.3 se muestran los mejores y peores resultados obtenidos para cada secuencia y método de ajuste, lo que corresponde con el mínimo y máximo error máximo (MaxErr) respectivamente. Se presentan solamente los errores obtenidos en la dirección horizontal igual que se ha hecho para las pruebas anteriores con las figuras desplazadas digitalmente para no duplicar resultados, ya que son muy similares en ambas direcciones.

Tabla 3.3. Mínimo y máximo error máximo (mejor y peor resultado respectivamente) utilizando los tres ajustes para cada secuencia (Sec.) junto con los valores de píxeles alrededor del pico de correlación (Nbd) y desenfoque gaussiano (r_b) que lo producen, así como la media (μ) y desviación estándar (σ) de los resultados de la secuencia.

| Sec. | | Ajuste splines | | Ajuste polinómico | | Ajuste gaussiano | |
|------|------------------|---|-------------------|-------------------------------|-------------------|--|---------------|
| | | Mínimo | Máximo | Mínimo | Máximo | Mínimo | Máximo |
| 1 | $\mu \pm \sigma$ | -0.001 ± 0.001 | -0.004 ± 0.09 | -0.001 ± 0.001 | -0.009 ± 0.1 | -0.001 ± 0.001 | 0.8 ± 0.9 |
| | MaxErr | 0.0037 | 0.1105 | 0.0041 | 0.1658 | 0.0022 | 2.3500 |
| | Nbd | 9×9 | 3×3 | 3×3 | 11×11 | 11×11 | 3×3 |
| | r_b | 5 | 0 | 5 | 0 | 0 | 3 |
| 2 | $\mu \pm \sigma$ | -0.005 ± 0.006 | -0.01 ± 0.09 | -0.004 ± 0.004 | -0.02 ± 0.07 | -0.008 ± 0.006 | 0.9 ± 0.8 |
| | MaxErr | 0.0160 | 0.1290 | 0.0128 | 0.1282 | 0.0206 | 2.3893 |
| | Nbd | 9×9 | 3×3 | 3×3 | 11×11 | 9×9 | 3×3 |
| | r_b | 1 | 5 | 0 | 0 | 0 | 3 |
| 3 | $\mu \pm \sigma$ | $2 \times 10^{-4} \pm 7 \times 10^{-4}$ | -0.008 ± 0.09 | $4 \times 10^{-4} \pm 0.002$ | -0.006 ± 0.04 | $-2 \times 10^{-4} \pm 7 \times 10^{-4}$ | 0.4 ± 0.6 |
| | MaxErr | 0.0013 | 0.1143 | 0.0027 | 0.0652 | 0.0014 | 1.9000 |
| | Nbd | 9×9 | 3×3 | 5×5 | 11×11 | 7×7 | 3×3 |
| | r_b | 4 | 0 | 4 | 0 | 1 | 5 |
| 4 | $\mu \pm \sigma$ | $-2 \times 10^{-4} \pm 0.007$ | -0.008 ± 0.1 | 0.003 ± 0.006 | -0.01 ± 0.1 | $4 \times 10^{-4} \pm 0.002$ | 0.7 ± 0.6 |
| | MaxErr | 0.0113 | 0.1547 | 0.0145 | 0.1533 | 0.0047 | 2.0000 |
| | Nbd | 5×5 | 3×3 | 3×3 | 11×11 | 7×7 | 3×3 |
| | r_b | 1 | 0 | 2 | 0 | 0 | 5 |
| 5 | $\mu \pm \sigma$ | $-4 \times 10^{-4} \pm 0.002$ | -0.007 ± 0.1 | $-2 \times 10^{-4} \pm 0.002$ | -0.02 ± 0.1 | $-5 \times 10^{-4} \pm 7 \times 10^{-4}$ | 0.5 ± 0.7 |
| | MaxErr | 0.0032 | 0.1276 | 0.0031 | 0.1677 | 0.0014 | 1.9000 |
| | Nbd | 9×9 | 3×3 | 3×3 | 11×11 | 11×11 | 3×3 |
| | r_b | 5 | 0 | 5 | 0 | 0 | 5 |

Según estos valores, los mejores resultados para todos los métodos proporcionan errores medios por debajo de 0.008 px con una desviación estándar por debajo de 0.007 px. Además, el valor del error máximo para el mejor resultado de cada función de ajuste (el cual se puede tomar como medida del efecto de atrapamiento de pico al ser la mayor diferencia entre la posición real y la detectada del objeto de interés) es al menos un orden de magnitud menor que el movimiento de la secuencia. Esto significa que el seguimiento de las muestras es muy bueno, siempre y cuando los parámetros de ajuste estén bien seleccionados.

En la figura 3.8 se muestra el error obtenido al realizar la correlación cruzada como la diferencia entre el movimiento medido y el real que conocemos que tiene la secuencia. Se ha representado este error para el movimiento horizontal con todos los entornos alrededor del pico de correlación evaluados y seleccionando un desenfoque gaussiano de 2 en todos los casos. Con esta representación gráfica se puede observar claramente el efecto de atrapamiento de pico mencionado anteriormente que es común cuando se utiliza la correlación cruzada junto con un ajuste del pico de correlación obtenido. Para facilitar la comparación entre gráficas, se han fijado los ejes por tipo de ajuste del pico de correlación, aunque no ha sido posible fijar los ejes de todas las gráficas ya que se hubiera perdido mucha información al proporcionar los ajustes gaussiano y splines errores más pequeños que el ajuste polinómico. Asimismo, también ha sido necesario eliminar los valores que proporcionaban unos errores demasiado altos, tratándose de todos los resultados obtenidos con entornos del pico de correlación pequeños (Nbd = 3×3) para los ajustes splines y gaussiano y, en el caso del ajuste gaussiano, también los resultados con un entorno de Nbd = 5×5 para las secuencias de 3 a 5 y los del entorno Nbd = 11×11 para la secuencia 4.

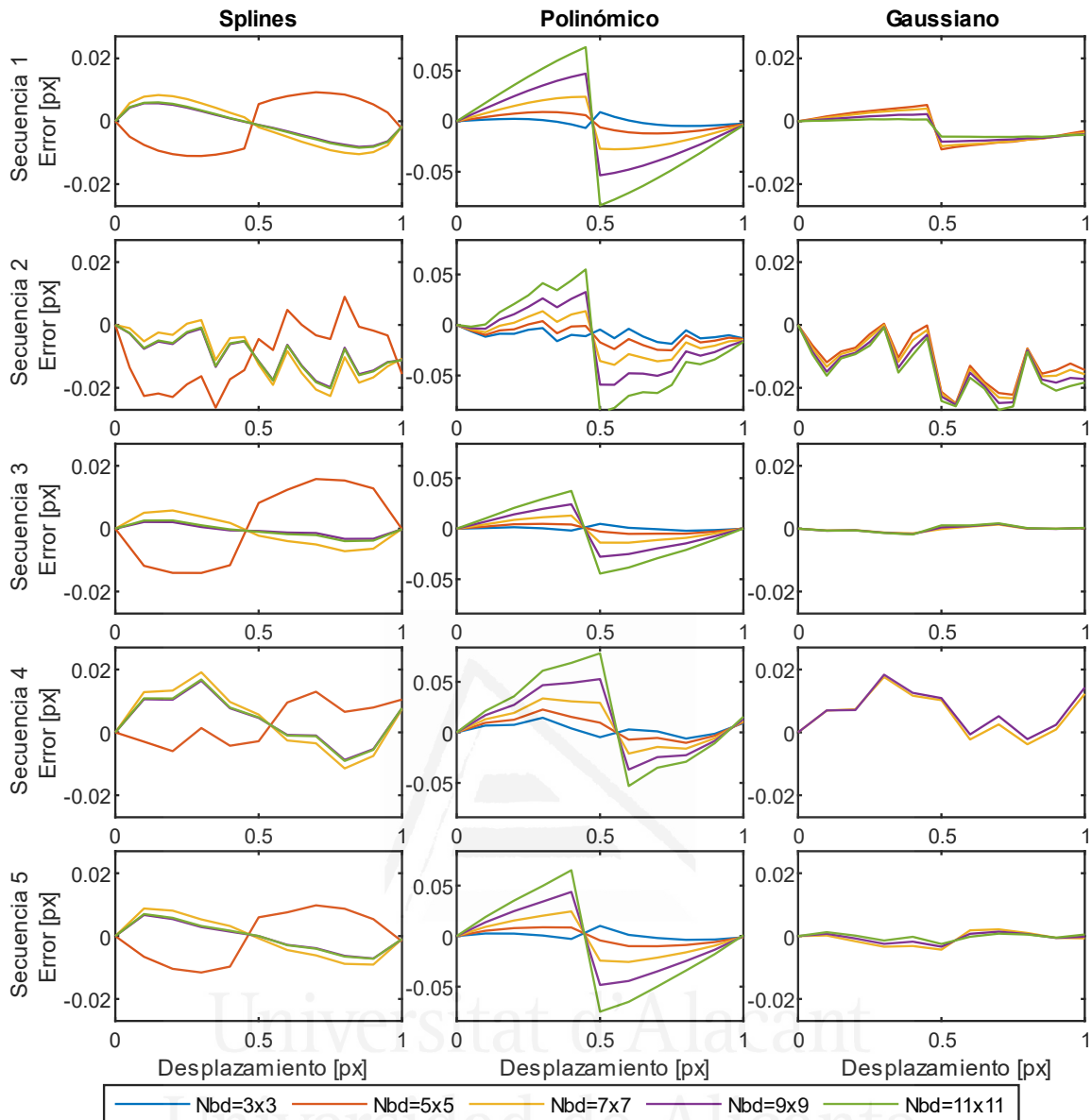


Figura 3.8. Error para un desenfoque gaussiano $r_b = 2$ cambiando el entorno del pico de correlación analizado (Nbd) para las 5 secuencias mostradas en la tabla 3.2 y con los tres ajustes estudiados. Para los ajustes splines y gaussiano se han eliminado de la gráfica los tamaños Nbd = 3×3 y para el ajuste gaussiano, también se han eliminado los resultados de Nbd = 5×5 para las secuencias 3 a 5 y los de Nbd = 11×11 para la secuencia 4 para mejorar la visualización de los resultados.

Se puede observar que, tanto para el ajuste gaussiano como con splines, el error no es sensible al tamaño del entorno escogido siempre que éste sea mayor de 5×5, siendo los errores similares en todos los casos. Sin embargo, en el caso de la interpolación polinómica se aprecia claramente que con un entorno más reducido el error es menor.

En la figura 3.9 se han representado los errores máximos obtenidos por tipo de ajuste del pico de correlación y secuencia para poder visualizar todos los resultados de un vistazo y determinar fácilmente las combinaciones de r_b y Nbd que proporcionan los mejores y peores resultados. De nuevo, se han fijado los ejes de las gráficas por ajuste del pico de correlación para poder compararlas entre sí, pero no ha sido posible fijar los ejes de todas

las gráficas ya que se hubieran perdido demasiados datos en la representación debido a que el ajuste gaussiano proporciona errores más pequeños que los otros dos ajustes. Además, para visualizar mejor los resultados, ha sido necesario eliminar ciertos valores que proporcionaban errores demasiado altos. En el caso del ajuste splines, se han eliminado todos los resultados obtenidos con entornos del pico de correlación pequeños ($Nbd = 3 \times 3$) y en el caso del ajuste gaussiano, se han eliminado los resultados con entornos pequeños ($Nbd = 3 \times 3$) para todas las secuencias y los de entornos de $Nbd = 5 \times 5$ para las secuencias de 3 a 5.

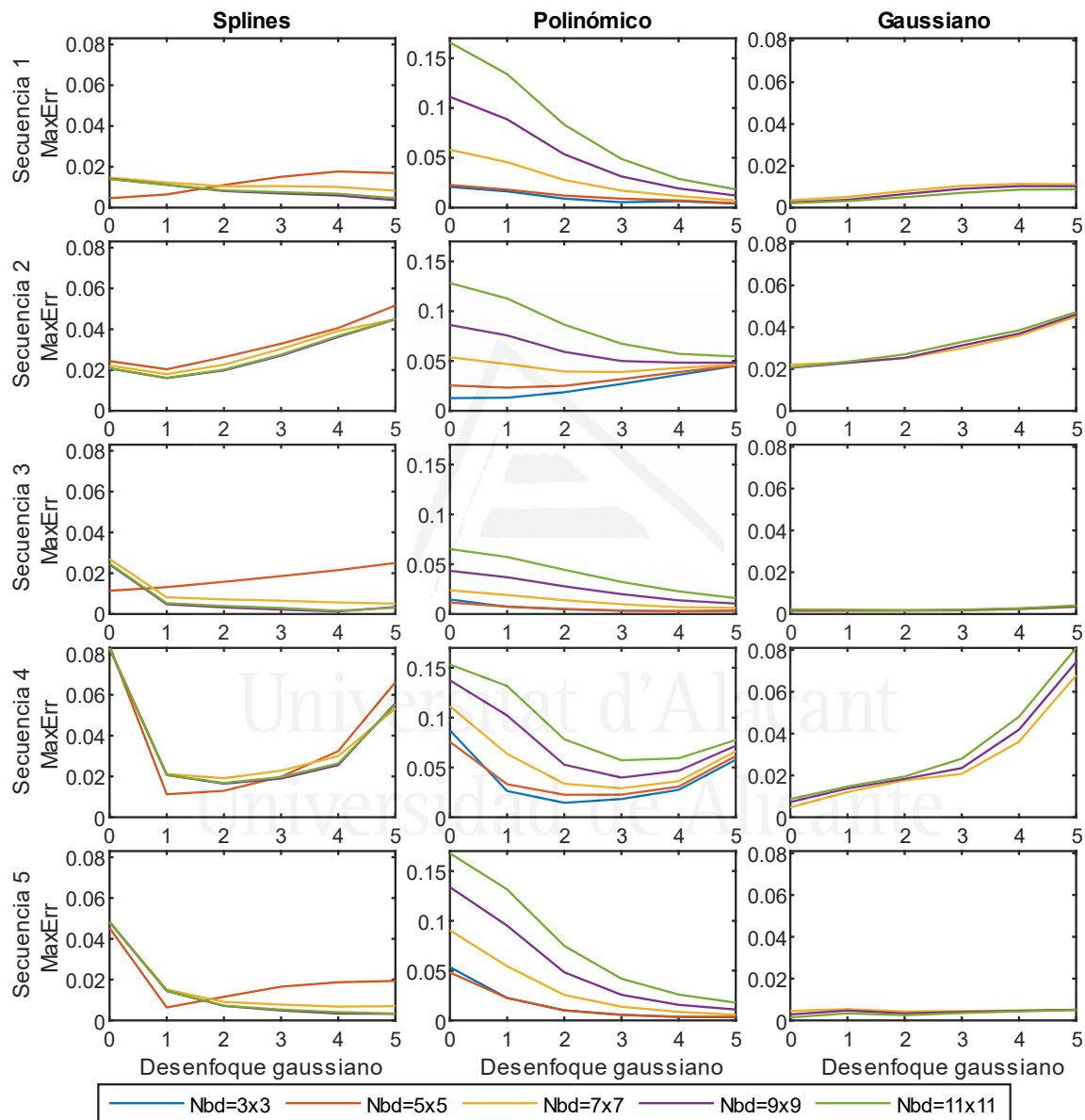


Figura 3.9. Error máximo para los distintos niveles de desenfoco gaussiano (r_b) cambiando el entorno del pico de correlación analizado (Nbd) para las cinco secuencias mostradas en la tabla 3.2 y con los tres ajustes estudiados. Para los ajustes splines y gaussiano se han eliminado de la gráfica los tamaños $Nbd = 3 \times 3$ y en las secuencias 3 a 5 también se han eliminado los resultados de $Nbd = 5 \times 5$ en el ajuste gaussiano para mejorar la visualización de los resultados.

En esta gráfica, el análisis respecto al tamaño del entorno de ajuste alrededor del pico es similar al de la gráfica anterior. En cuanto a la influencia del desenfoco, vemos que éste

sí tiene cierto efecto en el caso del ajuste por un polinomio cuadrático pero en el resto de casos, no es demasiado relevante, produciéndose una ligera mejora de los resultados con splines para desenfoques pequeños.

Los resultados completos se pueden consultar en la Publicación 1 del Anexo IV. Si bien, a modo de resumen podemos decir que para el ajuste polinómico de segundo grado es importante utilizar una matriz de ajuste muy pequeña (solamente de 3×3 px) para obtener un ajuste óptimo independientemente de la muestra evaluada y, para los ajustes splines y gaussiano, este número de píxeles variará en función de las características del objeto, pero normalmente se obtienen mejores resultados con valores medios-altos (entre 7×7 y 11×11 px). Por otra parte, el desenfoque gaussiano ha mejorado significativamente los resultados del ajuste polinómico, pero no ha tenido una tendencia clara en los resultados de las otras dos funciones estudiadas.

Finalmente, para comprender fácilmente la adaptabilidad del tipo de función de correlación al pico, lo que hace que para unos ajustes sea mejor utilizar un entorno grande y para otros un entorno pequeño, se ha representado en la figura 3.10 el pico de correlación y los perfiles de las tres funciones de ajuste correspondientes a un desplazamiento de 0.4 px para la secuencia 1 sin desenfocar y utilizando unos tamaños de matriz de ajuste de 3×3 y 11×11 px. Aunque estas curvas pueden variar mucho de un fotograma a otro y para diferentes secuencias, sirven para ilustrar el problema que se pretende explicar.

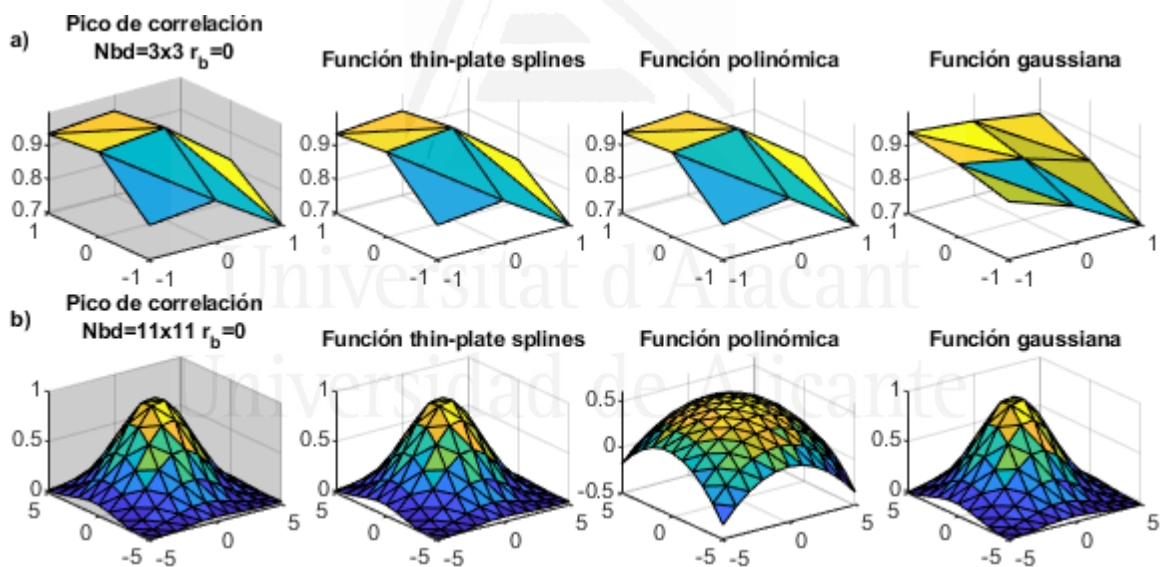


Figura 3.10. Pico de correlación y superficies de las funciones de ajuste especificadas en la gráfica para la secuencia 1 con un desplazamiento de 0.4 px, sin desenfocar y con vecindarios de **a)** 3×3 y **b)** 11×11 px.

Como se puede observar, al utilizar un tamaño de matriz pequeña, la función polinómica de segundo grado se ajusta bastante bien al perfil del pico mientras que la gaussiana no puede adaptarse a la forma del área local alrededor del máximo. En el caso del tamaño de matriz de ajuste grande, la función cuadrática solo reproduce una cúpula paraboidal mientras que la gaussiana puede reproducir el cambio de curvatura que sucede en el pico. Finalmente, la función splines puede adaptarse a ambas situaciones, por lo que proporciona unos resultados aceptables independientemente del tamaño de la matriz de ajuste utilizada. Se puede ver

claramente en la tabla 3.3 esta tendencia en cuanto al número de vecinos alrededor del pico de correlación para las funciones polinómica y gaussiana.



Universitat d'Alacant
Universidad de Alicante

CAPÍTULO 4: Aplicaciones del método subpíxel por ajuste del pico de correlación

Como se ha comentado previamente, el rango de aplicaciones de las técnicas subpíxel es amplio y los campos llegan a ser muy variados, incluyendo la construcción, el diseño industrial, el sonido e incluso la ergonomía y la salud.

En este capítulo se van a explicar dos aplicaciones prácticas desarrolladas utilizando el método subpíxel por ajuste del pico de correlación aplicadas al estudio de materiales. En el primer caso (Publicación 2 del Anexo IV, (Ferrer et al., 2021)), se ha medido la expansión por hidratación de una roca arenisca al sumergirse parcialmente en agua mientras que en el segundo caso, se ha evaluado la idoneidad de la textura superficial de varias probetas de hormigón (fabricadas con distintas técnicas para cubrir todas las posibles texturas) cuando se varía la iluminación y el desenfoque en condiciones de laboratorio.

En ambos casos, se toma como hipótesis inicial que las texturas propias tanto de la roca arenisca como del hormigón se pueden utilizar como texturas de tipo pseudospeckle debido a su gran similitud. Para el caso de la roca arenisca, su composición de cuarzo y feldespato forma una matriz moteada con diferentes intensidades en su superficie que simulan un patrón de speckle (figura 4.1). Por otra parte, en la superficie del hormigón suelen aparecer pequeños agujeros debidos a las burbujas que quedan pegadas al encofrado al poner el hormigón en obra, así como irregularidades de fabricación, de forma que se genera una textura en la superficie del hormigón en la que el fondo tiene una intensidad bastante uniforme y hay zonas aisladas con intensidades muy diferentes a las del fondo (figura 4.2). Estas características son localmente similares a las observadas en el patrón de speckle diferenciándose en la densidad y espaciado entre puntos.

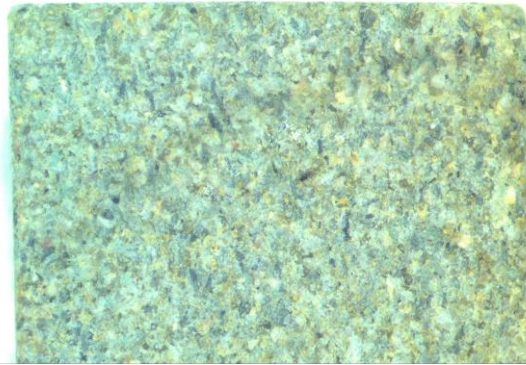


Figura 4.1. Imagen de la roca arenisca donde se puede apreciar su similitud con un patrón de speckle.

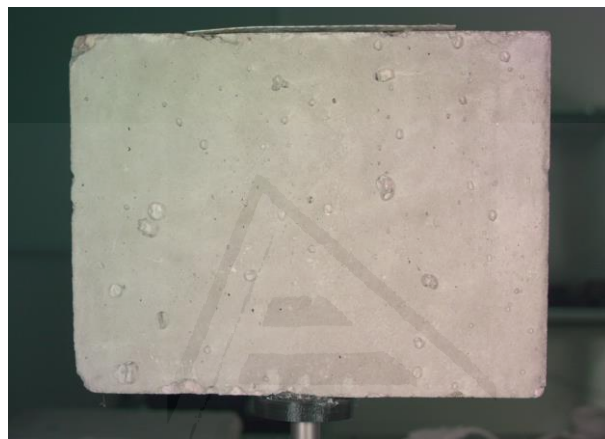


Figura 4.2. Imagen de una probeta de hormigón en donde se pueden apreciar los defectos locales en su superficie.

Si esta hipótesis es correcta, se podrían realizar los cálculos de desplazamientos de la superficie de interés utilizando la correlación cruzada y sin tener que emplear dianas artificiales, siempre y cuando los errores obtenidos sean asumibles. Por ello, se han estudiado y mejorado los métodos existentes con el fin de superar los inconvenientes que puedan presentarse al tratar de utilizar estos tipos de texturas como dianas con la correlación cruzada y ajuste del pico para detectar sus movimientos.

A continuación, se explican ambas aplicaciones detalladamente y por separado.

4.1. Medida de la expansión hídrica de una roca arenisca

En este primer caso (Publicación 2 del Anexo IV, (Ferrer et al., 2021)), se pretendía medir la expansión por hidratación de una roca arenisca con arcillas expansivas entre sus componentes al estar parcialmente sumergida en agua utilizando la propia textura de la roca como diana.

Algunas rocas areniscas incrementan su volumen al sumergirse en agua, incluso aunque solo una parte de ellas esté sumergida, ya que el agua avanza por las zonas no sumergidas por capilaridad e hidrata los componentes expansivos de la roca, provocando la expansión de toda la muestra. En una construcción realizada con un material expansivo, este proceso produce aumentos de volumen y desplazamientos en unas partes de la construcción que son incompatibles con los movimientos admisibles, tanto para la estructura como para los acabados, por lo que es importante medir este parámetro en el campo de la restauración de edificios históricos (Di Benedetto et al., 2015), por ejemplo, para evaluar el estado de deterioro de las construcciones en zonas con variaciones de humedad importantes durante los ciclos diarios o anuales o en zonas inundables.

El método más utilizado para medir esta expansión consiste en sumergir parcialmente en agua una muestra de la roca y medir su desplazamiento vertical con un transformador de desplazamiento lineal (LVDT, por sus siglas en inglés) que es un sensor de posición lineal (Siegesmund and Dürrast, 2011). El problema que presenta este sistema de medida es que es capaz de medir la expansión vertical, pero para medirla horizontalmente utilizando el sensor LVDT se necesitarían montajes complejos que no siempre son apropiados para el estudio en cuestión, además de requerir muestras de gran tamaño. Por tanto, el objetivo de realizar este trabajo utilizando métodos basados en imagen es superar los inconvenientes de las medidas que presenta el sensor LVDT.

En nuestro caso, para comprobar el uso de técnicas de imagen para monitorizar la expansión de la roca, se han utilizado 4 muestras de la misma roca arenisca compuesta principalmente por cuarzo y feldespato con unos tamaños de $11 \times 11 \times 30 \text{ mm}^3$. Las muestras se situaron en un recipiente y se llenó de agua hasta cubrir aproximadamente un tercio de la altura de la roca, tal y como se puede observar en la figura 4.3. A una distancia de 176 mm se colocó una cámara BASLER acA4600-10 uc. Se trata de una cámara a color con sensor tipo CMOS cuyo tamaño es de $6.5 \times 4.6 \text{ mm}$ y tiene una resolución máxima de $4608 \times 3288 \text{ px}$. El tamaño de cada píxel es de $1.4 \times 1.4 \text{ }\mu\text{m}$, la velocidad máxima a la que puede grabar son 10 fps y la profundidad de color son 12 bits. Se puede consultar más información sobre las características y uso de las cámaras y en los Anexos I y II.



Figura 4.3. Montaje experimental para realizar las medidas de la expansión de las rocas areniscas en el laboratorio.

Como se puede observar con las características mencionadas de la cámara utilizada, la velocidad de grabación máxima de esta cámara no es muy alta (solo 10 fps). No obstante, para este estudio no es necesario obtener una gran resolución temporal ya que la expansión de la roca al sumergirse parcialmente en agua va a ser un proceso lento que implicará pequeñas deformaciones, por lo que la resolución espacial necesariamente deberá de ser alta para detectar ese pequeño movimiento. De hecho, la velocidad de grabación utilizada ha sido un fotograma cada 2 minutos durante 3 horas.

Para este experimento se ha utilizado un objetivo telecéntrico VTL0513 de la marca *μ-tron* (figura 4.4) que funciona junto con una lente adicional VTL05FC que puede proporcionar aumentos entre $0.25\times$ y $1.0\times$, utilizándose en este caso con $0.5\times$. Las ventajas de utilizar un sistema telecéntrico son, principalmente:

- Que tiene muy poca distorsión, por lo que las imágenes apenas estarán deformadas por los bordes.
- Que la profundidad de foco es muy pequeña, así que, en cuanto a la distancia en la línea focal, solo unas pocas micras estarán enfocadas y será donde se deba situar el objeto de interés. Esto permite obtener la escala de la imagen (relación digital/real) utilizando tests calibrados tipo USAF (“1951 USAF resolution test chart,” 2021), ya que en la imagen de la roca no existirá ninguna distancia conocida que sirva de referencia para obtener esta escala a no ser que se utilice la propia dimensión de la probeta, en cuyo caso se estaría perdiendo precisión dadas las irregularidades propias de un corte en roca arenisca. La ventaja de un sistema telecéntrico es que, una vez realizada esta escala con el test calibrado tipo USAF, el test puede ser reemplazado por el objeto a medir sin variar la distancia al objeto, únicamente teniendo en cuenta el enfoque de la nueva imagen.
- Que el factor de aumento es constante debido a que la pupila de entrada está virtualmente en el infinito (“Introducción a la óptica telecéntrica,” 2018).



Figura 4.4. Objetivo telecéntrico VTL0513.

El sistema óptico de la cámara y el objetivo se han centrado con la parte superior de la muestra evitando grabar parte del recipiente contenedor del espécimen y visualizando toda la parte superior de la roca incluyendo los bordes laterales. Con una distancia de 176 mm desde la primera lente del objetivo hasta el objeto, la relación digital/real es de $0.36 \text{ px}/\mu\text{m}$. Para obtener una iluminación uniforme y no dependiente de la iluminación cambiante del medio ambiente, el experimento se ha instalado sobre una mesa mecanizada y dentro de una cabina de iluminación, tal y como se muestra en la figura 4.5. Se puede consultar más información sobre los sistemas de iluminación y las cabinas de iluminación en el Anexo III.

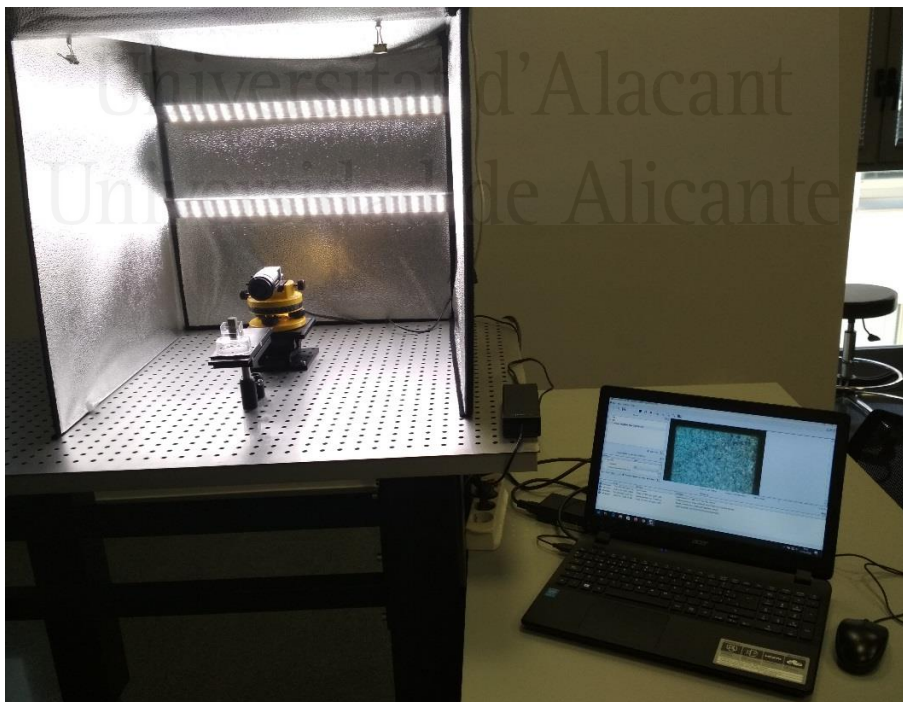


Figura 4.5. Montaje experimental sobre una mesa mecanizada y dentro de una cabina de iluminación.

Dado que sólo parte de la roca está sumergida, es de esperar que las deformaciones no sean homogéneas, por lo que se deberán analizar regiones independientes entre sí. Éstas deben ser lo suficientemente pequeñas como para asegurar que la deformación en su interior sea lo más uniforme posible, pero a su vez, lo más grandes posible para que la textura analizada sea representativa de la roca y equivalente en todas las regiones que se comparan.

Un modo de obtener el tamaño medio del detalle en una escena es calcular el tamaño del pico de autocorrelación (Russ, 2011). Así, para saber qué tamaño de ROI (región de interés) tomar para obtener un buen análisis de la deformación, se ha ido analizando la autocorrelación con ROIs de tamaños entre 5×5 y 200×200 px posicionadas aleatoriamente en 50 lugares diferentes de toda la superficie de la muestra. Para estimar el tamaño medio de las partículas en cada ROI, se ha calculado la altura a media anchura del pico de autocorrelación (conocido por sus siglas en inglés como FWHM). Como se puede apreciar en la figura 4.6, existe una estabilidad en los tamaños con ROIs entre 80×80 y 120×120 px. Por ello, se han seleccionado 6 ROIs de 100×100 px en las posiciones mostradas en la figura 4.7 para cada una de las cuatro muestras de la roca. Para el seguimiento de la deformación tomaremos cada ROI de la primera imagen de la secuencia como referencia y se comparará con las respectivas ROIs obtenidas a lo largo del tiempo. En las imágenes de la secuencia se ha reducido cada ROI 20 px a cada lado para evitar así que la imagen desplazada se mueva por fuera de la imagen de referencia al realizar la correlación cruzada.

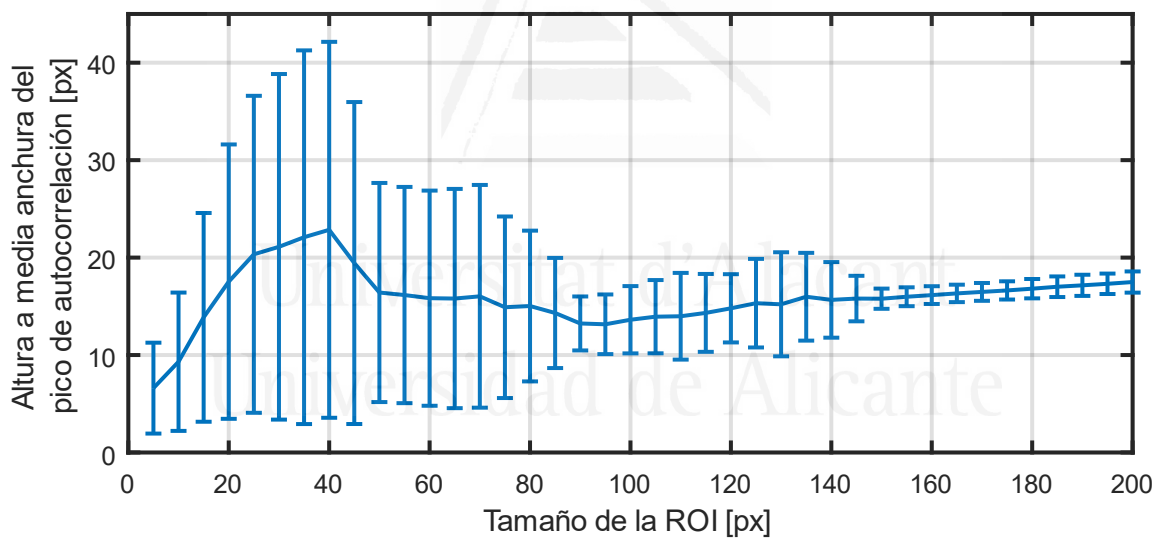


Figura 4.6. Altura a media anchura del pico de autocorrelación para estimar el tamaño medio de las partículas en cada ROI y así seleccionar los tamaños de ROIs adecuados durante el análisis experimental de las rocas.

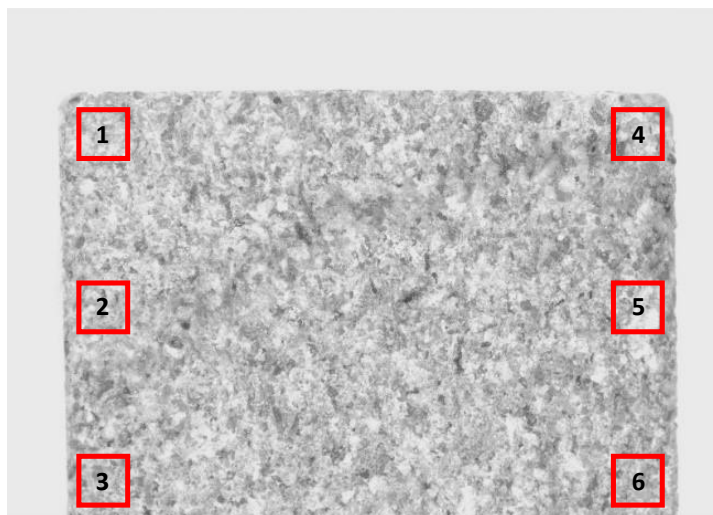


Figura 4.7. Tamaño y posición de las 6 ROIs analizadas para cada secuencia de las 4 muestras de la roca arenisca.

Para estudiar los parámetros óptimos para el seguimiento de la deformación, es decir, el método de ajuste del pico, el entorno y el uso de desenfoque, se ha realizado una simulación numérica, generando un desplazamiento progresivo vertical de 0.1 px hasta llegar a 10 px a partir de cada ROI de la primera imagen de la secuencia grabada para la muestra 1 de la roca. Este desplazamiento se ha generado utilizando la propiedad de traslación de la Transformada de Fourier (Gaskill, 1978). A continuación, se ha realizado la correlación cruzada con ajuste del pico mediante las funciones polinómica de segundo grado con un entorno de 3×3 px y la gaussiana con un entorno de 11×11 px para las 6 ROIs mostradas en la figura 4.7. Los errores obtenidos para cada ROI y tipo de ajuste se muestran en la figura 4.8.

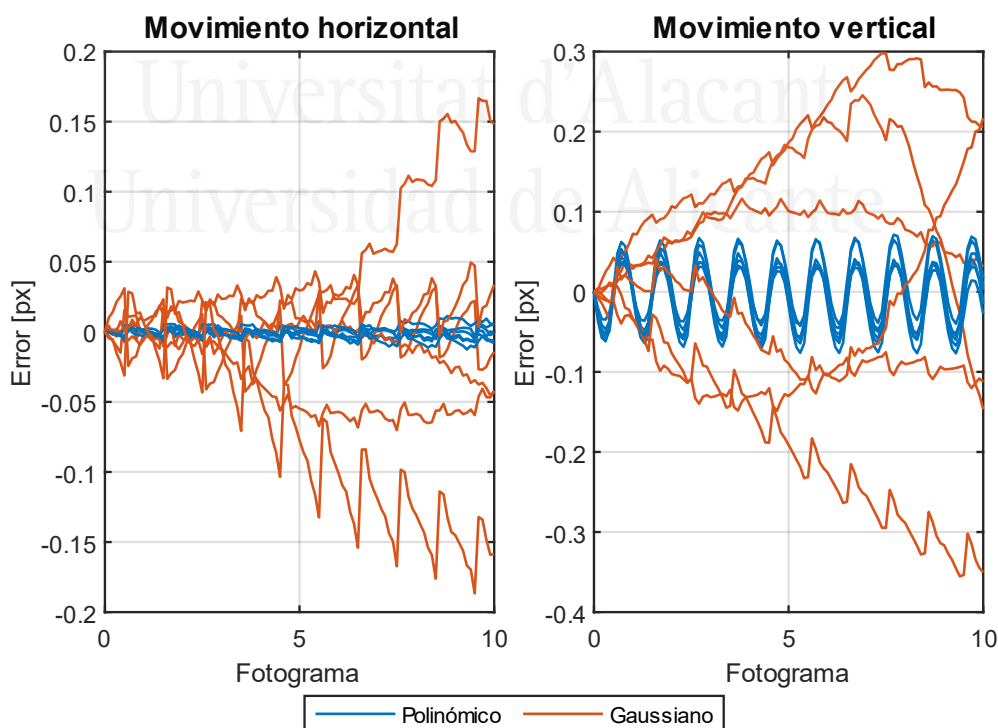


Figura 4.8. Errores horizontales y verticales producidos al realizar la correlación cruzada con los ajustes polinómico y gaussiano para un desplazamiento generado digitalmente.

Se ve que, para el movimiento horizontal, que es inexistente, el método con ajuste polinómico se ajusta bastante bien al movimiento previsto, con errores atribuibles al ruido de cálculo, mientras que el ajuste gaussiano detecta movimientos “fantasma” con errores completamente erráticos. En el caso del movimiento vertical, el comportamiento es similar. Se observa que el ajuste polinómico presenta ciclos coincidentes con cada píxel del desplazamiento, pero, en general, los errores son bajos. Por tanto, para la textura particular de la roca arenisca aquí estudiada, es mejor utilizar un ajuste polinómico con un entorno alrededor del pico de correlación pequeño, ya que proporciona unos errores menores que los del ajuste gaussiano con un entorno alrededor del pico de correlación grande. No se apreció que un ligero desenfoco afectara de modo importante al error, ya que las imágenes en sí no presentan bordes marcados.

De acuerdo con estos resultados, para estudiar la deformación de la roca arenisca al hincharse por hidratación se ha usado la correlación cruzada normalizada junto con un ajuste del pico de correlación polinómico utilizando una matriz de 3×3 px, manteniendo el pico de correlación centrado en la matriz y sin introducir ningún desenfoco. En la figura 4.9 se muestra el diagrama de flujo del programa utilizado para evaluar la correlación cruzada en distintos puntos de la roca arenisca al sumergirse parcialmente en agua.

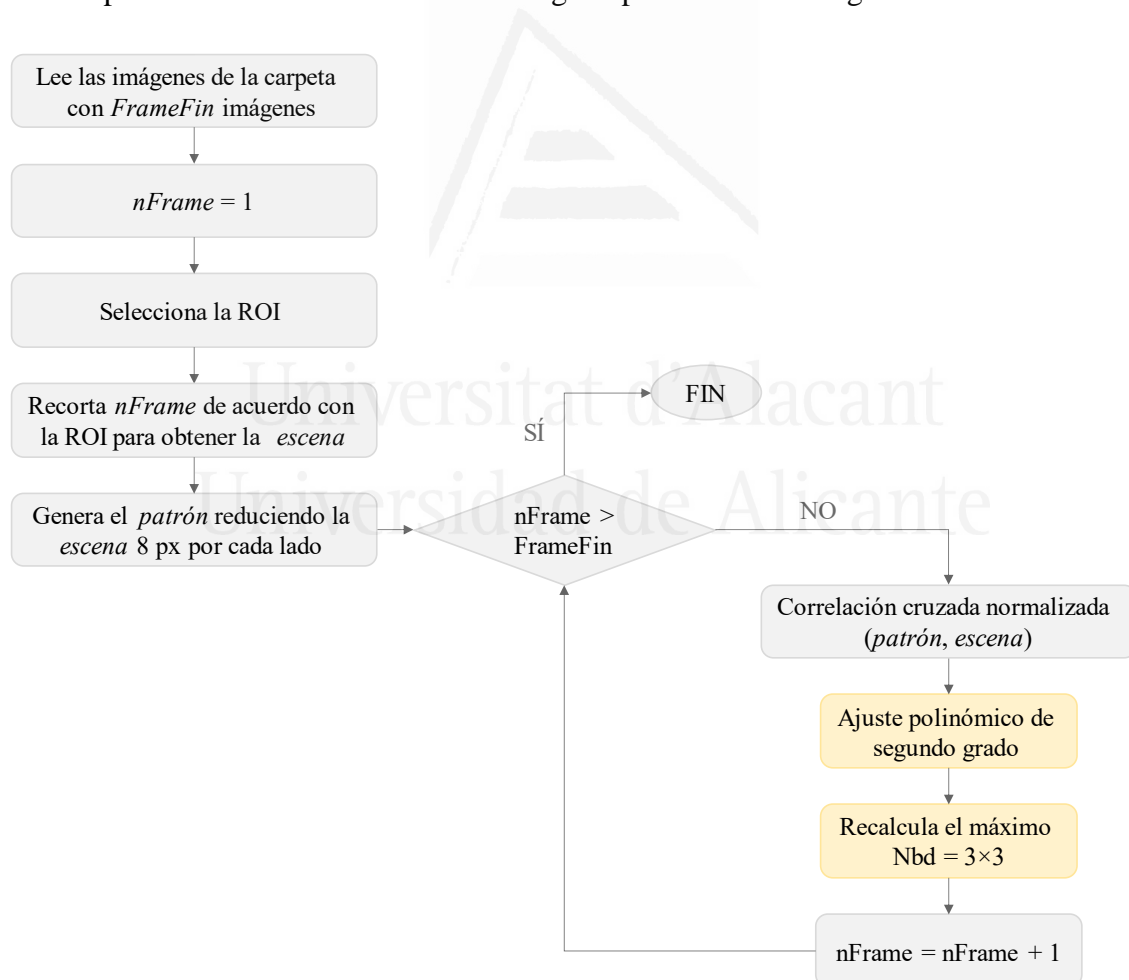


Figura 4.9. Diagrama de flujo del programa utilizado para evaluar diferentes puntos de la roca mediante la correlación cruzada.

En la figura 4.10 se muestra una gráfica con las posiciones absolutas de cada ROI en horizontal y en vertical para una medida estática en seco de la muestra 1. Para las otras tres muestras de la roca se obtienen resultados similares.

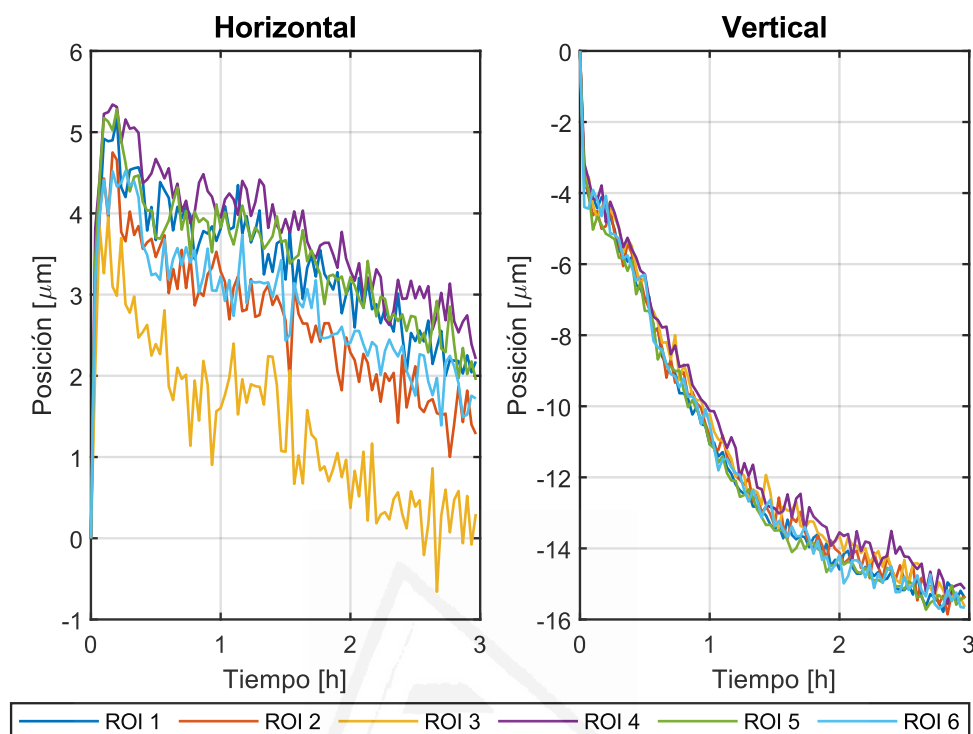


Figura 4.10. Posiciones absolutas de las 6 ROIs en horizontal y vertical para una medida estática en seco de la muestra 1.

Se puede apreciar que todas las ROIs presentan un movimiento negativo, es decir, hacia la izquierda y hacia abajo. Este movimiento sólo se puede explicar por derivas en la cámara, probablemente térmicas, por lo que se ha optado por calcular la expansión de la roca (diferencia entre posiciones absolutas de dos puntos) tanto en horizontal como en vertical para eliminar las derivas producidas por inestabilidades de la cámara o del objeto que se habían detectado. Estas expansiones se calculan como las diferencias de posiciones de las ROIs analizadas por correlación del siguiente modo:

- En horizontal:
 - ROI 1 – ROI 4
 - ROI 2 – ROI 5
 - ROI 3 – ROI 6
- En vertical:
 - ROI 1 – ROI 2
 - ROI 2 – ROI 3
 - ROI 1 – ROI 3
 - ROI 4 – ROI 5
 - ROI 5 – ROI 6
 - ROI 4 – ROI 6

Los resultados de las expansiones para la muestra 1 de la roca se muestran en la figura 4.11. En las otras tres muestras se obtienen unos resultados similares que no se presentan aquí para no mostrar resultados redundantes, pero se pueden visualizar en el artículo completo (Publicación 2 del Anexo IV, (Ferrer et al., 2021)).

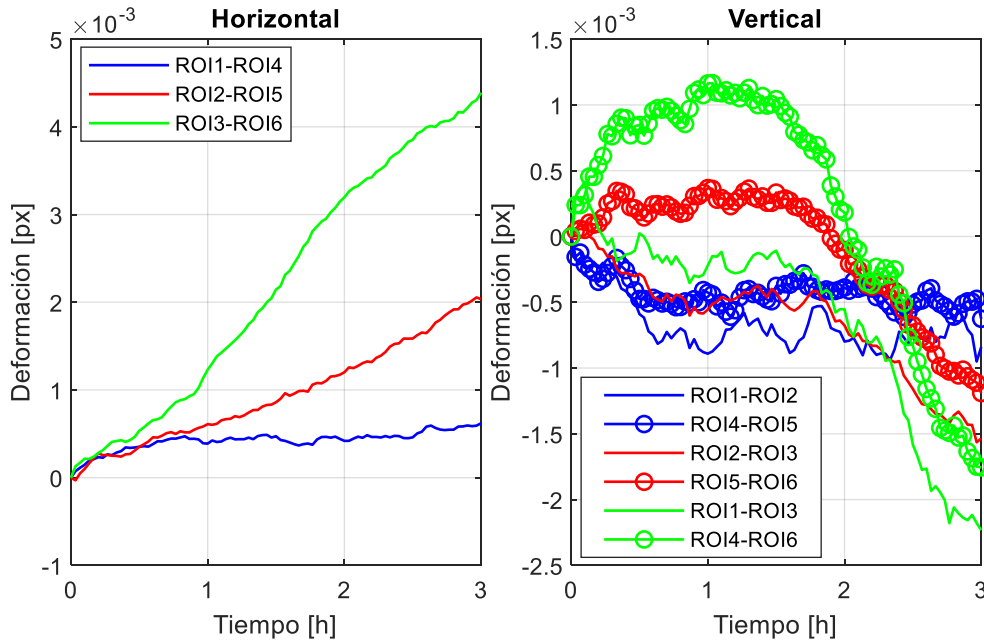


Figura 4.11. Expansiones horizontales y verticales para la muestra 1.

Se aprecia que la parte inferior de la roca (en verde en ambas gráficas) sufre una mayor expansión, lo cual era esperable ya que es la parte que está más cerca del agua. La variación de la posición de las ROIs en horizontal muestra una expansión constante de la roca, en cambio, en la dirección vertical se produce una expansión rápida al principio, una saturación y una aparente contracción al cabo de una hora y media. Este comportamiento anómalo se repite en las cuatro muestras analizadas por lo que puede deberse a un efecto de saturación o a la disolución de parte de los componentes de la roca. No obstante, comprobar el origen de este comportamiento se saldría de nuestro campo de estudio ya que no conocemos los comportamientos de las rocas areniscas y nuestro objetivo consistía en determinar si se podía medir la expansión de estas rocas utilizando imágenes y no estudiar el comportamiento de esta roca en particular.

Por otro lado, en la parte superior de la muestra (en azul en ambas gráficas), se puede ver que la expansión es menor, ya que el agua probablemente no consiga hidratar completamente los componentes más alejados de la línea del agua a partir de la que ésta sube por capilaridad.

Una vez conocidas las expansiones, es posible medir la deformación relativa producida en la roca a lo largo del tiempo dividiendo cada expansión entre la distancia inicial entre los puntos analizados, esto es, la distancia obtenida en el primer fotograma de la secuencia, de la forma en la que se muestra en la ecuación 4.1.

$$\varepsilon_{ij} = \frac{Pos_{ROIi}(t) - Pos_{ROIj}(t)}{Pos_{ROIi}(0) - Pos_{ROIj}(0)} \quad (4.1)$$

donde $Pos_{ROIi}(t)$ es la posición de la ROI i para el tiempo t y $Pos_{ROIi}(0)$ la posición de la ROI i en el fotograma inicial.

Con el fin de visualizar de modo global las deformaciones relativas en la muestra se ha procedido a un análisis local de toda la superficie de la muestra 1, representando en la figura 4.12 las deformaciones horizontales y verticales tras 1, 2 y 3 horas de haber sumergido la roca parcialmente en agua. Para ello se ha dividido la roca en ROIs de 100×100 px superpuestas 50 px con las ROIs adyacentes y se ha analizado la deformación relativa en cada una de estas zonas. Esta deformación relativa se ha obtenido restando dos ROIs alternadas en la dirección evaluada y dividiendo entre la distancia inicial entre las ROIs implicadas. Además, para evitar valores atípicos se ha aplicado un filtro de mediana de 3×3 al resultado.

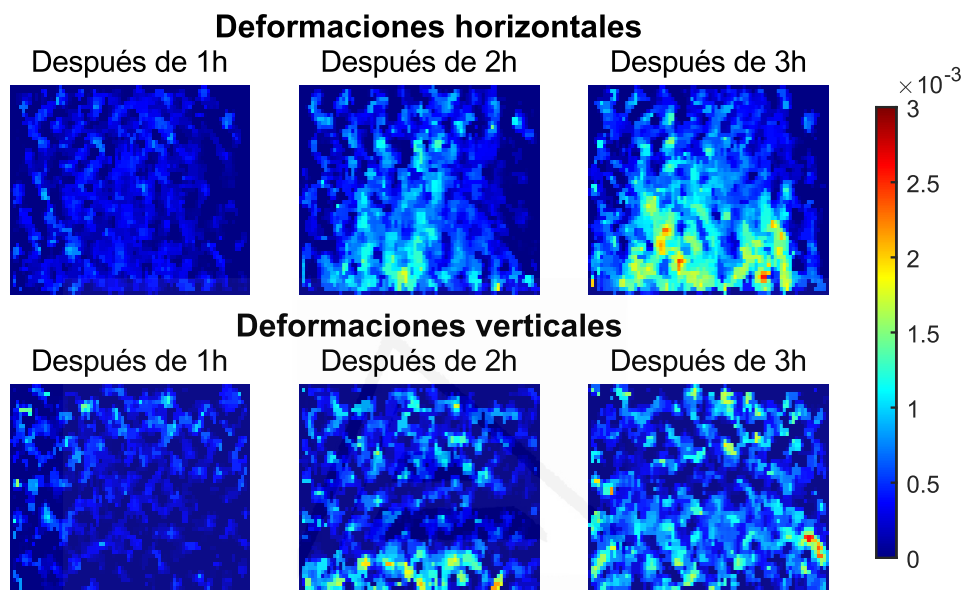


Figura 4.12. Mapas de deformaciones relativas obtenidos para la muestra 1 en horizontal y en vertical en diferentes momentos tras la inmersión de la roca en el agua.

Estos mapas de deformaciones relativas, de nuevo muestran unos valores mayores en las zonas más cercanas a la parte sumergida de la roca, resaltando la esperable mayor dilatación de la roca en estas zonas inferiores por las propias características expansivas de los componentes de la roca. Además, se pueden observar heterogeneidades en la distribución de las deformaciones en ambas direcciones, lo que evidencia una composición heterogénea de los componentes expansivos de la roca.

El error de cálculo aquí obtenido ha sido menor de 0.1 px o, equivalentemente, menor de $0.3 \mu\text{m}$. En este caso, una implementación experimental requiere secuencias muy largas, de varias horas, para poder observar el efecto completo de la hidratación de la muestra, lo que implica estabilizar de forma mecánica y térmica las condiciones para poder realizar toda la grabación con las mismas condiciones. En este caso, no se consiguió una buena estabilización de la cámara, observándose una deriva en las imágenes, aunque al evaluar magnitudes relativas como son las expansiones y las deformaciones en horizontal y en vertical y no magnitudes absolutas como sería la posición exacta de una ROI en particular de la roca, se han compensado estas inestabilidades de la cámara y otras posibles inestabilidades de la roca completa.

Como conclusión, se puede decir que este método es adecuado para medir desplazamientos locales y, por tanto, la deformación de pequeñas muestras de roca tanto en la dirección horizontal como en la vertical. El método utilizado, a pesar de las limitaciones descritas, es más simple de implementar que el LVDT y podría ser de interés para analizar los movimientos de elementos constructivos realizados a partir de piedra natural, incluso *in situ*, sin necesidad de obtener muestras de estos materiales. Así se evaluarían las expansiones producidas por los cambios de humedad que ocurren en su ubicación habitual, evitando alterar las construcciones históricas para evaluar las tensiones generadas por expansiones hídricas de la piedra natural.

Además, exceptuando el objetivo telecéntrico que podría sustituirse por un objetivo estándar con una calibración de la distorsión adecuada, el equipo que hemos utilizado es relativamente barato, lo que hace esta técnica accesible a estudiantes o tests preliminares en laboratorios de investigación. También queda demostrado que se pueden utilizar las texturas naturales de las rocas areniscas como la estudiada como diana para procesados en los que se utilicen técnicas de correlación de imagen. En los trabajos futuros sería necesario analizar más texturas naturales para comprobar que en todas ellas se obtienen resultados similares a los presentados en este trabajo. Aparte de este resumen de los resultados obtenidos, los resultados completos de este experimento se encuentran en la Publicación 2 del Anexo IV.

Como continuación de este trabajo y tratando de resolver las derivas térmicas de la cámara observadas al realizar grabaciones de larga duración, se ha abierto una línea de trabajo nueva dedicada a la búsqueda del origen de estos errores y al análisis de sus posibles soluciones sin incrementar el coste final del procedimiento y la complejidad del equipo. Para realizar esta investigación se ha utilizado un objeto binario formado por círculos negros sobre fondo blanco que ocupan todo el campo de visión de la cámara (figura 4.13).

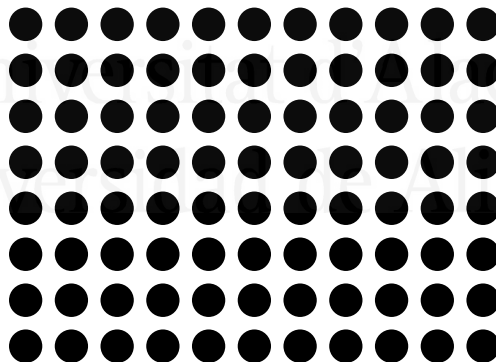


Figura 4.13. Objeto binario formado por círculos negros sobre fondo blanco utilizado para evaluar la estabilidad de la cámara al grabar un objeto estático.

Como las causas de las derivas en la cámara se supusieron térmicas, se colocó un disipador de calor sobre la cámara para ayudar a reducir su temperatura al estar en funcionamiento. Además, también se dejó la cámara en funcionamiento antes de comenzar la grabación para permitir su estabilización. Tras varias pruebas, se llegó a la conclusión de que era necesario mantener la cámara encendida y grabando durante al menos 2 horas antes de comenzar con el experimento en cuestión y utilizando un disipador para conseguir una buena estabilización térmica de la cámara. En la figura 4.14 se muestran los resultados obtenidos inicialmente (sin emplear el disipador de calor sobre la cámara ni dejarla encendida antes de comenzar con la grabación) y los obtenidos con la mejor combinación

de parámetros (utilizando el disipador y dejando la cámara encendida 2 horas antes de comenzar con la medida). En ambos casos se ha realizado una grabación de 2 horas pero tomando una imagen cada 2 minutos en el primer caso y una imagen por minuto en el segundo. El objeto binario utilizado para las grabaciones permaneció estático en todo momento con el fin de evitar que el error del propio sistema de desplazamiento incrementara el error registrado tras el procesado de las imágenes, además, sin posibilidad de cuantificar por separado el error debido a la cámara y el debido al sistema de desplazamiento. Para las dos secuencias se han medido las posiciones del círculo central y de los cuatro de las esquinas utilizando la correlación cruzada normalizada con un ajuste del pico polinómico y un vecindario de 3×3 px. Al medir estas posiciones, se está proporcionando un dato del error obtenido ya que se trata de una grabación estática.

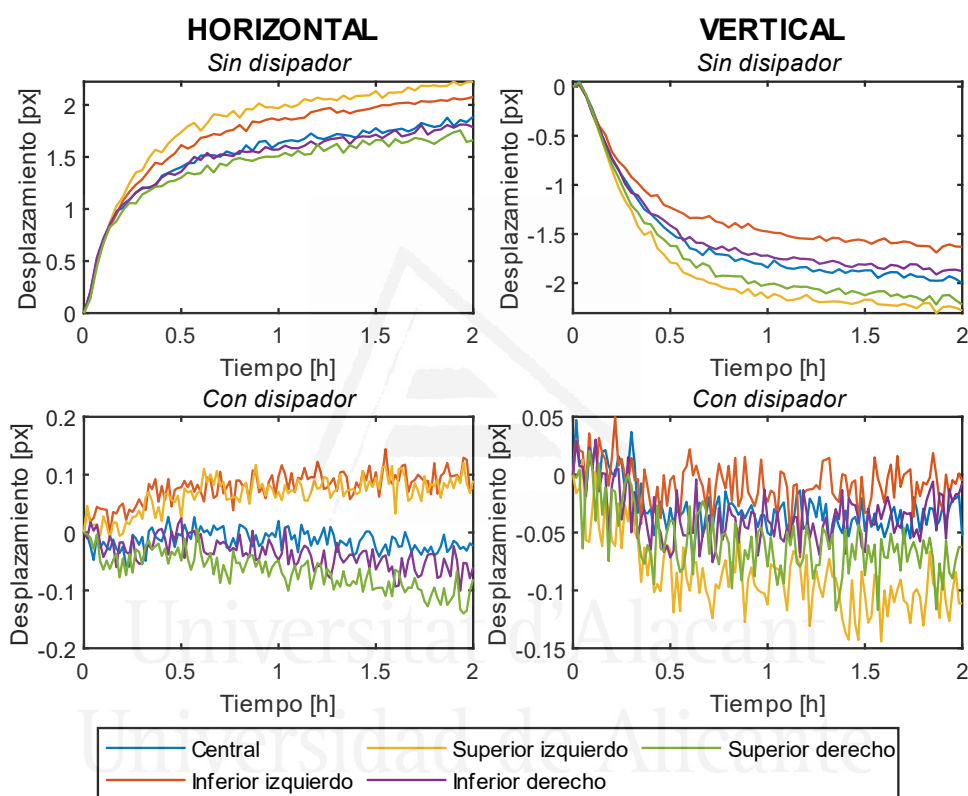


Figura 4.14. Desplazamientos medidos en horizontal y vertical para el círculo central y los cuatro de las esquinas sin utilizar disipador ni precalentar la cámara (*Sin disipador*) y utilizando un disipador además de dejar la cámara encendida 2 horas para mejorar su estabilidad (*Con disipador*).

Para proporcionar un valor más fiable del error obtenido al utilizar un disipador de calor y precalentar la cámara durante 2 horas, se ha realizado el mismo experimento tres veces y se ha calculado la media del error medio y de la desviación estándar para las tres medidas, obteniéndose un error de 0.02 ± 0.02 px. Este valor será el que se utilizará como referencia del error introducido por la cámara.

Para finalizar, con este estudio se consigue proporcionar una solución económica y sencilla de implementar para la estabilización térmica de la cámara al realizar medidas de larga duración.

4.2. Evaluación de la textura natural del hormigón como diana en medidas a través de la correlación cruzada

El objetivo de este experimento es analizar la posibilidad de utilizar la correlación cruzada para monitorizar los movimientos en superficies de hormigón utilizando su propia textura como diana natural y controlando las condiciones de iluminación y distancia en el laboratorio. Como se ha mencionado anteriormente, la textura del hormigón podría asemejarse a la del pseudospeckle, aunque con una densidad de detalle mucho más ligera. Así pues, queremos comprobar hasta qué punto esta asimilación es viable y la correlación proporciona información suficiente como para realizar un seguimiento fiable.



















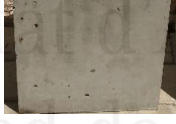


Se trata de un objetivo intermedio respecto del objetivo final que consiste en analizar la viabilidad de este procedimiento en condiciones reales, esto es, al aire libre, a gran distancia del objeto a medir y sin necesidad de acceder al punto de medida. Puesto que este objetivo final tiene muchas incertidumbres, ya que en condiciones reales no es posible controlar la iluminación, el ruido, el viento sobre el trípode donde está la cámara, la distorsión atmosférica, etc., se ha comenzado evaluando la parte que puede desarrollarse en el laboratorio para continuar en trabajos posteriores con los experimentos al aire libre.

Al encofrar el hormigón *in situ* es bastante común que quede una textura superficial como la que se ha mostrado en la figura 4.2 y que recuerda a las típicas texturas tipo speckle que se utilizan habitualmente como dianas en procedimientos estándar o de medida a través de la correlación cruzada como GOMCorrelate (“GOM Correlate Pro,” 2021) o Ncorr (Blaber et al., 2015). Hay que tener en cuenta que la presencia de burbujas en la superficie no empeora la calidad del producto, siempre que los tamaños sean pequeños y la densidad de agujeros no sea alta, solo se trata de un problema estético, motivo por el que no se pone demasiado empeño en evitar su aparición. En hormigones prefabricados con alta calidad, estos pequeños agujeros no aparecerán debido a la alta calidad de su proceso de vibración, sin embargo, su superficie nunca será completamente uniforme ya que habrá cambios en la humedad y temperatura por zonas, una distribución desigual de los agregados que lo componen o características propias del material utilizado para el encofrado, apareciendo patrones irregulares de diferentes tonos de gris que podrían servir también como textura superficial para el seguimiento de las imágenes. Asimismo, en hormigones encofrados *in situ* también pueden aparecer marcados los bordes o la textura de los paneles de encofrado.

Para comprobar si las texturas típicas del hormigón pueden servir como diana natural al utilizar la correlación cruzada junto con un ajuste del pico, se han preparado unas probetas de hormigón cúbicas de unos 20 cm de lado con texturas representativas en función del encofrado de las probetas y de los métodos de curación del hormigón para su posterior medida en el laboratorio. Para la fabricación de las probetas de hormigón imitando las texturas típicas que suele presentar en la obra, se han combinado las 3 técnicas o características del hormigón fresco que más influyen en el resultado final de su textura. Por tanto, se han utilizado moldes de madera y metálicos, se han preparado los hormigones con consistencias seca, plástica, blanda y fluida y se han utilizado como tipos de compactación del hormigón el picado con barra de acero, el golpeo con maza de goma y el vibrador

eléctrico. Como resultado, se han obtenido las 21 probetas de hormigón mostradas en la tabla 4.1, teniendo en cuenta que la combinación de hormigón de consistencia seca y encofrado de madera no se ha realizado porque resulta muy difícil de compactar y no suele utilizarse en la obra.

Tabla 4.1. Probetas realizadas en función del tipo de encofrado (E), compactación (C) y consistencia del hormigón utilizados.

| E | C | CONSISTENCIA DEL HORMIGÓN | | | |
|--------|----------|---|---|--|---|
| | | SECA | PLÁSTICA | BLANDA | FLUIDA |
| MADERA | PICA | |  |  |  |
| | MAZA | |  |  |  |
| | VIBRADOR | |  |  |  |
| ACERO | PICA |  |  |  |  |
| | MAZA |  |  |  |  |
| | VIBRADOR |  |  |  |  |

Para aligerar las probetas y facilitar así su utilización en el laboratorio, se han cortado en piezas de unos 2 cm de espesor, conservando la textura original en una de sus caras para poder ensayarla y desechando el núcleo interno de la probeta. De este modo, se aumenta también la variabilidad de las texturas para cada forma de fabricación al obtener 4 caras ensayables por cada probeta de hormigón fabricada. Se utilizan solamente las 4 caras laterales ya que en la cara superior el acabado es muy diferente por no estar en contacto con el encofrado y tendrá un acabado más rugoso. Por otra parte, en el caso de la cara inferior que sí que está en contacto con el encofrado, estará afectada por todo el peso de la probeta y tendrá un acabado más compacto y con menos agujeros, variando su textura respecto de las otras 4 caras. Además, al recortar las caras laterales, esta cara inferior quedaría bastante más pequeña que el resto, por lo que se prefiere desechar.

Finalmente, las probetas utilizadas en los ensayos tienen unos tamaños de aproximadamente $15 \times 20 \times 2 \text{ cm}^3$. En la figura 4.15 se muestran algunas de las probetas de

hormigón ensayadas en el laboratorio etiquetadas en su parte superior para facilitar su identificación.



Figura 4.15. Algunas de las muestras finales utilizadas para ensayar con su etiqueta correspondiente.

Para poder valorar si la textura del hormigón es adecuada para su utilización como diana natural es necesario evaluar los errores de localización y comprobar que no son excesivos comparados con los que se obtienen con otras texturas convencionales como las binarias. Se han tomado una serie de secuencias estáticas con las probetas de hormigón mostradas en la tabla 4.1 para no introducir errores al desplazar las muestras con un motor de desplazamiento continuo, cosa que complicaría la obtención de los errores directamente relacionados con la textura, que es nuestro objetivo. Además, se han grabado 6 secuencias de medidas diferentes para cada probeta de hormigón cambiando la iluminación y el desenfoque, evaluando así los errores producidos con iluminaciones lateral y bidireccional y con 3 niveles diferentes de desenfoque (sin desenfocar, desenfoque leve y desenfoque moderado). En todos los casos se proporciona el porcentaje de iluminación respecto del máximo que puede proporcionar la lámpara utilizada (MOMAN Led Video Light Battery: 3.7V/2800mAh) seleccionando siempre una temperatura de color de 5000 °K:

- 1) 1 foco de luz a la izquierda con un 30% de iluminación y sin desenfocar la imagen.
- 2) 2 focos de luz (uno a la derecha y otro a la izquierda) con un 25% de iluminación y sin desenfocar la imagen.
- 3) 2 focos de luz (uno a la derecha y otro a la izquierda) con un 15% de iluminación y sin desenfocar la imagen.
- 4) 2 focos de luz (uno a la derecha y otro a la izquierda) con un 5% de iluminación y sin desenfocar la imagen.
- 5) 2 focos de luz (uno a la derecha y otro a la izquierda) con un 25% de iluminación y con un desenfoque leve de la imagen.
- 6) 2 focos de luz (uno a la derecha y otro a la izquierda) con un 25% de iluminación y con un desenfoque moderado de la imagen.

Se han utilizado estas intensidades ya que, un solo foco al 30% de luz proporciona una buena iluminación lateral y al utilizar la iluminación bidireccional, un 25% de luz

proporciona la iluminación óptima para iluminar la escena a simple vista, mientras que un 5% corresponde con una iluminación más bien pobre.

Es importante evaluar el cambio en la iluminación para conocer si diferentes intensidades de luz, que producirían un contraste distinto en la imagen, pueden influir en los resultados de la correlación al utilizar la textura propia del hormigón. Además, como se sabe que un ligero desenfoco de la imagen puede mejorar la precisión de la medida al utilizar la correlación debido a que se incrementa el número de niveles de gris de la imagen, se ha evaluado también su efecto de manera experimental en este tipo de textura.

Con la evaluación teórica del desenfoco expuesta en el Capítulo 3 de esta memoria, no se puede conocer con certeza el nivel de desenfoco real que se está introduciendo, porque se introduce de forma teórica tras la grabación de la secuencia utilizando una función gaussiana con diferentes sigmas. Por ello, para medir los desenfocos utilizados en esta aplicación práctica, se ha empleado una diana circular cuasi binaria (círculo negro sobre fondo blanco impreso en papel con una impresora convencional) para la que se ha determinado el número de píxeles necesarios para pasar de negro a blanco en condiciones de enfoque máximo. Si la diana fuera completamente binaria (blanco y negro puros), este salto se produciría únicamente con 1 px. Sin embargo, al ser una diana impresa en papel, los colores no son puros y eso explica que hayan sido necesarios 8 px para dar ese salto de negro a blanco, incluso estando completamente enfocada. Para el desenfoco leve, se ha seleccionado un salto de 18 px mientras que para el desenfoco moderado, el salto ha sido de 28 px. En la figura 4.16 se muestran las gráficas de estos 3 tipos de desenfoco junto con la diana con la que han sido generados y con una muestra de hormigón grabada en las mismas condiciones para poder visualizar el nivel de desenfoco introducido en una imagen real.

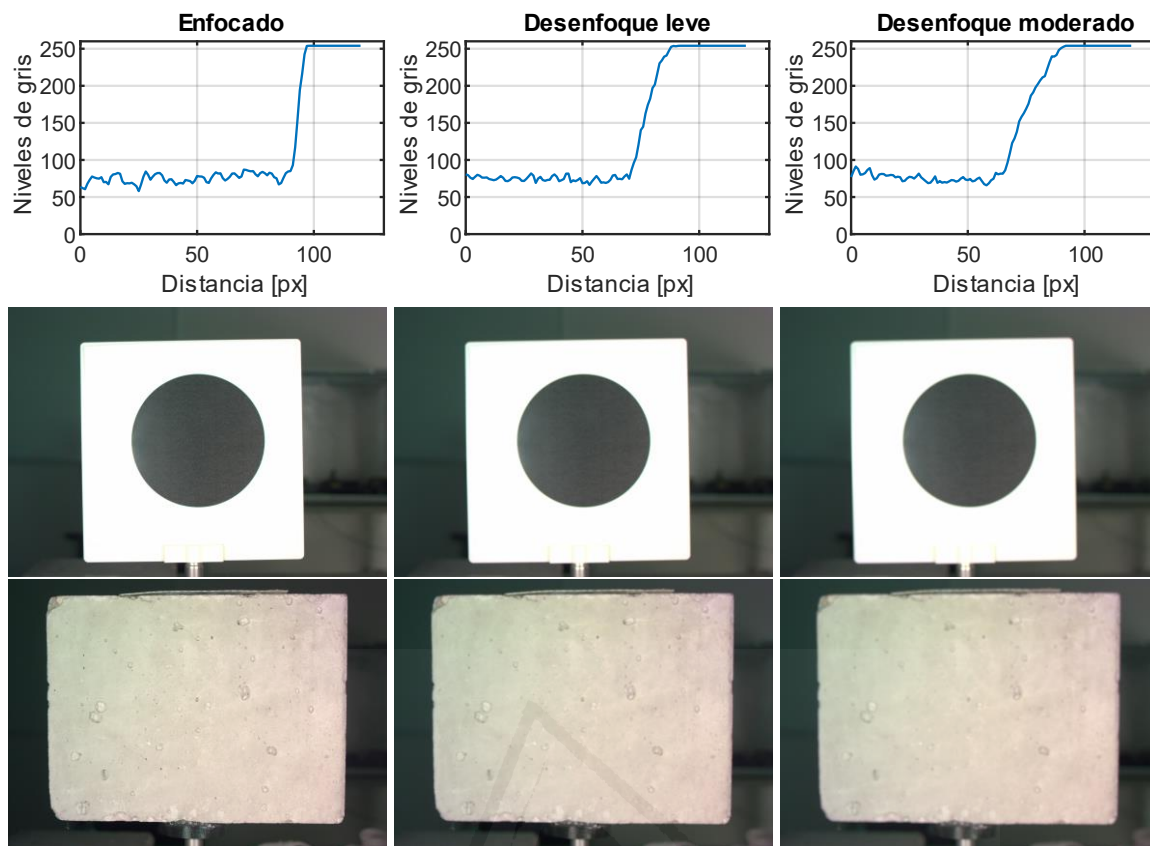


Figura 4.16. Gráficas de los tres tipos de enfoque junto con las dianas circulares cuasi binarias que los han producido y una probeta de hormigón grabada con los mismos enfoques.

Para monitorizar la posición de la textura de hormigón, se ha utilizado la correlación cruzada normalizada junto con un ajuste gaussiano del pico de correlación utilizando una matriz de píxeles alrededor del pico de 11×11 px y sin inducir desenfoco en el programa utilizado para realizar la correlación, ya que la evaluación del desenfoco se realiza introduciéndolo durante la grabación de las secuencias de imágenes. En la figura 4.17 se observa el montaje realizado para la grabación de estas secuencias de vídeo. A cada lado de la cámara hay una lámpara para iluminar la probeta de hormigón con iluminación lateral o bidireccional y según las diferentes intensidades mencionadas anteriormente.

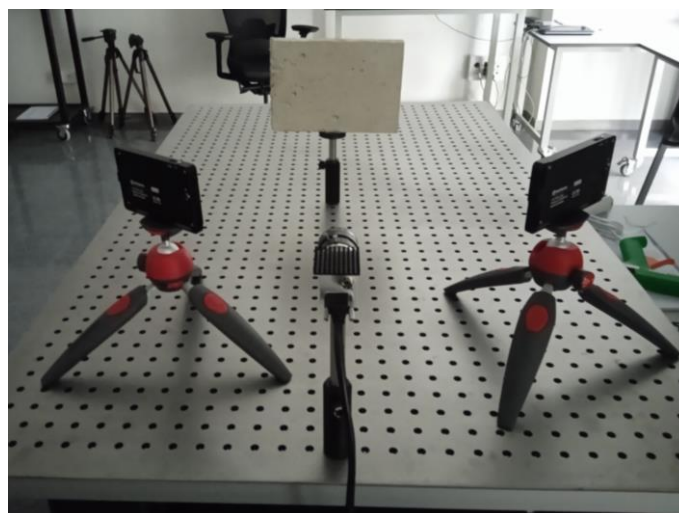


Figura 4.17. Montaje realizado para la grabación de las secuencias.

La cámara utilizada en este experimento ha sido una BASLER acA4600-10 uc, que es la misma cámara empleada para el experimento de la roca arenisca y para la calibración de la cámara utilizando el objeto compuesto por círculos negros sobre fondo blanco. Se puede consultar en los Anexos I y II más información sobre las características y uso de las cámaras y los objetivos y en el Anexo III sobre los sistemas de iluminación.

Como se ha visto que en medidas de larga duración aparecían derivas en las imágenes debidas al calentamiento de la cámara y que se pueden minimizar utilizando un disipador de calor sobre la cámara y grabando durante 2 horas antes de comenzar con el experimento en cuestión, en este caso, se ha grabado un fotograma por minuto durante 3 horas, desechando las imágenes de las 2 primeras horas y analizando solamente los últimos 60 minutos.

El objetivo empleado en este experimento y que se muestra en la figura 4.18 junto con la cámara, ha sido el BASLER C-Mount 1:1.6/12mm 2/3". En la misma figura, se observa al fondo uno de los focos de luz utilizados. El objeto está situado a una distancia de 475 mm desde la primera lente del objetivo, proporcionando una relación digital/real de 18.28 px/mm.

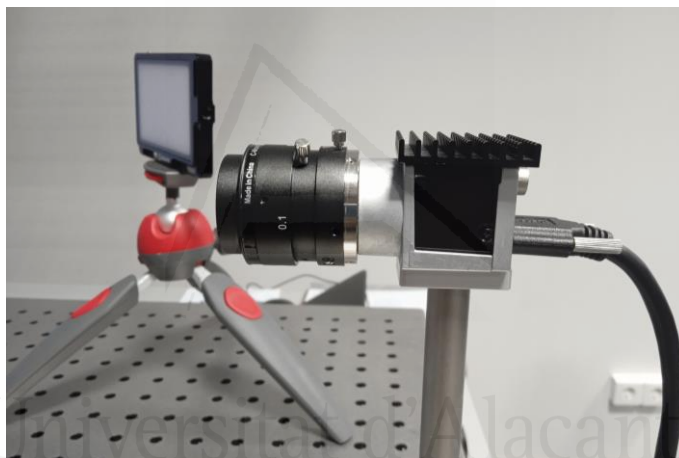


Figura 4.18. Cámara junto con el objetivo utilizados para medir las secuencias de hormigón.

En la figura 4.19 se muestra el diagrama de flujo del programa empleado para evaluar si la textura propia del hormigón es adecuada para ser utilizada como diana natural al utilizar la correlación cruzada normalizada. La ROI (región de interés) seleccionada para cada probeta de hormigón ha sido aquella que proporciona la mayor parte de hormigón posible sin incluir los bordes de la probeta. Además, para las 6 secuencias de cada probeta de hormigón se ha evaluado la misma ROI. De este modo, se pueden comparar los resultados obtenidos, ya que los únicos cambios están producidos por la distinta iluminación y enfoque, y no por cambios en la zona de la probeta evaluada que podrían contener distintas texturas.

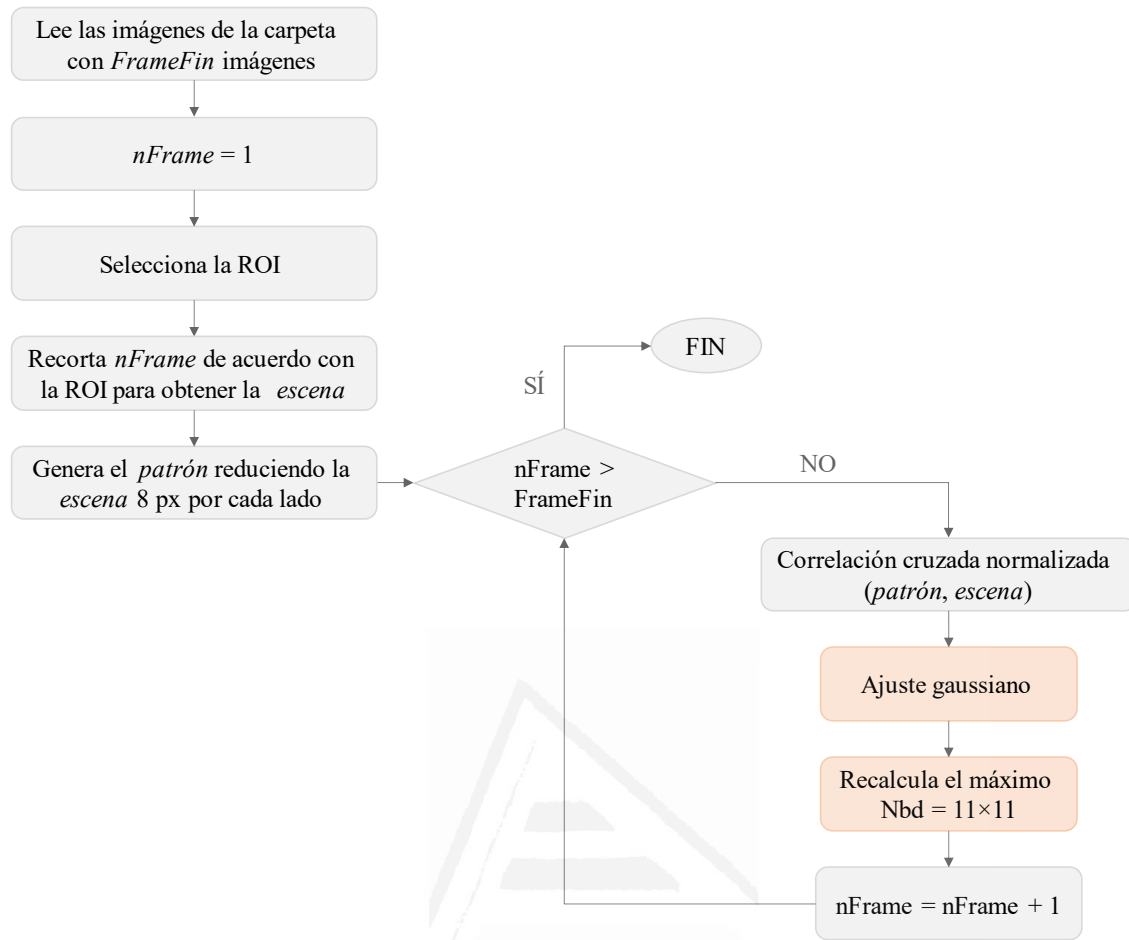


Figura 4.19. Diagrama de flujo del programa utilizado para evaluar si la textura propia del hormigón es adecuada para ser utilizada como diana natural al emplear la correlación.

A continuación, en la figura 4.20 se muestran las gráficas de los movimientos “fantasma” o error de posicionamiento detectados tanto en horizontal como en vertical para las 6 secuencias medidas con la probeta de hormigón fluida/maza/acero. Se ha añadido también la línea de tendencia del movimiento para ver si hay una tendencia clara en el movimiento y comprobar así si se han reducido las derivas de la cámara observadas en el estudio de la roca arenisca. Además, la línea de tendencia permite comparar los resultados de las distintas gráficas de manera sencilla. Seguidamente, en la tabla 4.2, se muestran la media (μ) y la desviación estándar (σ) de los movimientos medidos con la correlación para proporcionar información sobre el error de la medida por tratarse de una grabación estática, así como la pendiente de la línea de tendencia. Tanto en la figura 4.20 como en la tabla 4.2, solo se muestran los resultados para la probeta de hormigón fluida/maza/acero como muestra representativa de los resultados obtenidos, ya que los resultados son similares en todas las probetas de hormigón y mostrar 126 gráficas (6 secuencias por cada una de las 21 probetas estudiadas) sería demasiado largo y pesado.

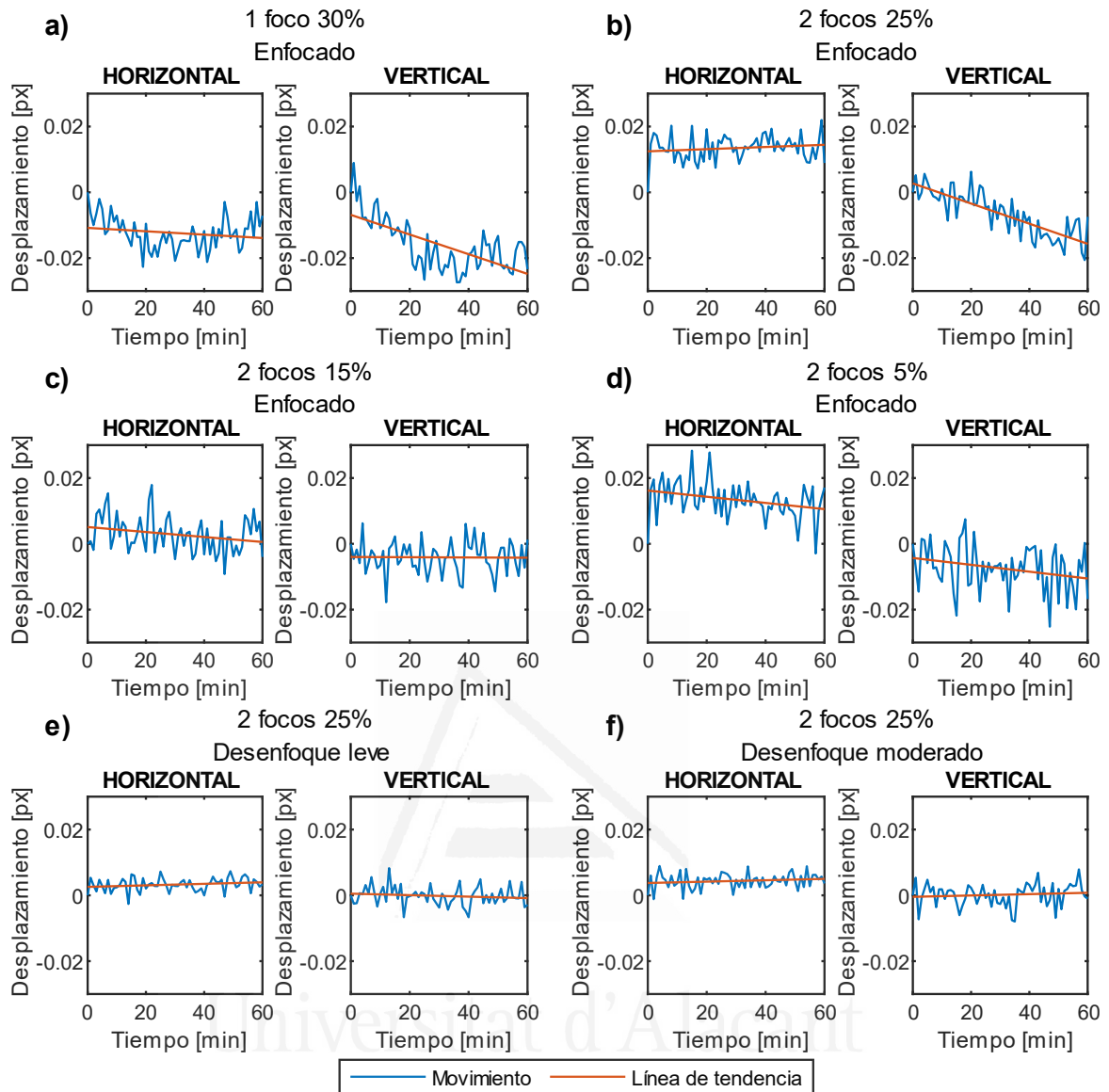


Figura 4.20. Movimientos horizontal y vertical (azul) junto con la línea de tendencia (rojo) para las secuencias **a)** luz lateral con un 30% de iluminación y enfocado, **b)** luz bidireccional con un 25% de iluminación y enfocado, **c)** luz bidireccional con un 15% de iluminación y enfocado, **d)** luz bidireccional con un 5% de iluminación y enfocado, **e)** luz bidireccional con un 25% de iluminación y un ligero desenfoque, **f)** luz bidireccional con un 25% de iluminación y un desenfoque moderado.

Tabla 4.2. Resultados de la media (μ) y la desviación estándar (σ) del movimiento y de la pendiente de la línea de tendencia tanto en horizontal como en vertical para las 6 secuencias mostradas en la figura 4.20 de la probeta de hormigón fluida/maza/acero.

| | Horizontal | | Vertical | |
|--------------------------------|-----------------------|------------------------|----------------------------------|------------------------|
| | $\mu \pm \sigma$ [px] | pendiente | $\mu \pm \sigma$ [px] | pendiente |
| 1 foco 30% enfocado | -0.012 ± 0.005 | -5.12×10^{-5} | -0.016 ± 0.008 | -2.99×10^{-4} |
| 2 foco 25% enfocado | 0.013 ± 0.004 | 3.30×10^{-5} | -0.065 ± 0.007 | -3.80×10^{-4} |
| 2 foco 15% enfocado | 0.003 ± 0.006 | -7.57×10^{-5} | -0.004 ± 0.005 | -3.81×10^{-6} |
| 2 foco 5% enfocado | 0.013 ± 0.006 | -9.40×10^{-5} | -0.007 ± 0.007 | -1.04×10^{-4} |
| 2 foco 25% desenfoque leve | 0.003 ± 0.002 | 2.31×10^{-5} | $-1.68 \times 10^{-4} \pm 0.003$ | -2.26×10^{-5} |
| 2 foco 25% desenfoque moderado | 0.004 ± 0.002 | 2.09×10^{-5} | $1.63 \times 10^{-4} \pm 0.003$ | 2.04×10^{-5} |

Al utilizar la propia textura del hormigón se han conseguido unos errores menores a los obtenidos previamente con las imágenes compuestas por círculos binarios que se usaron para evaluar el error de la cámara y del método de procesado (ver figura 4.14 *Con dissipador*). En este caso, el error ha sido un orden de magnitud menor independientemente de las características de fabricación de la probeta y de las condiciones de iluminación y desenfoque utilizadas en la grabación, lo que demuestra la hipótesis de partida en la que se sugería un buen comportamiento de la correlación para medir movimientos en texturas de hormigón por las características particulares de su superficie, similares a las texturas de tipo speckle.

Además, se puede observar en las gráficas mostradas para el caso de la probeta de hormigón fluida/maza/acero, que desenfoque la imagen mejora los resultados, aunque no hay una gran diferencia entre el desenfoque ligero y el moderado. No obstante, no se han realizado grabaciones con un desenfoque alto porque se ha partido de la premisa de que con un desenfoque fuerte los resultados obtenidos con la correlación cruzada empeorarían, ya que así se demostró con un círculo teórico en la figura 3.1 del Capítulo 3. Este empeoramiento de los resultados con un gran desenfoque puede deberse a que la textura del objeto deja de apreciarse y, por tanto, deja de ser adecuada para su uso con la correlación cruzada.

Como conclusión final, es importante resaltar que los resultados muestran que las texturas propias del hormigón presentan mejores resultados que los obtenidos para una diana binaria, por lo que la hipótesis inicial se confirma, es decir, la textura propia del hormigón presenta características que son particularmente buenas para el seguimiento del movimiento a través de la correlación cruzada. Por tanto, no sería necesario introducir dianas artificiales en la escena si el elemento cuyos movimientos se van a medir está construido con hormigón. Este hecho proporciona una gran mejora en el proceso de medida ya que no será necesario acceder al punto de media para colocar una diana, mejorando así la seguridad de los trabajadores, reduciendo costes y facilitando esta técnica a empresas pequeñas con presupuestos económicos limitados.

No obstante, todavía sigue siendo necesario continuar avanzando en esta línea de investigación debido a que la implementación práctica de este método en estructuras reales al aire libre todavía no se ha realizado. Como se ha comentado anteriormente, demostrar que se puede utilizar la correlación cruzada normalizada con imágenes de hormigón para determinar el movimiento o vibración de la estructura en la que se encuentra, es solo un

objetivo intermedio y los siguientes trabajos deberán centrarse en analizar el procedimiento aquí mostrado en condiciones reales, es decir, en estructuras de hormigón situadas al aire libre y a gran distancia del punto de medida. Para ello, es necesario analizar los inconvenientes que se presentan en ese tipo de toma de imágenes, como las distorsiones atmosféricas debidas al calor, los movimientos de la cámara y el trípode por el viento o la variación de la iluminación durante la toma, entre otros. Tras realizar una búsqueda bibliográfica sobre cómo minimizan estos errores presentes en las grabaciones al aire libre otros grupos de investigación, se intentarían implementar esas recomendaciones en nuestras medidas. Además, en el caso de que las imágenes grabadas no fueran lo suficientemente estables en cuanto a ruido y contraste para un análisis de ellas mediante la correlación cruzada, se deberá trabajar en un procesado de imágenes anterior al análisis de las mismas en el que se estabilizaran las imágenes para su posterior tratamiento.



Universitat d'Alacant
Universidad de Alicante

CAPÍTULO 5: Métodos subpíxel por análisis de luminancia

A raíz de la publicación de un trabajo del profesor J. Sheridan y sus colaboradores (Wan et al., 2020), se ha intentado reproducir el nuevo método propuesto para evaluar su utilidad y facilidad de uso con el tipo de imágenes que empleamos en nuestros estudios. Para ello, se ha trabajado en colaboración con ellos de modo que se han intercambiado ideas y resultados para obtener un método más eficiente por parte de ambos equipos de trabajo.

El método de la correlación cruzada junto con un ajuste del pico obtenido analizado en el Capítulo 3 es, posiblemente, el método más utilizado para el seguimiento de objetos con precisión subpíxel. Desafortunadamente, el cálculo de la posición del pico de correlación a partir de un ajuste local, introduce un sesgo en la posición hacia el entero más cercano, lo que produce errores en la localización que pueden ser mayores a un 20% (Murray et al., 2017). No obstante, el método basado en el análisis de la luminancia midiendo las variaciones de los niveles de gris acumulados es simple y no requiere ajuste, por lo que, a priori, no introducirá sesgos de localización. Por el contrario, la variación de niveles de gris debe ser calibrada previamente para obtener la relación de esta variación con el movimiento y obtener así la posición del objeto.

Nuestra propuesta consiste en la comparación del método de la variación de los niveles de gris acumulados (a partir de ahora, lo llamaremos VarCGL) con el método típico de la correlación cruzada. Lamentablemente, esta comparación no se puede realizar de modo directo, ya que, mientras que la correlación es capaz de proporcionar la posición relativa de un objeto respecto a una diana de referencia, el método VarCGL proporciona un parámetro indirecto que deberá ser calibrado para, a partir de ahí, obtener la posición del objeto.

En muchas aplicaciones reales, la calibración a partir de una secuencia real previa no será posible, ya sea por la irrepetibilidad del experimento o por ser una prueba destructiva. Así pues, resultaría interesante poder generar secuencias de movimiento para la calibración a partir de una imagen inicial, ya sea real o sintética. La generación de estas secuencias con desplazamientos subpíxel de modo directo, sólo es posible para figuras geométricas simples, por lo que se tendrá que recurrir a diferentes técnicas de interpolación.

En principio, los dos métodos más utilizados para generar un desplazamiento digital arbitrario son:

- Interpolación directa utilizando el algoritmo *griddedInterpolant* (o interpolación de rejilla) de Matlab. Este algoritmo desplaza la imagen de interés un determinado desplazamiento que puede ser diferente en horizontal y en vertical.
- Utilizar la propiedad de traslación de la Transformada de Fourier (conocida en inglés como Shifting Property) que establece que la Transformada de Fourier de un objeto desplazado es igual a la Transformada de Fourier del objeto original multiplicada por una fase (Gaskill, 1978).

$$\mathcal{F}\{f(x - x_0)\} = e^{-j2\pi x_0 \xi} F(\xi)$$

De acuerdo con esto, el objeto del primer fotograma se multiplica por la exponencial compleja adecuada para obtener el desplazamiento deseado. Después, se obtiene la parte real de la inversa de la transformada, siendo ésta una versión desplazada de la entrada.

El problema que puede aparecer a la hora de implementar un desplazamiento digital es que las interpolaciones realizadas distorsionen el objeto, produciendo como resultado un error en el desplazamiento generado. Para evaluar esta posible distorsión se usará el método VarCGL ya que no introduce distorsión y las diferencias que se calculan entre imágenes son muy sensibles a pequeñas variaciones. Así, generaremos secuencias sintéticas con objetos simples desplazados de manera analítica. A continuación, generaremos las mismas secuencias a partir de un fotograma inicial, calculando los objetos en las posiciones desplazadas usando los dos métodos de interpolación descritos. Así evaluaremos el parámetro VarCGL en cada secuencia para detectar las diferencias entre las secuencias y determinar el mejor método de generación de desplazamientos de entre los dos propuestos.

Recordemos que para calcular el VarCGL se calcula, en primer lugar, la matriz de la diferencia de los niveles de gris entre la imagen del objeto y la trasladada (5.1), después se suman los valores obtenidos por columnas y por filas (5.2) y, finalmente, se calculan las sumas absolutas de estos resultados (5.3) tal y como se explicó con más detalle en el Capítulo 2.4.

$$D = I_0 - I_{\delta_x} \quad (5.1)$$

$$C_i = \sum_{j=1}^N D_{ij} \quad R_j = \sum_{i=1}^N D_{ij} \quad (5.2)$$

$$T_C = \sum_i |C_i| \quad T_R = \sum_j |R_j| \quad (5.3)$$

Tras esto, se comparará el método de la correlación cruzada con el método VarCGL en imágenes sintéticas sencillas y se evaluarán las ventajas e inconvenientes de cada uno. Finalmente, se compararán ambos métodos del mismo modo, pero utilizando secuencias reales medidas en el laboratorio con condiciones de iluminación controladas.

5.1. Evaluación de métodos de generación de secuencias de calibración digitales

Para la calibración se han generado los tres objetos sencillos mostrados en la figura 5.1 (un cuadrado binario, un círculo binario y una gaussiana) de manera numérica, de modo que el desplazamiento se incluye en la propia ecuación de generación del objeto. De este modo, se conoce su posición en todo momento y no se introducen errores de desplazamiento o de medida como ocurriría en los experimentos reales. Posteriormente, a partir de la primera imagen de la secuencia, se han generado dos secuencias adicionales, cada una con un método de los comentados anteriormente y se han comparado los resultados de los parámetros T_C y T_R del método VarCGL obtenidos con cada una de las tres secuencias generadas.



Figura 5.1. Objetos sencillos utilizados para comparar los métodos VarCGL y correlación cruzada de forma teórica.

Los tres objetos se han creado en una matriz de soporte de 512×512 px con diferentes tamaños de objeto desde 2 hasta 400 px, siendo este tamaño la anchura en el caso del cuadrado, el diámetro en el caso del círculo y la desviación típica (sigma) en el caso de la gaussiana. La precisión se ha fijado en 16 bits para evitar un ruido excesivo de cuantización. Desde el fotograma inicial, la diana se desplaza un píxel completo en pasos de 0.005 px en la dirección horizontal y sin movimiento vertical. Este desplazamiento se ha implementado de diferentes formas en función del objeto en cuestión.

El caso del cuadrado binario es un poco complejo ya que, aunque se estén usando desplazamientos subpíxel, al estar alineado con la cuadrícula del sensor, la función sólo podrá tomar valores en las coordenadas enteras, según se describe en las ecuaciones 5.4 y 5.5:

$$f(x, y) = \text{rect}\left(\frac{x - x_0}{w}\right) \times \text{rect}\left(\frac{y}{w}\right) \quad (5.4)$$

con

$$\text{rect}\left(\frac{x-x_0}{w}\right) = \begin{cases} 1 & x \in \left[-\frac{w}{2} + x_0, \frac{w}{2} + x_0\right] \\ 0 & \text{en otro caso} \end{cases} \quad (5.5)$$

siendo w la anchura del rectángulo y x_0 el desplazamiento impuesto en el eje “x”. De modo equivalente se puede definir para la variable “y”.

Obsérvese que, con esta definición clásica, en el caso discreto, un objeto rectangular no mostrará cambios hasta que se haya desplazado un píxel completo. No obstante, en la vida real con dianas rectangulares, cuando el objeto se mueve, la versión capturada sufre una degradación del borde: si el movimiento es hacia la derecha, la primera columna brillante representando el borde izquierdo se va oscureciendo mientras ocurre lo contrario en la primera columna del lado derecho como efecto del área de iluminación saliendo o entrando en las nuevas áreas del sensor de CCD/CMOS. Así:

$$\text{rect}\left(\frac{x-\delta_x}{w}\right) = \begin{cases} 1 & x \in \left[-\frac{w}{2} + \delta_x + 1, \frac{w}{2} + \delta_x - 1\right] \\ 1 - \delta_x & x = \left[-\frac{w}{2} + \delta_x\right] \\ \delta_x & x = \left[\frac{w}{2} + \delta_x\right] \\ 0 & \text{en otro caso} \end{cases} \quad (5.6)$$

siendo δ_x el parámetro de desplazamiento subpíxel definido en el intervalo (0, 1).

A continuación y para evitar las dificultades que presenta el cuadrado al tener los bordes alineados con la matriz de soporte y con la dirección del movimiento, cosa que podría presentar unos resultados distorsionados, se ha evaluado un círculo binario con diferentes radios (r). Para ello, se ha utilizado la expresión:

$$f(x, y) = \begin{cases} 1 & (x-\delta_x)^2 + y^2 \leq r^2 \\ 0 & \text{en otro caso} \end{cases} \quad (5.7)$$

Finalmente, hemos analizado una gaussiana, que presenta bordes suaves. Para generarla se ha utilizado la ecuación 5.8.

$$f(x, y) = \exp\left(-\frac{(x-\delta_x)^2 + y^2}{2\sigma^2}\right) \quad (5.8)$$

Con los tres objetos generados sintéticamente según las ecuaciones descritas anteriormente, se han calculado los valores de T_C y T_R para la secuencia generada

directamente y para las desplazadas a partir de la primera imagen con la interpolación de rejilla utilizando el algoritmo *griddedInterpolant* de Matlab y con la propiedad de traslación de la Transformada de Fourier. En las figuras 5.2 a 5.4 se representa la diferencia de estos valores entre las imágenes sintéticas y las obtenidas por ambos métodos de interpolación.

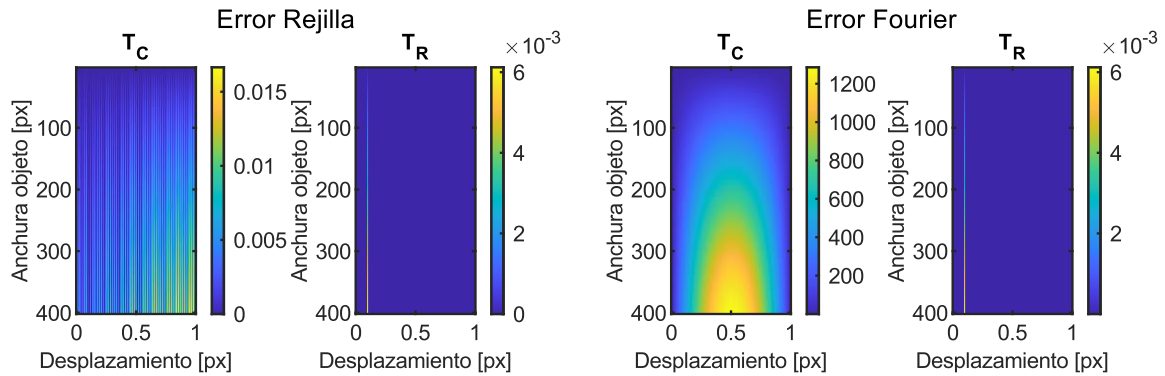


Figura 5.2. Mapas de error de los parámetros T_C y T_R obtenidos al comparar los valores medidos mediante el cálculo directo del cuadrado desplazado y la imagen recalculada con la interpolación de rejilla (izquierda) y con el método de Fourier (derecha).

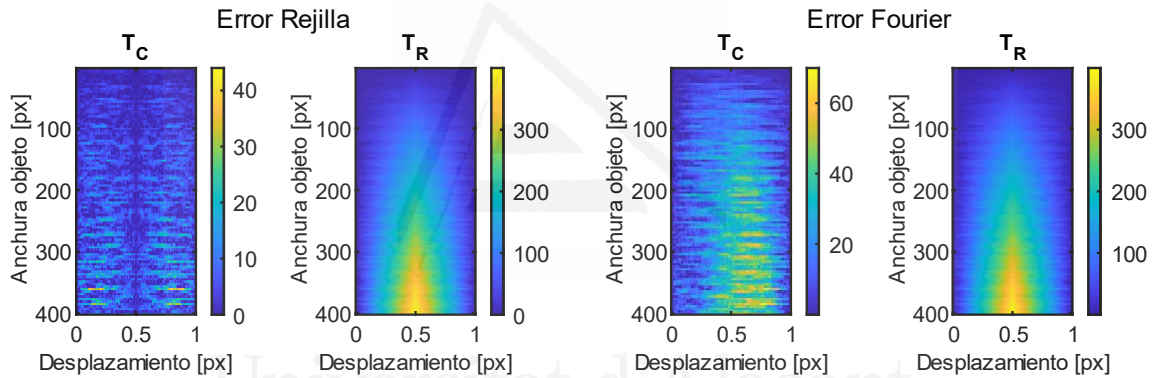


Figura 5.3. Mapas de error de los parámetros T_C y T_R obtenidos al comparar los valores medidos mediante el cálculo directo del círculo desplazado y la imagen recalculada con la interpolación de rejilla (izquierda) y con el método de Fourier (derecha).

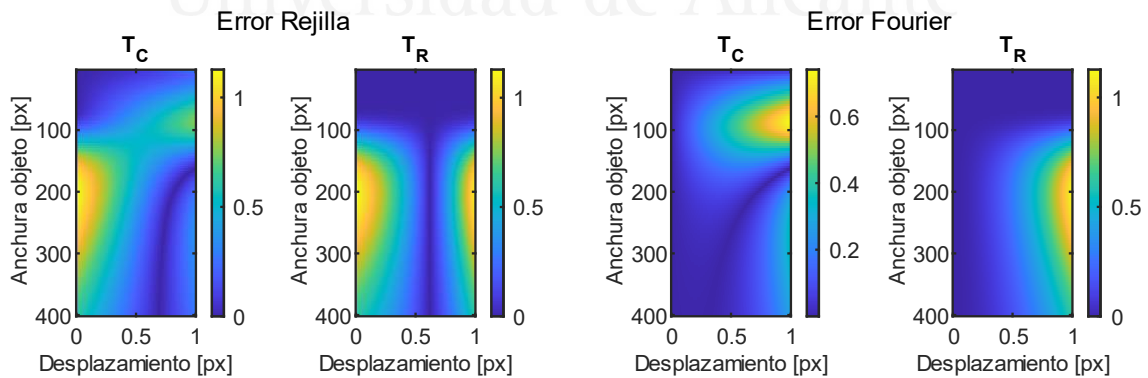


Figura 5.4. Mapas de error de los parámetros T_C y T_R obtenidos al comparar los valores medidos mediante el cálculo directo de la gaussiana desplazada y la imagen recalculada con la interpolación de rejilla (izquierda) y con el método de Fourier (derecha).

Como solo se ha generado un movimiento en la dirección horizontal, en T_R se debería observar un valor de 0 en todos los casos, mientras que en T_C sí que podría haber alguna

diferencia más acusada. En el caso del cuadrado se puede ver esa tendencia, pero no así en los otros dos objetos.

En el cuadrado (figura 5.2), se aprecia que el método de cálculo de imagen desplazada utilizando la interpolación de rejilla proporciona mejores resultados que el método de Fourier, en el sentido de que la versión desplazada mediante interpolación de rejilla es muy similar a la calculada de modo directo según la ecuación 5.6. Esto probablemente se debe a que la Transformada de Fourier aplicada sobre una función con bordes abruptos provoca oscilaciones en los bordes por los cambios bruscos de luminancia, lo que distorsiona las sumas. Esto puede apreciarse en la figura 5.5 donde se muestran las diferencias en valor absoluto entre un cuadrado de anchura 100 px generado analíticamente desplazado 0.35 px y sus versiones generadas mediante los dos métodos propuestos.

En el caso del círculo (figura 5.3), los errores son similares en las dos formas de generación de movimiento de la secuencia, aunque siguen siendo menores para el método de interpolación de rejilla. En este caso se observa un fenómeno curioso. Como ya se ha comentado antes, dado que el movimiento es horizontal y las imágenes son sintéticas, el valor de T_R debería ser muy bajo o nulo de la misma forma que sucede en el cuadrado. En cambio, los errores son muy altos en ambos métodos de interpolación con el círculo, por lo que pensamos que esto es un efecto de bordes combinado con el efecto de la simetría del círculo en el que se tienen bordes orientados en todas las direcciones. En la figura 5.6 se muestran las diferencias en valor absoluto entre los métodos y se puede comprobar este efecto para un círculo de diámetro 100 px desplazado 0.35 px.

Finalmente, para el perfil gaussiano (figura 5.4), ambos métodos de generación de desplazamiento producen unos resultados similares, siendo ligeramente mejor el caso del método de la Transformada de Fourier. En la figura 5.7, calculado con un sigma de 100 px y un desplazamiento de 0.35 px, se puede ver que, como los bordes son suaves, las diferencias entre métodos no son tan acusadas como en los dos objetos binarios anteriores.

Además, en las distribuciones de errores por tamaño de objeto vemos que, aunque el caso gaussiano presenta una distribución de error sin un patrón definido, en general, el error aumenta cuanto mayor sea el tamaño, lo cual refuerza la idea de que el efecto de bordes y el aliasing influyen en el resultado.

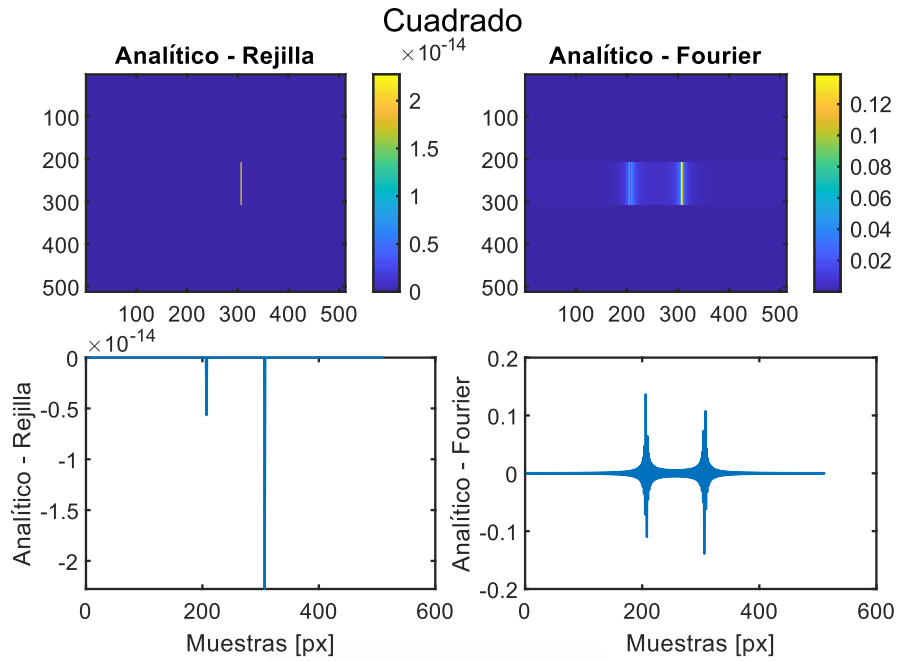


Figura 5.5. Diferencias en valor absoluto entre un cuadrado de anchura 100 px generado analíticamente desplazado 0.35 px y sus versiones generadas mediante ambos métodos de desplazamiento; punto a punto (fila superior) y perfil en la línea central (fila inferior).

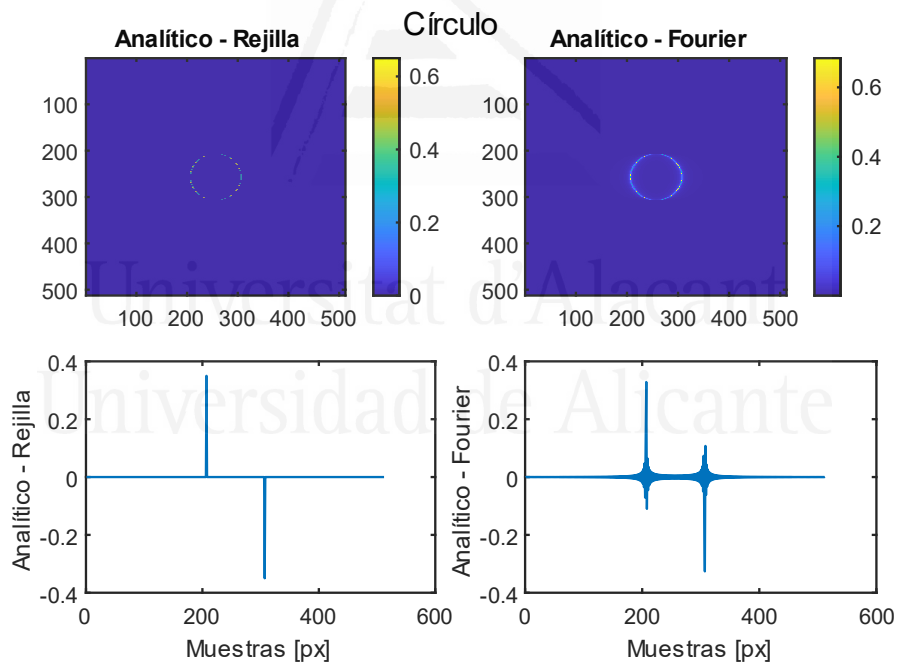


Figura 5.6. Diferencias en valor absoluto entre un círculo de 100 px de diámetro generado analíticamente desplazado 0.35 px y sus versiones desplazadas generadas mediante ambos métodos de desplazamiento; punto a punto (fila superior) y perfil en la línea central (fila inferior).

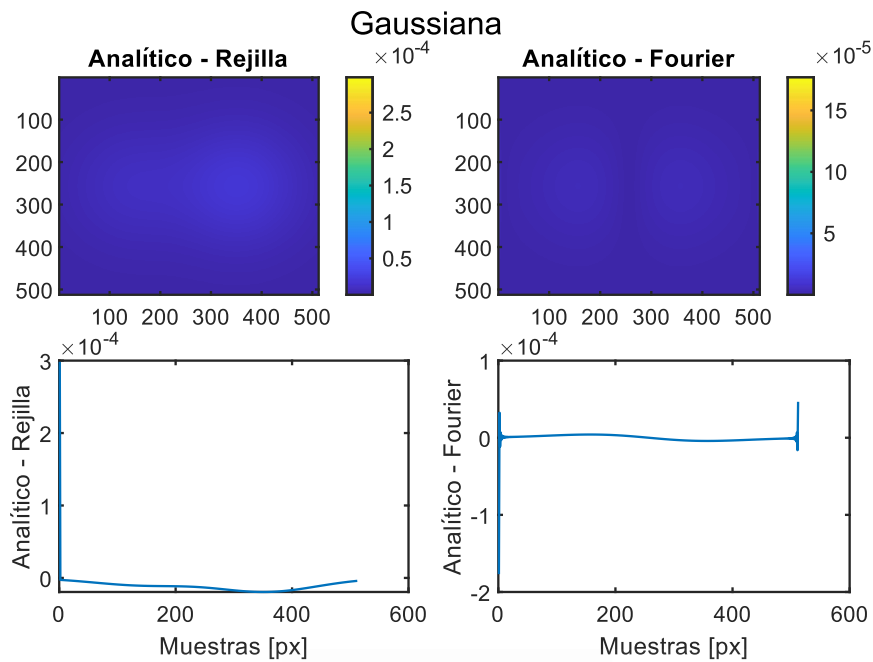


Figura 5.7. Diferencias en valor absoluto entre una gaussiana de $\sigma = 100$ px generada analíticamente desplazada 0.35 px y sus versiones desplazadas generadas mediante ambos métodos de desplazamiento; punto a punto (fila superior) y perfil en la línea central (fila inferior).

Con el fin de comprobar si el error de los métodos se debe a la presencia de bordes o a un error intrínseco del método, se han repetido las figuras anteriores con los objetos desenfocados. Del mismo modo que se hizo en los capítulos anteriores, el desenfoco se consigue aplicando un filtro gaussiano sobre el objeto original. En este caso, el filtro gaussiano introducido tiene un $\sigma = 3$, lo que significa aplicar un desenfoco ligero. Se debe destacar que el desenfoco sólo se ha aplicado para el cuadrado (figura 5.8) y el círculo (figura 5.9), ya que no tendría mucho sentido desenfocar una gaussiana.

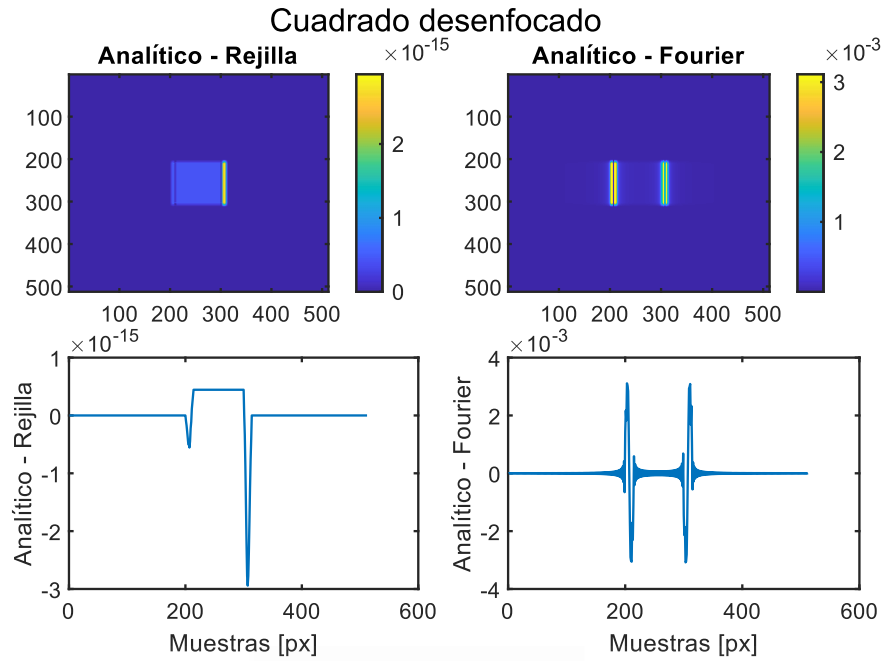


Figura 5.8. Representación de la figura 5.5 pero con el cuadrado desenfocado para comprobar si el error de los métodos se debe a la presencia de bordes o a un error intrínseco del método.

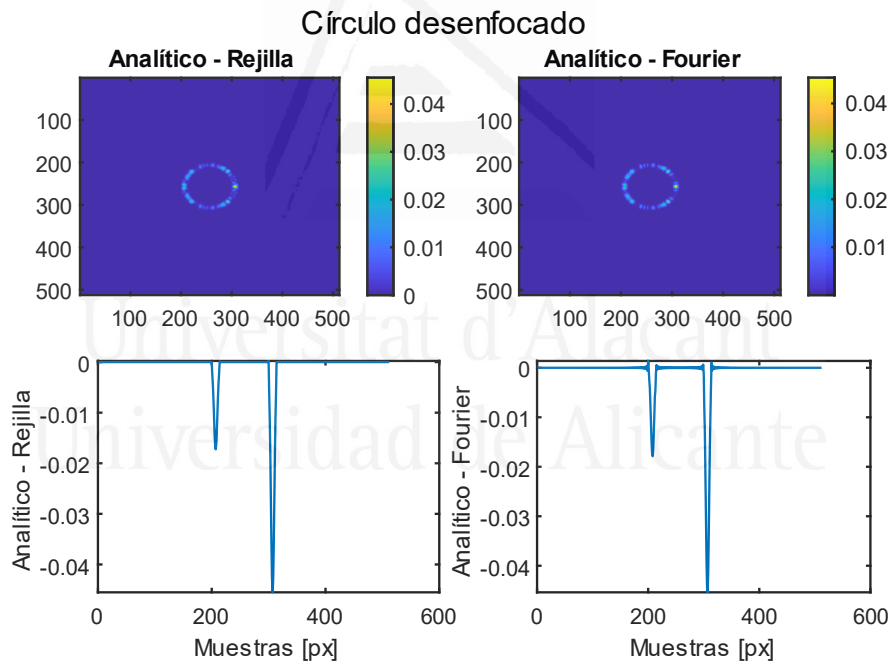


Figura 5.9. Representación de la figura 5.6 pero con el círculo desenfocado para comprobar si el error de los métodos se debe a la presencia de bordes o a un error intrínseco del método.

Se puede ver que, aunque sigue existiendo error en los bordes, éste ha disminuido en dos órdenes de magnitud. Además, se puede observar que las diferencias entre los dos métodos ya no son tan acusadas. Por tanto, el error está causado por la presencia de bordes y se puede solucionar fácilmente desenfocando ligeramente las imágenes.

5.2. Comparación numérica de los métodos VarCGL y correlación cruzada

Para seguir con la evaluación del método VarCGL, se ha comparado el funcionamiento de este frente al método de correlación usando una matriz alrededor del pico de 9×9 px con un ajuste gaussiano, de acuerdo con lo establecido en el Capítulo 3. Esta comparación se ha realizado para las diferentes secuencias sintéticas generadas con los dos métodos de generación de desplazamiento analizados anteriormente.

Recordemos que en el método VarCGL se calcula un parámetro ligado a la variación de niveles de gris que, en teoría, aumenta linealmente con el desplazamiento. Dicho parámetro se ha ajustado linealmente a la posición conocida, obteniéndose así una correspondencia directa entre la variación de los niveles de gris y la posición del objeto. En este caso, dado que el movimiento simulado es lineal y uniforme en sentido horizontal, con una amplitud máxima de 1 px, bastará con normalizar el parámetro T_C a su valor máximo, de modo que la correspondencia entre su valor y la posición estimada sea directa. El valor de T_R también se ha normalizado al valor máximo de T_C con el fin de facilitar la comparativa en ambas direcciones.

En las figuras 5.10 a 5.12 se muestran los resultados obtenidos con el método VarCGL y con la correlación cruzada para los dos tipos de generación de secuencias utilizados, tanto con objetos de anchura 100 px nítidos como ligeramente desenfocados ($\sigma = 3$).

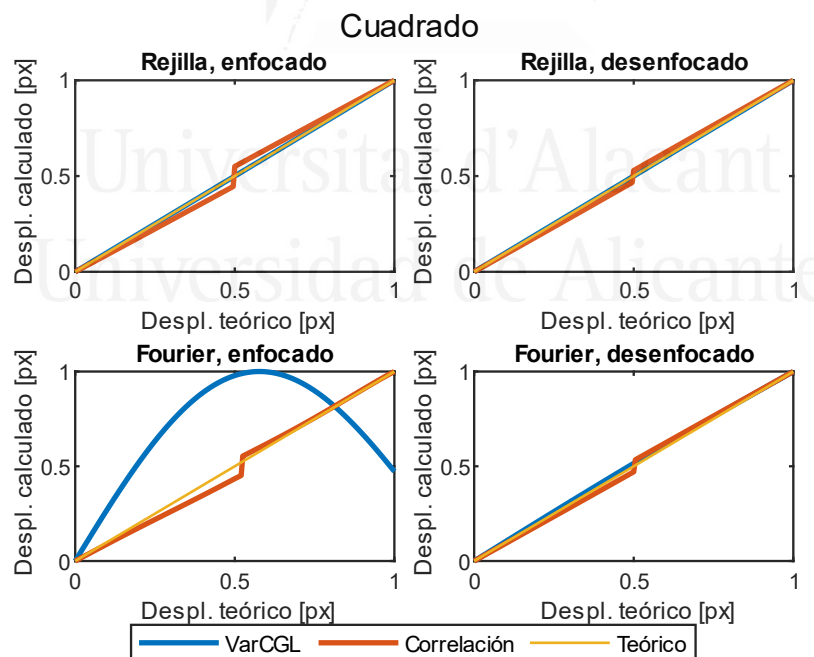


Figura 5.10. Resultados con los métodos VarCGL y correlación cruzada para los dos tipos de generación de secuencias con un cuadrado de anchura 100 px enfocado y desenfocado.

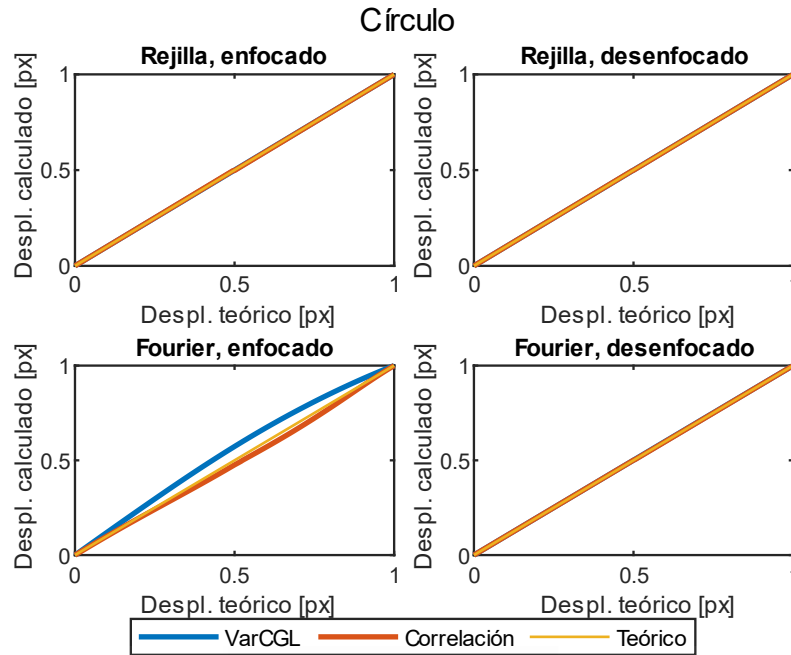


Figura 5.11. Resultados con los métodos VarCGL y correlación cruzada para los dos tipos de generación de secuencias con un círculo de diámetro 100 px enfocado y desenfocado.

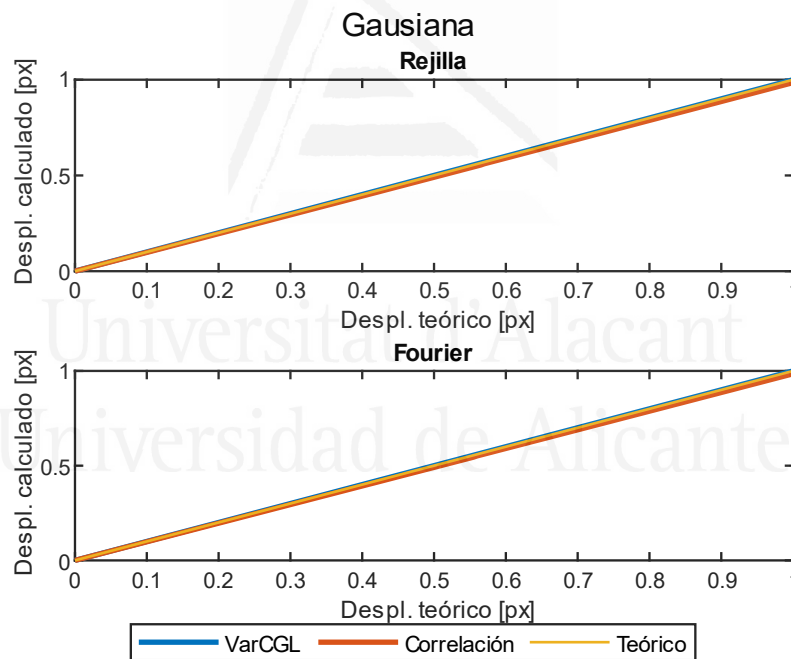


Figura 5.12. Resultados con los métodos VarCGL y correlación cruzada para los dos tipos de generación de secuencias con una gaussiana enfocada de sigma 100 px.

A la vista de los resultados para el cuadrado y el círculo, se puede decir que, en primer lugar, el método de Fourier para generar desplazamientos puede introducir demasiado error en el cálculo. En los casos en los que el resultado es preciso, éste no es significativamente mejor que el método de interpolación de rejilla, con lo que recomendamos descartar este método para el cálculo de secuencias desplazadas.

En cuanto al método de localización, se puede ver que, en el cuadrado, aparece muy claramente un efecto de atrapamiento de pico cuando se usa la correlación, mientras que el

efecto no se aprecia en el caso del método VarCGL. En general, también se ve que el desenfoque ayuda a mejorar el resultado en el caso de la localización por correlación, con resultado neutro en el caso del VarCGL.

En el caso de la diana con perfil gaussiano, se ve que el resultado es muy similar para ambos métodos, aunque ligeramente mejor para el VarCGL que para la correlación.

Finalmente, se ha analizado el funcionamiento de los diferentes métodos con el objeto del jinete (figura 5.13) que tiene los bordes más complejos. El resultado muestra un comportamiento similar al del cuadrado, quizá por el efecto del aliasing por la presencia de bordes muy definidos.



Figura 5.13. Objeto del jinete con bordes más complejos.

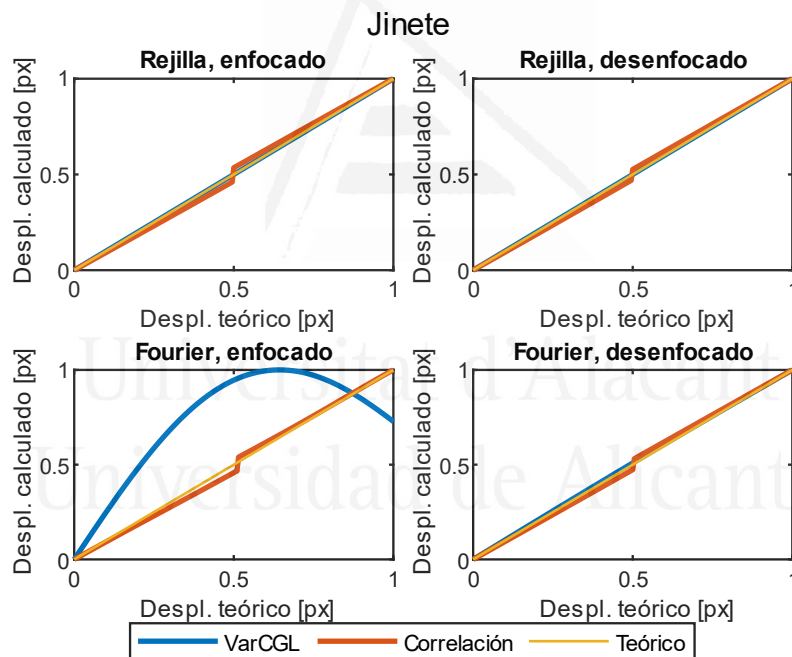


Figura 5.14. Resultados con los métodos VarCGL y correlación cruzada para los dos tipos de generación de secuencias con un objeto complejo enfocado y desenfocado.

Con el fin de completar las pruebas se ha analizado la existencia de movimientos fantasma en el eje vertical dado que el movimiento impuesto ha sido únicamente horizontal. El movimiento detectado tiene una estructura aleatoria con una amplitud del orden de 10^{-5} píxeles, lo que demuestra que se trata de ruido de cálculo y no de errores de localización. Este comportamiento no se cumple, sin embargo, con la gaussiana, que muestra un error de localización importante en el eje “y” (figura 5.15) cuando se usa la combinación de interpolación de rejilla con localización usando VarCGL. No obstante, se ha comprobado que el efecto desaparece con una gaussiana más estrecha ($\sigma = 50$), con lo que pensamos que

el problema podría deberse a un recorte de bordes de la gaussiana en la ventana de trabajo y no a un error del método.

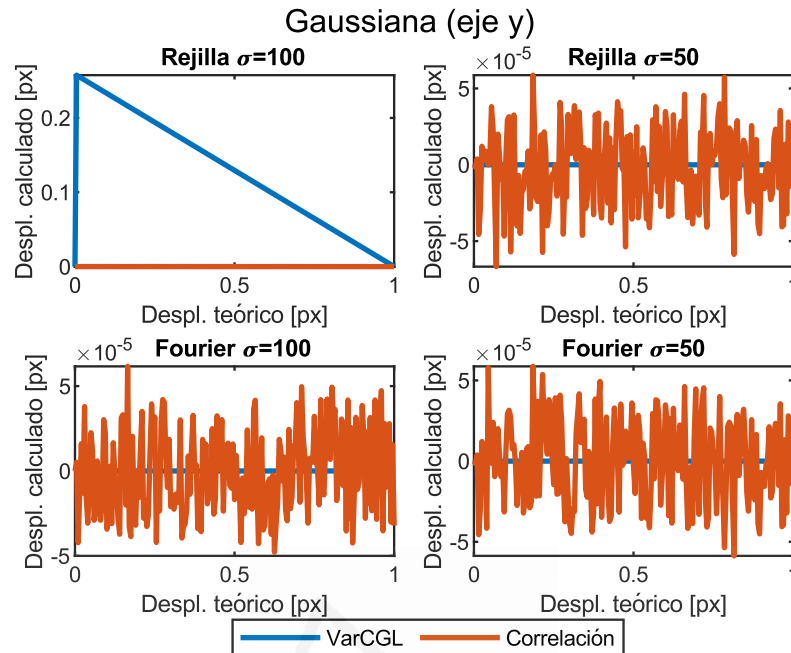


Figura 5.15. Error de desplazamiento de la gaussiana enfocada de sigmas 100 px y 50 px en el eje "y".

En los casos anteriores se ha comprobado que el método VarCGL no sufre ciertos inconvenientes que sí aparecen con la correlación, aunque es importante contar con una buena simulación de movimientos para calibrar el método correctamente. Para ello, deberemos tener ciertas precauciones con el tamaño relativo entre la diana y el cuadro de imagen, así como con los bordes muy definidos, siendo preferible suavizar los perfiles. No obstante, los errores son relativamente pequeños, con lo que podría suceder que experimentalmente estas diferencias no fuesen relevantes frente a errores de enfoque, ruido térmico o de otro tipo.

Por tanto, procederemos a implementar ambos métodos de localización en el laboratorio.

5.3. Comprobación del método VarCGL utilizando secuencias reales

La comparación del método VarCGL y de la correlación cruzada se ha realizado experimentalmente moviendo objetos binarios capturados con una cámara de 8 bits. Las dianas utilizadas para las secuencias reales grabadas en laboratorio han sido las mismas utilizadas previamente en el análisis digital para poder comparar los resultados. Se trata, por tanto, de las mostradas en las figuras 5.1 y 5.13 impresas en cartulina blanca y pegadas sobre

un desplazador lineal movido por un motor de desplazamiento continuo (DSM100/M de Thorlabs) con un incremento de movimiento mínimo de $0.5 \mu\text{m}$ y una repetibilidad de $\pm 1.5 \mu\text{m}$.

La plataforma desplaza horizontalmente el objeto una distancia de 4 mm a una velocidad de 0.4 mm/s. La secuencia ha sido capturada con una cámara BASLER acA800-510 um con una resolución espacial de $800 \times 600 \text{ px}$ y una resolución temporal de 300 fps. La cámara se ha colocado a 5 m del objeto, proporcionando una relación digital/real de 1.4270 px/mm. De acuerdo con la cámara y el desplazador, el desplazamiento entre fotogramas es de 0.0024 px o, equivalentemente, de $1.6 \mu\text{m}$. Se puede consultar más información sobre las características y uso de las cámaras y en los Anexos I y II.

Para garantizar una iluminación uniforme del objeto, se ha instalado dentro de una cabina de iluminación portátil con paredes reflectantes como se muestra en la figura 5.16. Se puede consultar más información sobre los sistemas de iluminación y las cabinas de iluminación en el Anexo III.

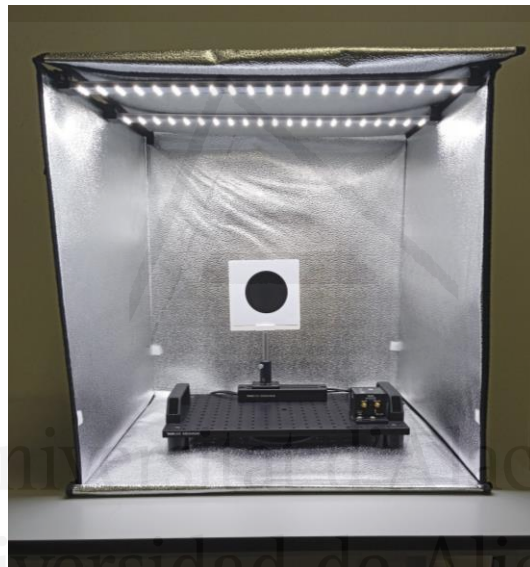


Figura 5.16. Montaje experimental del objeto sobre el motor de desplazamiento continuo dentro de una cabina de iluminación con paredes reflectantes para garantizar una iluminación uniforme del objeto.

Es importante resaltar que las imágenes se han impreso en tinta negra, pero todos los fotogramas se han invertido antes de realizar cualquier cálculo o procesado, quedando el objeto blanco sobre fondo negro para poder calcular de forma adecuada tanto la correlación cruzada como el método VarCGL.

El tiempo de captura ha sido ligeramente mayor que el de desplazamiento, por lo que en todas las secuencias horizontales se puede observar una fase plana al principio y al final de la representación y unos codos marcando el inicio y el final de la fase de movimiento en los que la velocidad es variable.

En las gráficas de la figura 5.17 se muestran el desplazamiento en mm obtenido para la correlación y calculado para el método VarCGL. La correspondencia entre los parámetros

del método VarCGL y el desplazamiento se ha realizado del mismo modo que en los cálculos teóricos.

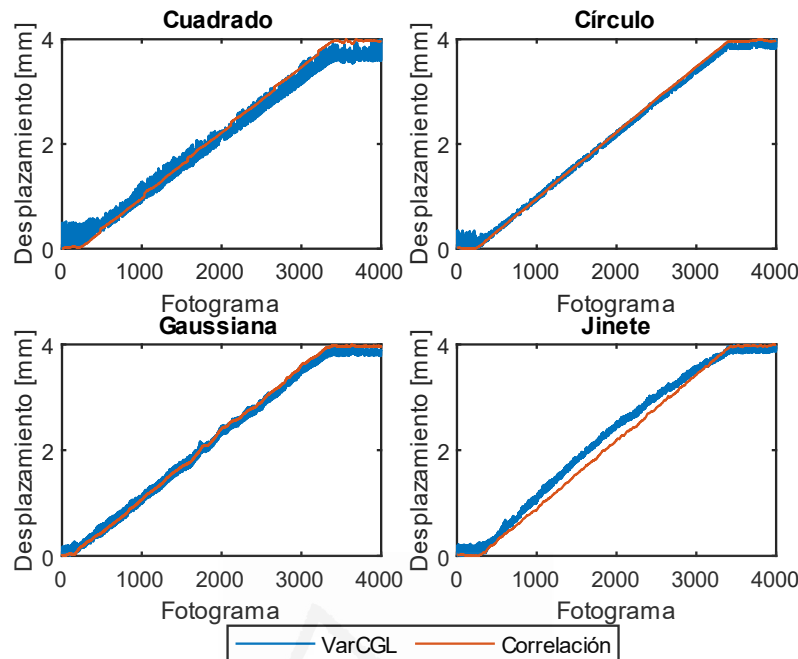


Figura 5.17. Desplazamientos calculados mediante el método VarCGL y la correlación cruzada para los 4 objetos grabados.

Lo primero que se nota es que, con independencia de la precisión del método VarCGL, los resultados obtenidos son más ruidosos que en la correlación. Esto podría deberse a las altas frecuencias del parpadeo de la luz de los LEDs que podrían distorsionar los resultados del método VarCGL.

Para compensar este efecto y descartar que el origen del ruido sea debido a la frecuencia de parpadeo de la lámpara de iluminación, se han vuelto a medir las secuencias corrigiendo el contraste en la imagen antes de comenzar la grabación de forma que el fondo sea lo más blanco posible y la figura sea lo más negra posible. En el caso de la gaussiana, el contraste se ha ajustado para que el centro alcance el valor máximo.

Con los ajustes de contraste mencionados, se grabaron nuevas secuencias cuyos resultados se muestran en la figura 5.18.

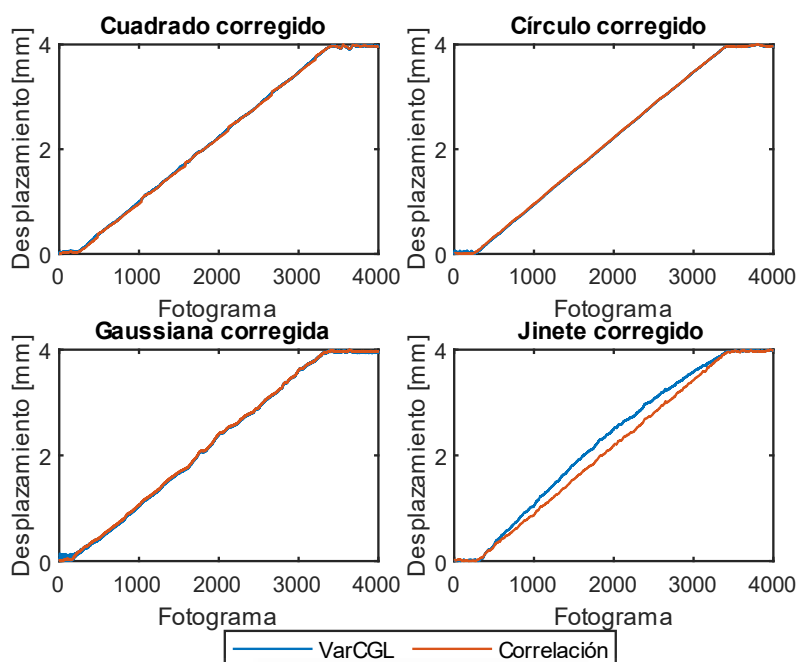


Figura 5.18. Desplazamientos calculados mediante el método VarCGL y la correlación cruzada para los 4 objetos grabados ajustando el contraste en las secuencias.

Ahora se puede observar que casi todas las variaciones desaparecen y que ambas curvas están prácticamente superpuestas. Se nota que la curva del VarCGL para el jinete no sigue una línea recta, aunque el movimiento grabado es completamente rectilíneo y a velocidad constante, igual que en el resto de objetos.

En la tabla 5.1 se muestran los resultados del error de localización. Para evitar los efectos del cambio de velocidad al principio y al final del movimiento, se ha restringido el análisis a los 3 mm centrales de la trayectoria y se ha calculado la diferencia entre la posición teórica y la medida. Además, para evitar diferencias en las posiciones de inicio, se han considerado desplazamientos relativos en vez de localizaciones exactas.

Tabla 5.1. Resultados de la localización del error (media \pm desviación estándar) para los 3 mm centrales y considerando desplazamientos relativos.

| | VarCGL | Correlación |
|------------------|----------------------|----------------------|
| Cuadrado | 0.0023 \pm 0.0129 | -0.0018 \pm 0.0182 |
| Círculo | -0.0123 \pm 0.0072 | -0.0075 \pm 0.0068 |
| Gaussiana | -0.0106 \pm 0.0295 | -0.0158 \pm 0.0283 |
| Jinete | -0.1265 \pm 0.0664 | 0.0043 \pm 0.0142 |

Los resultados obtenidos muestran que para dianas simples, la precisión de ambos métodos es similar, mientras que para una figura con bordes complejos como la del jinete, el error del método VarCGL da un error muy superior al obtenido por correlación. Por otra parte, se observa un comportamiento extraño en la diana cuadrada que no puede apreciarse debido a la escala de la gráfica, por lo que, se muestra a continuación una vista ampliada de 1 mm en el centro de la trayectoria donde se puede apreciar claramente un salto de la curva de correlación para la corrección de la trayectoria. En la figura 5.18 se puede intuir que ese comportamiento se repite a lo largo de la curva hasta 4 veces, coincidiendo con el desplazamiento de un píxel completo. De acuerdo con nuestras experiencias y resultados

previos, este efecto se debe al atrapamiento de pico amplificado por la forma particular de la diana.

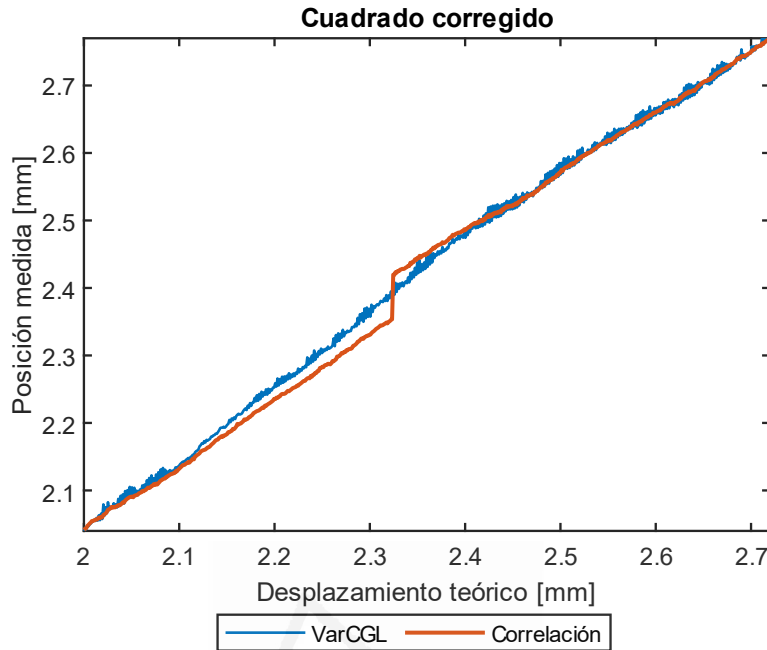


Figura 5.19. Vista ampliada de 1 mm central de la diana cuadrada para observar el comportamiento extraño que presenta.

En dianas simples, el ajuste del pico puede producir desviaciones suaves debidas al atrapamiento de pico. Sin embargo, en figuras complejas, dado que los cambios de luminancia en los bordes suceden en diferentes puntos y tiempos de la figura, la trayectoria calculada se corrige suavemente. Esto no ocurre en el cuadrado, en donde los bordes están alineados con la matriz de píxeles de la cámara y cada cambio de píxel ocurre para todos los puntos del borde a la misma vez, lo que implica una corrección rápida de la trayectoria.

En el caso de la figura del jinete, es posible que la forma del objeto y el ruido experimental provoquen que la variación de los parámetros usados en el método VarCGL no sea lineal con el desplazamiento, lo que provocará errores de posicionamiento. Por tanto, se hace necesario un estudio más pormenorizado del método en dianas de aspecto complejo.

En cualquier caso, a partir de este estudio preliminar se puede concluir que el método VarCGL:

1. No sufre de atrapamiento de pico.
2. Es más rápido computacionalmente.
3. En general, su error de posicionamiento es menor que el de la correlación.

Sin embargo, algunas posibles desventajas de este método son:

1. Insensibilidad en la dirección del movimiento.
Se trata de algo difícil de solucionar, al menos sin introducir más carga computacional, pero no es un gran problema ya que para el tipo de experimentos que realizamos, suele ser importante conocer el desplazamiento, pero no siempre su sentido.

2. Sensibilidad al ruido.

Se sabe que la correlación es óptima en términos de MSE (Maragos, 1989). Sin embargo, la integración directa de los niveles de gris puede estar mucho más afectada por cambios de iluminación, sombras o la presencia de ruido en la imagen. No obstante, muchos de estos problemas pueden superarse con procesados de imagen previos adecuados, lo cual puede añadir tiempo y complejidad en el cálculo.

3. Necesidad de calibración previa.

En algunos casos, realizar un pre-test experimental no es posible por la presencia de tests destructivos, procesos irreversibles, movimientos no controlados, deformaciones, etc. y una grabación experimental controlada permitiría conocer los niveles de gris que se corresponden con cada valor de desplazamiento. Este problema podría solucionarse realizando un desplazamiento teórico a la primera imagen de la secuencia experimental. Los movimientos simulados digitalmente en objetos a partir del primer fotograma han proporcionado unos resultados similares a los experimentales, mostrando que el precalibrado digital se puede realizar con garantías de éxito en experimentos reales.

Una posible aplicación de este nuevo método podría ser la calibración geométrica de cámaras, que generalmente se basa en la detección de los bordes en un tablero de ajedrez utilizando el método de la calibración de Zhang (Yu et al., 2017; Burger, 2016). Estos métodos no suelen tener en cuenta efectos de calentamiento de la cámara o derivas mecánicas que producen movimientos a nivel subpíxel. Además, las dianas utilizadas en forma de cuadrícula no son particularmente convenientes para la correlación, por lo que, en este caso, el método VarCGL puede ser una buena alternativa.

CAPÍTULO 6: Conclusiones

Tal y como se ha mencionado a lo largo de la presente tesis doctoral, hay diferentes técnicas subpíxel y su rango de aplicaciones es muy amplio, siendo común observar el empleo de una misma técnica en campos de investigación muy diferentes. En este caso, con los estudios llevados a cabo tanto teórica como experimentalmente, se han obtenido una serie de resultados muy útiles para nuestro equipo de investigación y para la comunidad científica en general.

A lo largo de los capítulos de la tesis, se han explicado en detalle varios métodos subpíxel y se han llevado a cabo experimentos para evaluar su funcionamiento en tareas tanto teóricas como experimentales. Específicamente, se han evaluado dos métodos: el primero de ellos es la correlación cruzada normalizada junto con un ajuste del pico de correlación e introduciendo un desenfoque en las imágenes grabadas y el segundo de ellos se trata del análisis de la luminancia comparando la variación de los niveles de gris acumulados entre imágenes.

En el primer caso, se han evaluado de forma teórica los mejores resultados obtenidos con la correlación cruzada utilizando tres tipos de ajuste del pico con diferentes tamaños de la matriz de ajuste, así como con distintos niveles de desenfoque gaussiano, obteniéndose que la combinación óptima de parámetros sería utilizar una matriz de ajuste del pico amplia y de tipo gaussiana sin introducir ningún desenfoque. No obstante, al utilizar esta técnica en experimentos reales, estos parámetros pueden verse ligeramente modificados en función de las características de las imágenes evaluadas. Aunque es posible que el método de ajuste más adecuado sea diferente en función del objeto, las condiciones en las que cada método funciona de manera óptima se mantienen.

Por otra parte y continuando con la evaluación del método de la correlación cruzada pero en experimentos reales, se ha utilizado por un lado una roca arenisca con una textura superficial similar al speckle para medir su expansión por hidratación al sumergirse parcialmente en agua. En este caso, el error de cálculo obtenido ha sido menor de $0.3 \mu\text{m}$, lo cual indica que el método es adecuado para medir desplazamientos locales en las direcciones horizontal y vertical utilizando la propia textura superficial de la roca arenisca. Además, para minimizar las derivas térmicas observadas en la cámara se realizó un experimento teórico utilizando una imagen compuesta por círculos binarios obteniéndose un error de $0.02 \pm 0.02 \text{ px}$ al utilizar un disipador de calor sobre la cámara y dejándola encendida 2 horas antes de comenzar con el experimento en cuestión. Por otro lado, se ha evaluado la textura superficial del hormigón para ser utilizada como diana natural para el seguimiento de desplazamientos en su superficie con una precisión subpíxel, sin necesidad de utilizar dianas artificiales y minimizando las derivas térmicas de la cámara. En este caso, el error obtenido ha sido un orden de magnitud menor que el obtenido teóricamente utilizando la imagen compuesta por círculos binarios. Por este motivo, se asume que la textura propia del hormigón es similar al speckle y proporciona información suficiente como para poder realizar un seguimiento de su superficie de forma fiable al utilizar la correlación cruzada.

En el segundo método evaluado, se han comparado los resultados obtenidos mediante el análisis de la luminancia con los obtenidos al utilizar la correlación cruzada empleando las mismas secuencias de imágenes para ambos métodos, obteniéndose que el método del análisis de la luminancia es más preciso que el de la correlación, no presenta atrapamiento de pico y es más rápido computacionalmente. Por el contrario, se trata de un método insensible a la dirección del movimiento, bastante variable en función del ruido o cambios en la iluminación de la escena y requiere de una calibración previa para poder utilizarse. Sin embargo, estos inconvenientes se pueden resolver con procesados de imágenes adecuados, cosa que añadiría tiempo de computación, por lo que en función de la escena donde se encuentre el objeto de interés, será recomendable un método de detección de movimientos con precisión subpíxel u otro. De todas formas, este nuevo método todavía está en estudio, por lo que puede ser que con futuras investigaciones se consigan mejorar sus inconvenientes llegando a ser posible su utilización en cualquier situación con resultados similares o mejores a los proporcionados por otros métodos de detección de movimientos o vibraciones con resoluciones subpíxel.

Mediante los trabajos desarrollados en la presente tesis doctoral, ha sido posible resolver los dos objetivos que se planteaban al inicio de la misma. Por una parte, se han estudiado teóricamente distintos métodos de seguimiento con resolución subpíxel, tanto mediante el estudio del estado del arte, como a través de los experimentos teóricos y prácticos planteados y resueltos. Por otra parte, se han implementado experimentalmente dos métodos subpíxel diferentes sin identificación previa del objeto, comparando además los resultados proporcionados por cada uno de ellos.

Resolviendo los objetivos planteados en esta tesis, se consigue adaptar las técnicas subpíxel a aplicaciones reales que puedan ser de interés y, además, se han mostrado aplicaciones muy distintas, lo que confirma que una misma técnica puede ser aplicada a situaciones reales diversas en las que se requiera medir un movimiento o vibración con precisiones por debajo de la resolución nominal de la cámara digital.

CAPÍTULO 7: Futuras líneas de investigación

En cuanto a las líneas de investigación que pueden continuarse a partir de la presente tesis doctoral, es importante continuar investigando en los siguientes tres puntos principales:

1. *En el nuevo método basado en el análisis de la luminancia*

Puesto que es bastante novedoso, todavía hay mucho que investigar para obtener mejores resultados en su implementación además de poder predecir a partir de más investigaciones el tipo de objetos o situaciones en los que puede proporcionar mejores resultados que los de otros métodos. En particular, habría que analizar el método con dianas complejas y en condiciones ambientales alejadas de las utilizadas en el laboratorio.

2. *El método de la correlación cruzada*

A pesar de que los resultados de esta tesis han sido novedosos y clarificadores, es importante seguir investigando en este campo para conseguir proporcionar los parámetros adecuados que se deberían seleccionar para otros tipos de objetos. También sería importante llevar a cabo un experimento con los mismos objetos utilizados teóricamente (objetos con speckle con distinto ruido y contraste) pero utilizando secuencias de movimientos reales medidas en condiciones controladas de laboratorio. Además, como se ha comentado en el Capítulo 4.2, para las superficies de hormigón todavía debe evaluarse que en situaciones reales al aire libre y a gran distancia del punto de medida es posible obtener errores de medición asumibles, ya que en estas situaciones existen otros inconvenientes como las distorsiones atmosféricas debidas al calor, los movimientos de la cámara y el trípode por el viento o la variación de la iluminación durante la medida.

3. *En nuevos métodos que puedan ir surgiendo para el seguimiento de objetos con precisión subpíxel*

Como siempre, es fundamental continuar leyendo la bibliografía actual para conocer las investigaciones que realizan otros grupos de investigación, cosa que puede ayudar a mejorar nuestros métodos de detección con precisión subpíxel o darnos ideas para comenzar otras líneas de investigación.



Universitat d'Alacant
Universidad de Alicante

BIBLIOGRAFÍA

- 1951 USAF resolution test chart, 2021. Wikipedia.
- Blaber, J., Adair, B., Antoniou, A., 2015. Ncorr: Open-Source 2D Digital Image Correlation Matlab Software. *Exp. Mech.* 55, 1105–1122. <https://doi.org/10.1007/s11340-015-0009-1>
- Bromiley, P.A., 2013. Products and convolutions of Gaussian probability density functions (Internal Report No. Tina Memo No. 2003-003). Medical School, University of Manchester, Manchester.
- Burger, W., 2016. Zhang's camera calibration algorithm: In-depth tutorial and implementation 56.
- Casas, J., 1994. Teoría de la imagen, in: *Óptica*. Librería PONS, Zaragoza (España), pp. 331–345.
- CCD o CMOS: ¿Es posible que los sensores CMOS sustituyan a los CCD en todos los casos?, 2016. *Revoluc. Artif.* URL <https://blog.infaimon.com/ccd-o-cmos-es-posible-sensores-cmos-sustituyan-ccd/> (accessed 12.4.20).
- Centroide, 2019. Wikipedia Encicl. Libre.
- Di Benedetto, C., Cappelletti, P., Favaro, M., Graziano, S.F., Langella, A., Calcaterra, D., Colella, A., 2015. Porosity as key factor in the durability of two historical building stones: Neapolitan Yellow Tuff and Vicenza Stone. *Eng. Geol.* 193, 310–319. <https://doi.org/10.1016/j.enggeo.2015.05.006>
- El sensor CMOS, 2016. *Top Fotogr.* URL <http://www.topfotografia.net/Fotografia/teoria-de-la-fotografia/sensor-cmos/sensor-cmos.html> (accessed 12.4.20).
- Encyclopedia of Laser Physics and Technology - Speckle, 2007. RP Photonics. URL <https://www.rp-photonics.com/speckle.html> (accessed 12.28.20).
- Espinosa, J., Ferrer, B., Mas, D., Pérez, J., Roig, A.B., 2014. Método y sistema para medir vibraciones.

- Espinosa, J., Perez, J., Ferrer, B., Mas, D., 2015. Method for targetless tracking subpixel in-plane movements. *Appl. Opt.* 54, 7760–7765. <https://doi.org/10.1364/AO.54.007760>
- Feng, D., Feng, M.Q., 2017. Experimental validation of cost-effective vision-based structural health monitoring. *Mech. Syst. Signal Process.* 88, 199–211. <https://doi.org/10.1016/j.ymssp.2016.11.021>
- Ferrer, B., Espinosa, J., Mas, D., 2019. A method to measure small local strains in concrete surfaces using its natural texture and image cross-correlation. *Struct. Control Health Monit.* 26. <https://doi.org/10.1002/stc.2410>
- Ferrer, B., Espinosa, J., Perez, J., Ivorra, S., Mas, D., 2011. Optical scanning for structural vibration measurement. *Res. Nondestruct. Eval.* 22, 61–75. <https://doi.org/10.1080/09349847.2010.519137>
- Ferrer, B., Espinosa, J., Roig, A.B., Perez, J., Mas, D., 2013. Vibration frequency measurement using a local multithreshold technique. *Opt. Express* 21, 26198. <https://doi.org/10.1364/OE.21.026198>
- Ferrer, B., Mas, D., 2018. Parametric evaluation of errors using isolated dots for movement measurement by image cross-correlation. *Sensors* 18, 525. <https://doi.org/10.3390/s18020525>
- Ferrer, B., Mas, D., García-Santos, J.I., Luzi, G., 2016. Parametric study of the errors obtained from the measurement of the oscillating movement of a bridge using image processing. *J. Nondestruct. Eval.* 35, 53. <https://doi.org/10.1007/s10921-016-0372-6>
- Ferrer, B., Tomás, M.-B., Mas, D., 2021. Use of image correlation to measure macroscopic strains by hygric swelling in sandstone rocks. *Appl. Sci.* 11, 2495. <https://doi.org/10.3390/app11062495>
- File:Resolution illustration.png, 2006. Wikipedia Free Encycl.
- Gaskill, J.D., 1978. The Fourier Transform, in: *Linear Systems, Fourier Transforms, and Optics*. John Wiley & Sons, Inc., United States of America, pp. 179–222.
- Goicochea, É.L., 2014. Centroides de cuerpos compuestos.
- GOM Correlate Pro: Strains and displacements from motion pictures, 2021. URL <https://www.gom.com/en/products/gom-suite/gom-correlate-pro> (accessed 2.28.22).
- Grehan, G., 2020. Efficient high-speed detection of sub-pixel motion. University College Dublin School of Electrical and Electronic Engineering.
- Hack, E., Leroy, D., 2005. Camera-based monitoring of the rigid-body displacement of a mandrel in superconducting cable production. *Opt. Lasers Eng.* 43, 455–474. <https://doi.org/10.1016/j.optlaseng.2004.05.004>
- Hayes, A., 2021. What is cross-correlation? Investopedia. URL <https://www.investopedia.com/terms/c/crosscorrelation.asp> (accessed 1.4.21).
- High-speed camera, 2021. Wikipedia Free Encycl.
- Introducción a la óptica telecéntrica, 2018. *Opto Eng.* URL <https://www.opto-e.es/resources/telecentric-lenses-tutorial> (accessed 12.23.20).
- Lava, P., 2015. Practical considerations in DIC measurements, in: *Image Correlation & Material Identification*. Presented at the 10th International Conference on Advances in Experimental Mechanics, MatchID - Metrology beyond colors, Edinburgh UK, p. 79.
- Lei, X., Jin, Y., Guo, J., Zhu, C., 2015. Vibration extraction based on fast NCC algorithm and high-speed camera. *Appl. Opt.* 54, 8198. <https://doi.org/10.1364/AO.54.008198>
- Los objetivos: Características principales, clases que disponemos y consejos útiles., 2016. *Carrete Digit.* URL <https://carredigital.com/los-objetivos-caracteristicas-principales-clases-disponemos-consejos-utiles/> (accessed 12.17.20).

- Maragos, P., 1989. Morphological correlation and mean absolute error criteria. Presented at the International Conference on Acoustics, Speech and Signal Processing, IEEE, Glasgow, UK, pp. 1568–1571.
- Mas, D., Espinosa, J., Roig, A.B., Ferrer, B., Pérez, J., 2012a. Use of subpixel techniques in pocket cameras to measure vibrations and displacements, in: Schelkens, P., Ebrahimi, T., Cristóbal, G., Truchetet, F., Saarikko, P. (Eds.), Presented at the SPIE Photonics Europe, Brussels, Belgium, p. 84360P. <https://doi.org/10.1117/12.922247>
- Mas, D., Espinosa, J., Roig, A.B., Ferrer, B., Perez, J., Illueca, C., 2012b. Measurement of wide frequency range structural microvibrations with a pocket digital camera and sub-pixel techniques. *Appl. Opt.* 51, 2664. <https://doi.org/10.1364/AO.51.002664>
- Mas, D., Ferrer, B., Acevedo, P., Espinosa, J., 2016a. Methods and algorithms for video-based multi-point frequency measuring and mapping. *Measurement* 85, 164–174. <https://doi.org/10.1016/m.measurement.2016.02.042>
- Mas, D., Perez, J., Ferrer, B., Espinosa, J., 2016b. Realistic limits for subpixel movement detection. *Appl. Opt.* 55, 4974–4979. <http://dx.doi.org/10.1364/AO.55.004974>
- MATLAB, R2020a. The MathWorks, Inc.
- Michaelis, D., Neal, D.R., Wieneke, B., 2016. Peak-locking reduction for particle image velocimetry. *Meas. Sci. Technol.* 27, 104005. <https://doi.org/10.1088/0957-0233/27/10/104005>
- Murray, C.A., Hoult, N.A., Take, W.A., 2017. Dynamic measurements using digital image correlation. *Int. J. Phys. Model. Geotech.* 17, 41–52. <https://doi.org/10.1680/jphmg.15.00055>
- Nobach, H., Damaschke, N., Tropea, C., 2005. High-precision sub-pixel interpolation in particle image velocimetry image processing. *Exp. Fluids* 39, 299–304. <https://doi.org/10.1007/s00348-005-0999-z>
- Nobach, H., Honkanen, M., 2005. Two-dimensional Gaussian regression for sub-pixel displacement estimation in particle image velocimetry or particle position estimation in particle tracking velocimetry. *Exp. Fluids* 38, 511–515. <https://doi.org/10.1007/s00348-005-0942-3>
- Nogueira, J., Lecuona, A., Nauri, S., Legrand, M., Rodríguez, P.A., 2011. Quantitative evaluation of PIV peak locking through a multiple Δt strategy: relevance of the rms component. *Exp. Fluids* 51, 785–793. <https://doi.org/10.1007/s00348-011-1094-2>
- O’Gorman, L., Bruckstein, A.M., Bose, C.B., Amir, I., 1990. Subpixel registration using a concentric ring fiducial, in: 10th International Conference on Pattern Recognition. Presented at the International Conference on Pattern Recognition, IEEE, Atlantic City, NJ, USA, pp. 249–253. <https://doi.org/10.1109/ICPR.1990.119365>
- Overmars, E.F.J., Warncke, N.G.W., Poelma, C., Westerweel, J., 2010. Bias errors in PIV: the pixel locking effect revisited. Presented at the 15th International Symposium on Applications of Laser Techniques to Fluid Mechanics, Lisbon, Portugal, p. 10.
- Ozana, N., Margalith, I., Beiderman, Y., Kunin, M., Campino, G.A., Gerasi, R., Garcia, J., Mico, V., Zalevsky, Z., 2015. Demonstration of a remote optical measurement configuration that correlates with breathing, heart rate, pulse pressure, blood coagulation, and blood oxygenation. *Proc. IEEE* 103, 248–262. <https://doi.org/10.1109/JPROC.2014.2385793>
- Pearson, J.J., Hines Jr., D.C., Golosman, S., Kuglin, C.D., 1977. Video-rate image correlation processor, in: Tescher, A.G. (Ed.), Presented at the 21st Annual Technical Symposium, San Diego, pp. 197–205. <https://doi.org/10.1117/12.955714>
- Pedrotti, F.L., Pedrotti, L.S., 1993. Optical Instrumentation, in: Introduction to Optics. Prentice Hall, New Jersey, pp. 109–150.

- Pérez Benito, C., 2015. Diseño de filtros para el procesamiento de imágenes basados en teoría de Grafos. Universitat Politècnica de València.
- Pixel, 2021. Wikipedia Free Encycl.
- Previous DIC Challenge 1.0 Data, 2020. IDICs Soc. Exp. Mech. URL <https://idics.org/challenge/> (accessed 10.20.21).
- Pueo, B., 2016. High speed cameras for motion analysis in sports science. *J. Hum. Sport Exerc.* 11. <https://doi.org/10.14198/jhse.2016.111.05>
- Reu, P.L., Toussaint, E., Jones, E., Bruck, H.A., Iadicola, M., Balcaen, R., Turner, D.Z., Siebert, T., Lava, P., Simonsen, M., 2018. DIC Challenge: Developing images and Gguidelines for evaluating accuracy and resolution of 2D analyses. *Exp. Mech.* 58, 1067–1099. <https://doi.org/10.1007/s11340-017-0349-0>
- Rieger, B., Nieuwenhuizen, R., Stallinga, S., 2015. Image processing and analysis for single-molecule localization microscopy: Computation for nanoscale imaging. *IEEE Signal Process. Mag.* 32, 49–57. <https://doi.org/10.1109/MSP.2014.2354094>
- Roche, R., 2013. Cámaras Alta Velocidad.
- Roesgen, T., 2003. Optimal subpixel interpolation in particle image velocimetry. *Exp. Fluids* 35, 252–256. <https://doi.org/10.1007/s00348-003-0627-8>
- Russ, J.C., 2011. The image processing handbook, Sixth edition. ed. CRC Press, United States of America.
- Ryle, J.P., Vohnsen, B., Sheridan, J.T., 2015. Simultaneous drift, microsaccades, and ocular microtremor measurement from a single noncontact far-field optical sensor. *J. Biomed. Opt.* 20, 027004. <https://doi.org/10.1117/1.JBO.20.2.027004>
- Schneider, C.A., Rasband, W.S., Eliceiri, K.W., 2012. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods, Historical Commentary* 9, 671–675. <https://doi.org/10.1038/nmeth.2089>
- Sensor CMOS, 2020. Wikipedia Encicl. Libre.
- Shortis, M.R., Clarke, T.A., Robson, S., 1995. Practical testing of the precision and accuracy of target image centering algorithms, in: El-Hakim, S.F. (Ed.), . Presented at the Photonics East '95, Philadelphia, PA, pp. 65–76. <https://doi.org/10.1117/12.220925>
- Shortis, M.R., Clarke, T.A., Short, T., 1994. A comparison of some techniques for the subpixel location of discrete target images, in: El-Hakim, S.F. (Ed.), Presented at the Photonics for Industrial Applications, Boston, MA, pp. 239–250. <https://doi.org/10.1117/12.189136>
- Siegesmund, S., Dürrast, H., 2011. Physical and mechanical properties of rocks, in: Siegesmund, S., Snethlage, R. (Eds.), *Stone in Architecture*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 97–225. https://doi.org/10.1007/978-3-642-14475-2_3
- Sjödahl, M., 1997. Accuracy in electronic speckle photography. *Appl. Opt.* 36, 2875. <https://doi.org/10.1364/AO.36.002875>
- Sjödahl, M., Benckert, L.R., 1994. Systematic and random errors in electronic speckle photography. *Appl. Opt.* 33, 7461–7472. <https://doi.org/10.1364/AO.33.007461>
- Speckle pattern, 2020. Wikipedia Free Encycl.
- Stanier, S., Dijkstra, J., Leśniewska, D., Hambleton, J., White, D., Muir Wood, D., 2016. Vermiculate artefacts in image analysis of granular materials. *Comput. Geotech.* 72, 100–113. <https://doi.org/10.1016/j.compgeo.2015.11.013>
- Sundararajan, D., 2017. *Digital Image Processing*. Springer Singapore, Singapore. <https://doi.org/10.1007/978-981-10-6113-4>
- Superresolución, 2019. Wikipedia Encicl. Libre.

- Sutton, M.A., Orteu, J.-J., Schreier, H.W., 2009. Image correlation for shape, motion and deformation measurements. Springer US, Boston, MA. <https://doi.org/10.1007/978-0-387-78747-3>
- Temperatura de color: (Cálido, Neutro, Frío), 2006. Muchos Leds. URL <https://muchosleds.com/es/content/6-temperatura-de-color-calido-neutro-frio> (accessed 12.22.20).
- Thin-plate smoothing spline - MATLAB tpaps - MathWorks España, 2006. URL <https://es.mathworks.com/help/curvefit/tpaps.html>
- Tipos de objetivos fotográficos: Guía teórica y práctica, 2010. DeCamaras - Fotogr. Digit. URL <https://www.decamaras.com/CMS/content/view/861/61-Tipos-de-objetivos-fotograficos-Guia-teorica-y-practica> (accessed 12.16.20).
- Tomás, M.-B., Ferrer, B., Mas, D., 2020. Influence of neighborhood size and cross-correlation peak-fitting method on location accuracy. *Sensors* 20, 6596. <https://doi.org/10.3390/s20226596>
- Wahbeh, A.M., Caffrey, J.P., Masri, S.F., 2003. A vision-based approach for the direct measurement of displacements in vibrating systems. *Smart Mater. Struct.* 12, 785–794. <https://doi.org/10.1088/0964-1726/12/5/016>
- Wan, M., Healy, J.J., Sheridan, J.T., 2020. Fast, sub-pixel accurate, displacement measurement method: optical and terahertz systems. *Opt. Lett.* 45, 6611. <https://doi.org/10.1364/OL.413011>
- Yang, J., Huang, T., 2017. Image super-resolution: Historical overview and future challenges, in: Milanfar, P. (Ed.), *Super-Resolution Imaging*. CRC Press, pp. 1–33. <https://doi.org/10.1201/9781439819319-1>
- Yu, W., Xie, J., Wu, X., Liu, K., 2017. Research of improved Zhang's calibration method 5.
- Zhou, Y., Sun, C., Song, Y., Chen, J., 2015. Image pre-filtering for measurement error reduction in digital image correlation. *Opt. Lasers Eng.* 65, 46–56. <https://doi.org/10.1016/j.optlaseng.2014.04.018>

ANEXO I: Cámaras y objetivos

AI.1. Resoluciones espacial y temporal

De acuerdo con la explicación dada en el Capítulo 1.1, el tamaño del píxel de una cámara digital es importante a la hora de detectar posiciones de objetos con exactitud, ya que, si los píxeles del sensor de la cámara son relativamente grandes, dos detalles del objeto de interés que estén cercanos entre sí podrían coincidir dentro del mismo píxel y no se podría resolver la distancia entre ambos detalles en la escena. Este tamaño suele estar relacionado con la resolución espacial, de modo que, a mayores resoluciones espaciales, menor será, generalmente, el tamaño del píxel. De esta forma, cámaras con mayor resolución espacial serán, en principio, mucho mejores ya que no solamente tendrán más cantidad de píxeles en el sensor, sino que al ser estos más pequeños, estarán más juntos unos de otros y la imagen digital obtenida tendrá una resolución muy superior.

Como cada píxel solo puede recoger una información homogénea en toda su superficie, en la figura AI.1 se puede ver gráficamente que, conforme va aumentando el número de píxeles manteniéndose el mismo tamaño del sensor (por tanto, disminuyendo el tamaño de los píxeles), el detalle del objeto se consigue apreciar mejor. Al tener la misma información en toda la superficie del píxel, si todo el objeto (en este caso, la letra *R*) coincide dentro del mismo elemento sensor de la cámara, la imagen digital obtenida será un único color uniforme. Conforme va aumentando el tamaño de la matriz de fotorreceptores y, de la misma forma, se van haciendo cada vez más pequeños, el objeto comienza a verse cada vez más nítido.

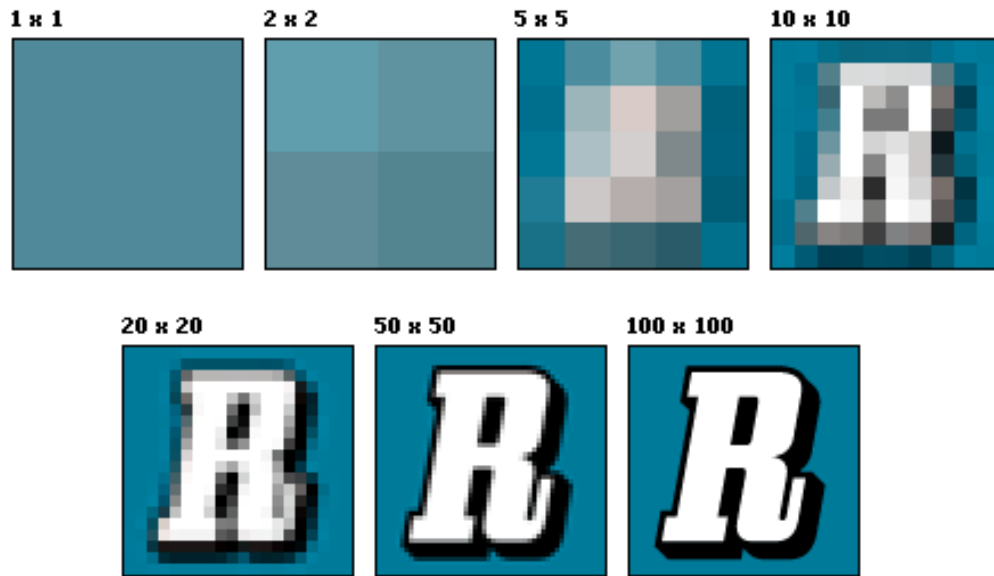


Figura AI.1. Sobre cada imagen se muestra el tamaño de la matriz de píxeles existente en el sensor. A mayor número de píxeles, manteniendo el tamaño del sensor de la cámara, se obtiene una imagen más nítida debido a que la información recogida por cada píxel es uniforme en toda su superficie (“File,” 2006).

En relación con la resolución temporal, cuando se realizan medidas de objetos cuyos movimientos o vibraciones son rápidos, es necesario tener una resolución temporal alta para capturar procesos rápidos y no perder información en cada momento.

No siempre es necesario tener ambas resoluciones altas. En ocasiones se tendrán objetos con detalles precisos pero cuyo posicionamiento exacto en cada momento no sea necesario conocer. En ese caso, la resolución espacial deberá de ser muy buena, pero no se requerirá una resolución temporal demasiado alta. Por otra parte, se puede tener la situación contraria en la que los detalles del objeto no sean importantes, pero sí que sea requisito indispensable conocer la posición exacta del mismo en cada momento. En esta ocasión, la resolución temporal deberá ser lo más alta posible, mientras que se podrá prescindir de una excelente resolución espacial.

Habitualmente, las resoluciones espacial y temporal de las cámaras están inversamente relacionadas debido a las limitaciones del sistema de transferencia, procesado y almacenamiento, por lo que antes de comenzar a realizar una medida para una detección subpíxel será necesario conocer las características del objeto y la velocidad del movimiento esperado para mejorar las resoluciones según interese sin obtener una cantidad de datos más alta de lo necesario. Esta parte es importante tenerla en cuenta porque las cámaras de alta velocidad generan imágenes que utilizan una gran cantidad de espacio en los ordenadores. Por ello, generar imágenes extra innecesarias (grabando a una frecuencia temporal más alta de lo indispensable) o más pesadas de lo imprescindible (grabando no solo el objeto de interés, sino también el fondo; utilizando una frecuencia espacial demasiado alta), va a consumir una gran cantidad de espacio que podría ralentizar los equipos. Además, producir más imágenes de las necesarias o demasiado pesadas puede generar problemas en las grabaciones debido a que el búfer de la cámara puede no ser suficientemente grande para almacenarlas todas y se perderían imágenes en el proceso de grabación y traspaso de datos

al ordenador. Por todo ello, también es importante conocer las características de la cámara y utilizar aquella que más se vaya a ajustar a las necesidades del objeto de interés de cada estudio.

AI.2. Características de las cámaras

En las cámaras digitales, es el sensor el que se encarga de capturar la información que le llega y transformarla en una imagen digital para poder ser utilizada en ordenadores y equipos digitales. Los sensores más comunes en las cámaras fotográficas y de vídeo son CCD (Charge-Coupled Device) y CMOS (Complementary Metal Oxide Semiconductor). Ambos están formados básicamente por semiconductores de metal-óxido y están distribuidos en forma de matriz. Sin embargo, la principal diferencia entre ellos es que, en el caso del CCD, primero se transporta la información capturada en el sensor fuera de él y después se transforma a información digital, mientras que en el CMOS, la conversión de la información se realiza en cada píxel dentro del mismo sensor. Esta diferencia implica que el sensor CCD obtenga una calidad de imagen mayor con un bajo nivel de ruido, mientras que el sensor CMOS consiga elevadas tasas de captura y tenga menos problemas de sobreexposición ya que cada píxel tiene su propia información y no se puede intercambiar con los píxeles vecinos, aunque reducir el ruido en estos sensores es más complejo. Además, los sensores CMOS son más baratos de fabricar y consumen menos energía que los CCD. Por todo ello, no es mejor un sensor u otro, sino que tienen distintas características y en cada estudio en particular pueden interesar más unas u otras (“CCD o CMOS,” 2016; “El sensor CMOS,” 2016; “Sensor CMOS,” 2020).

Asimismo, las cámaras digitales a color estándar utilizan una distribución de píxeles en el sensor como la que se muestra en la figura AI.2, denominada mosaico de Bayer, en la que cada píxel recoge la información de un color de los tres primarios (rojo, verde y azul). De este modo, con cuatro píxeles se obtiene la información total del color, quedando un píxel con la información del color rojo, otro con la del color azul y dos píxeles que recogen información para el color verde. Por este motivo, cuando sea interesante obtener la información grabada con una cámara a color en escala de grises para su posterior procesado de imágenes, si la imagen no tiene un color predominante, sino que tiene una distribución homogénea de colores tanto en el objeto como en el fondo, la mejor opción será seleccionar la información del canal verde ya que se obtendrá la mitad de la información total de la imagen, mientras que, si se seleccionaran los canales rojo o azul, se obtendría solamente un cuarto de toda la información de la escena.

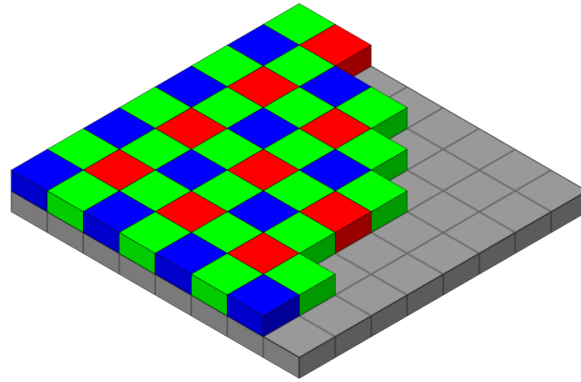


Figura AI.2. Mosaico de Bayer presente en los sensores de las cámaras a color.

En la figura AI.3 se han separado los tres canales de una imagen grabada a color para visualizar el efecto explicado sobre un test USAF normalizado (“1951 USAF resolution test chart,” 2021). Como se puede observar, en la imagen del canal verde se distinguen como tres bandas separadas algunos elementos con tamaños menores que para las imágenes de los canales rojo y azul. En cualquier caso, si la información del color no es relevante, siempre será mejor usar un sensor en escala de grises.

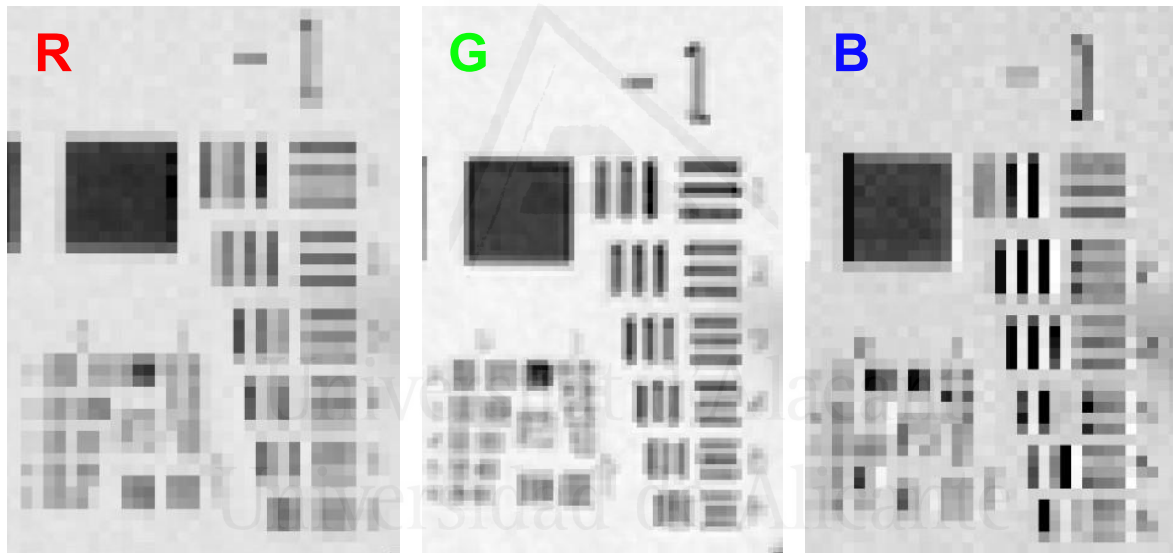


Figura AI.3. Imagen del test USAF separada por canales para mostrar que como el canal verde (G) recoge el doble de información que los canales rojo (R) y azul (B), se consigue una mayor resolución de la cámara en este canal.

En cuanto a los parámetros de las cámaras que suelen proporcionar los fabricantes se encuentran el tamaño del sensor, el tamaño del píxel y la resolución espacial máxima de la cámara. Estos tres parámetros están relacionados entre sí, ya que, conociendo dos de ellos, se puede calcular el tercero. El tamaño del sensor proporciona los tamaños horizontal y vertical en milímetros, el tamaño del píxel es lo que mide cada píxel en micrómetros y la resolución espacial máxima de la cámara es el número total de píxeles del sensor en horizontal y en vertical. La relación de aspecto de las cámaras suele ser 4:3, que indica la relación entre el alto y el ancho de la imagen.

Por otra parte, la resolución temporal también es un parámetro a conocer porque indica la máxima velocidad de grabación; los fotogramas por segundo máximos que podremos

obtener con nuestra cámara. En general, las especificaciones de la cámara dan la resolución temporal máxima a “*Full-frame*”, es decir, con la máxima resolución espacial. En la mayoría de cámaras de tipo científico podemos aumentar la resolución temporal a costa de disminuir el tamaño del fotograma con lo que podremos adaptar la cámara a diferentes tipos de experimentos.

En las cámaras a color, otro parámetro a tener en cuenta es la profundidad de color, que indica los bits máximos que puede almacenar un único píxel. A mayor profundidad de color, más información se tendrá del objeto medido porque se podrán almacenar 2^{PC} niveles diferentes en cada píxel (siendo PC la profundidad de color). Sin embargo, el almacenaje de estas imágenes será mucho más costoso a medida que aumentamos la profundidad de color ya que cada fotograma tendrá una cantidad de información mucho mayor y, por tanto, un peso también mayor. Hay que tener en cuenta que aumentar en un bit la profundidad de color, duplica el número de niveles de gris que se consiguen medir con la cámara.

Todas estas características se deben conocer y tener en cuenta antes de comenzar a realizar un experimento para, sabiendo las características del objeto a grabar, seleccionar la cámara que mejores resultados vaya a proporcionar.

AI.3. Características de los objetivos

También juegan un papel importante los objetivos utilizados. Estos son sistemas ópticos formados por un conjunto de lentes convergentes y divergentes que consiguen hacer converger los rayos de luz procedentes del objeto, formando una imagen óptica en el sensor de la cámara (Pedrotti and Pedrotti, 1993; “Tipos de objetivos fotográficos: Guía teórica y práctica,” 2010; “Los objetivos,” 2016).

Las características principales que definen un objetivo ópticamente son:

- *Distancia focal*
Es la distancia en milímetros desde el plano principal imagen del objetivo hasta el plano donde se van a enfocar los rayos que proceden del infinito (es decir, el plano donde se encuentra el sensor de la cámara). Al aumentar la distancia focal, la imagen capturada en el sensor de la cámara será más grande, pero el campo angular será menor, es decir, el campo de visión capturado por la cámara será menor.
- *Apertura Numérica (AN)*
Corresponde a la división de la focal entre el diámetro efectivo de la lente frontal, por lo que indica la luminosidad del objetivo. Se escribe como $f/4$ (aunque también puede verse escrito como $f:4$ o $1:4$), y, en este ejemplo, indicaría que el diámetro de la lente es una cuarta parte de la distancia focal. A mayor valor de AN, menor será la abertura en el objetivo, aunque también se podrá reducir cuando sea necesario con el diafragma iris que incorporan la mayoría de objetivos. No obstante, nunca se podrá aumentar la abertura de un objetivo a un

tamaño mayor del que viene indicado por la apertura numérica, ya que es su máximo.

- *Objetivo zoom*

Permite obtener un rango de focales y no una sola focal fija al posibilitar el movimiento de las lentes del objetivo de forma que, variando la separación entre ellas, se cambian las características del sistema óptico. De esta manera, se pueden hacer converger los rayos de luz procedentes de un objeto situado a otra distancia en el mismo elemento sensor de la cámara.

- *Profundidad de campo y profundidad de foco*

Se dice que un sistema óptico es estigmático cuando todos los rayos que salen de un punto objeto y pasan por el sistema se juntan exactamente en el punto imagen (Casas, 1994). Si el sensor de la cámara no está exactamente en el foco del objetivo, sino que está ligeramente delante o detrás de él o el sistema óptico no es perfecto y tiene aberraciones, en vez de obtener un punto imagen en el sensor, se obtendrá una mancha circular denominada círculo de confusión. Además, dado que los píxeles del sensor de la cámara tienen un tamaño finito, todas aquellas posiciones del objeto que generen un círculo menor que dicho tamaño producirán una imagen similar. Mientras el diámetro de este círculo de confusión sea lo suficientemente pequeño, la imagen continuará siendo aceptable; por tanto, hay un rango de distancias en las que el objeto aparecerá enfocado a pesar de no estar su imagen exactamente en el sensor de la cámara. Ese rango de distancias en las que el objeto aparece enfocado se denomina profundidad de campo en el espacio objeto y profundidad de foco en el espacio imagen (figura AI.4).

Aparte de con el tamaño del píxel, la profundidad de foco variará con el tamaño del diafragma, ya que, si el diafragma está bastante cerrado, los rayos que llegarán a la imagen procedentes del objeto solo serán aquellos que estén más cerca del eje óptico del sistema. Al formar estos rayos un cono muy estrecho, habrá un rango de distancias para las que la intersección del cono con el plano imagen producirá un círculo de desenfoque más pequeño que el tamaño del píxel de la cámara, estando el objeto enfocado para ese rango de distancias. Lo contrario ocurre cuando el diafragma del objetivo está más abierto y los rayos procedentes del objeto llegan más alejados del eje óptico del sistema. Por esta razón, se consigue aumentar la profundidad de foco del sistema al disminuir el diafragma, obteniendo un rango mayor de objetos enfocados en el espacio objeto, por lo que aumenta también la profundidad de campo (figura AI.5).

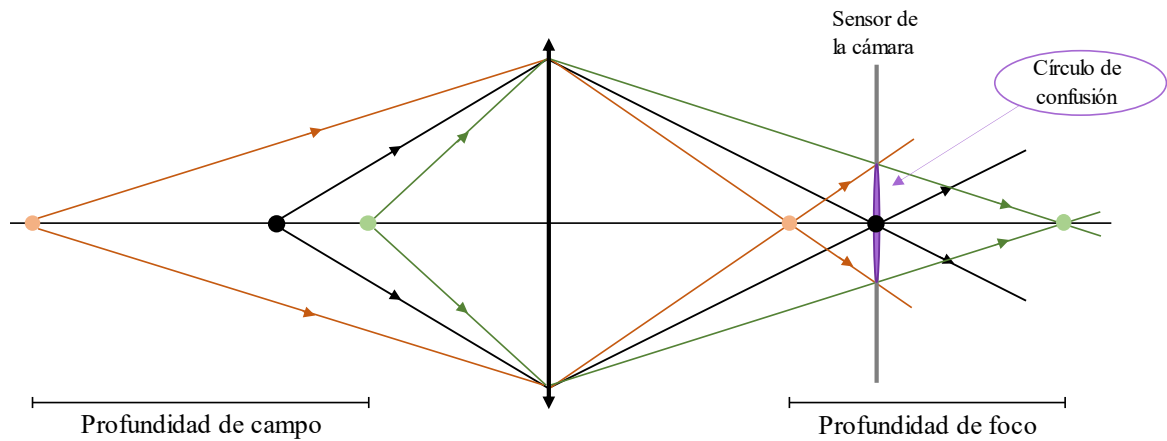


Figura AI.4. Descripción gráfica simplificada de la profundidad de campo y de foco. Se ha sustituido todo el objetivo de la cámara por una sola lente convergente para mejorar su visualización.

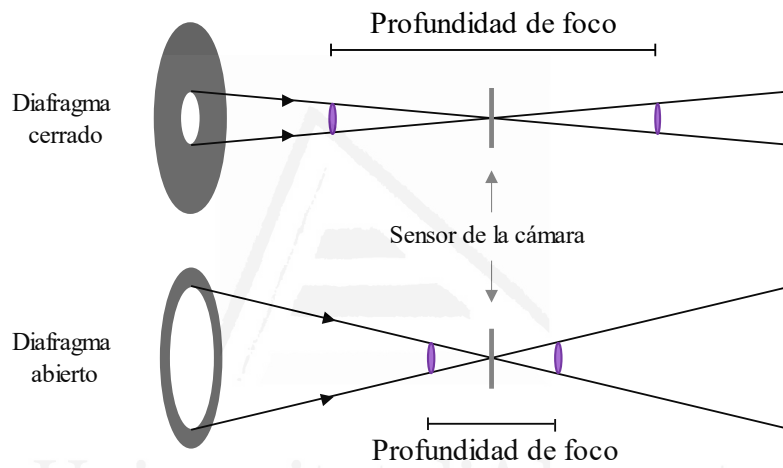


Figura AI.5. Descripción gráfica de la relación del tamaño del diafragma con la profundidad de foco (y, por tanto, también con la profundidad de campo en el espacio objeto).

Universidad de Alicante

ANEXO II: Características y uso de las cámaras

Los sensores de las cámaras pueden variar en tamaño y dos cámaras cuyas estructuras las hacen parecer iguales, pueden tener unas características muy diferentes. Por ejemplo, en la figura AII.1 se muestran dos cámaras por su parte delantera para poder comparar el tamaño de los sensores (tras la película roja) a pesar de que sus armazones son prácticamente del mismo tamaño. Por otro lado, en la parte trasera de la cámara (figura AII.2) se encuentran la conexión por cable para poder transferir los datos hacia el ordenador para su posterior almacenamiento. La parte circular es la que permite la conexión a la electricidad para que pueda ponerse en funcionamiento la cámara y la otra abertura permite la conexión con el ordenador para realizar la transferencia de información. Actualmente se suelen utilizar conexiones USB 3.0 (figura AII.2 derecha), pero también existen cámaras más antiguas que utilizan un sistema por cable Gigabit Ethernet (figura AII.2 izquierda) que es una conexión más lenta y puede producir más errores a la hora de medir y almacenar las imágenes en el ordenador.



Figura AII.1. Cámaras BASLER mostradas por la parte delantera por donde se puede observar la diferencia en el tamaño del sensor.



Figura AII.2. Cámaras BASLER con conexión Gigabit Ethernet (izquierda) y USB 3.0 (derecha) para la transferencia de datos al ordenador.

En los experimentos realizados en la presente tesis doctoral se han utilizado las dos cámaras BASLER cuyas características se describen en la tabla AII.1. La primera de ellas ha sido utilizada para la medida de la expansión hídrica de una roca arenisca (Capítulo 4.1) y para la evaluación de la textura natural del hormigón como diana en medidas a través de la correlación cruzada (Capítulo 4.2) y la segunda cámara se ha utilizado para la comprobación del método subpíxel por análisis de luminancia utilizando secuencias reales (Capítulo 5.3).

Tabla AII.1. Características de las cámaras BASLER utilizadas en los experimentos realizados en la presente tesis doctoral.

| | BASLER acA4600-10uc | BASLER acA800-510um |
|-------------------------------|---------------------|---------------------|
| Tipo de sensor | CMOS | CMOS |
| Tamaño del sensor | 6.5 × 4.6 mm | 3.8 × 2.9 mm |
| Resolución máxima | 4608 × 3288 px | 800 × 600 px |
| Tamaño del píxel | 1.4 × 1.4 μm | 4.8 × 4.8 μm |
| Máxima velocidad de grabación | 10 fps | 511 fps |
| Profundidad de color | 12 bits | 10 bits |

Para manejar estas cámaras y poder grabar con ellas, se necesita instalar el programa *pylon Viewer* en un ordenador. Así, se pueden modificar los diferentes parámetros de la cámara y permite visualizar el objeto a medir para poder enfocararlo adecuadamente antes de comenzar a grabar, centrar la imagen, superponer una cuadrícula para alinear correctamente el sistema, etc. En la figura AII.3 se muestra la interfaz principal del programa con un objeto enfocado. A continuación, en la figura AII.4 se muestra el cuadro de diálogo en el que se selecciona la ubicación donde guardar las imágenes y la frecuencia de grabación, así como el tiempo o las imágenes al cabo del cual se desea detener la grabación, aunque también se puede detener la grabación manualmente en cualquier momento.

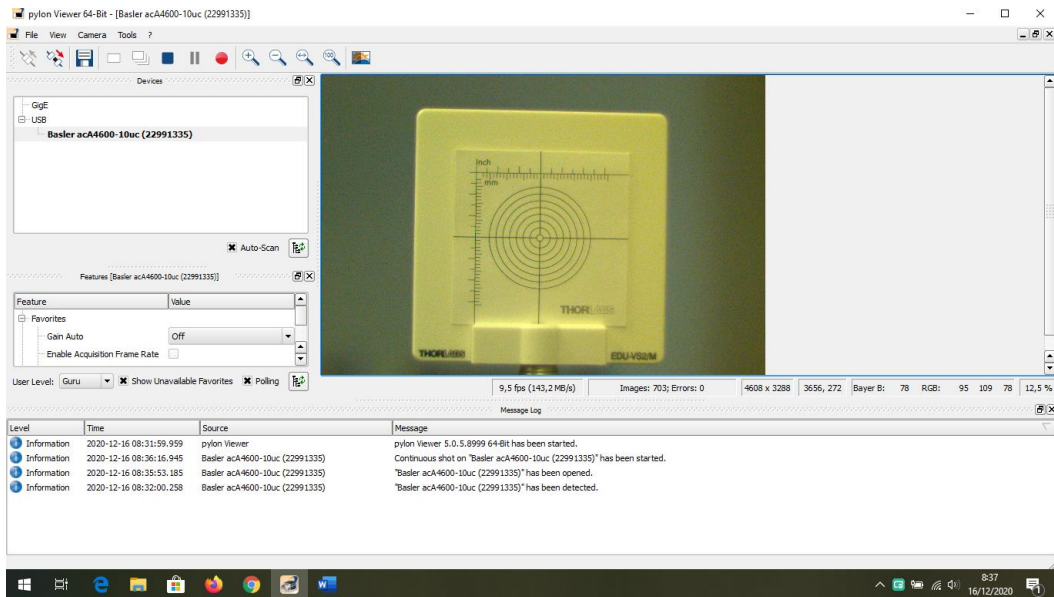


Figura AII.3. Interfaz principal del programa *pylon Viewer* para grabar con las cámaras BASLER, con un objeto enfocado.

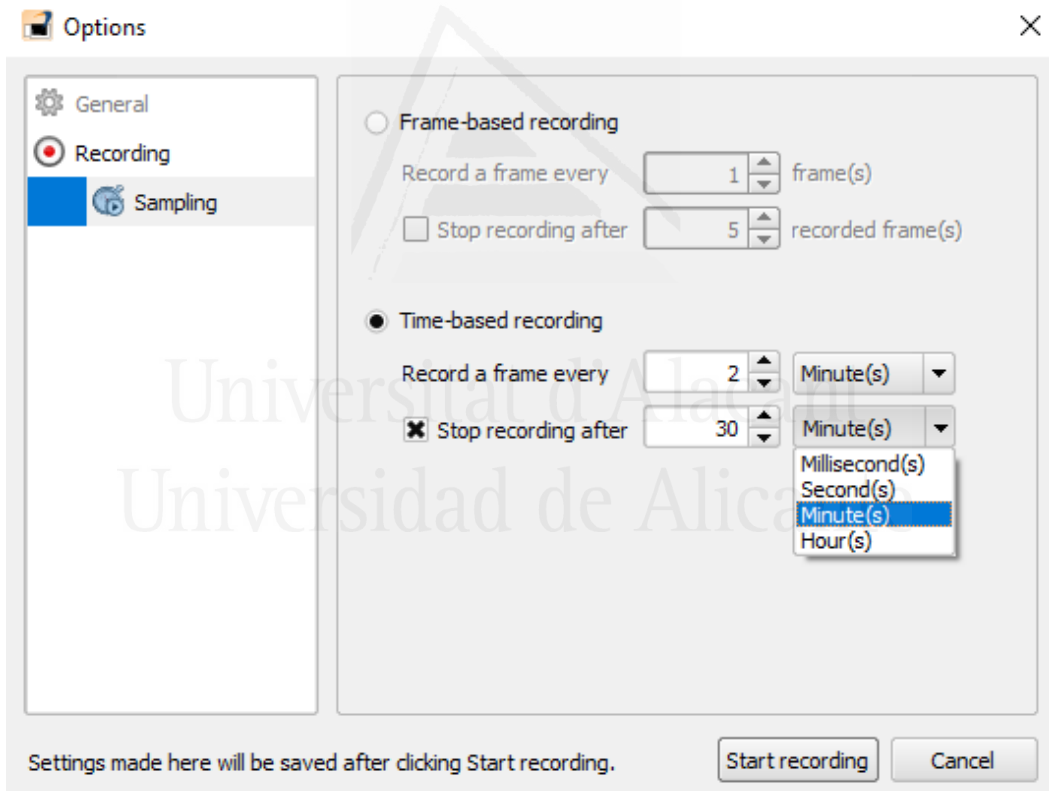


Figura AII.4. Selección de la ubicación donde guardar la grabación y sus características. Se puede seleccionar cada cuántos fotogramas o cada cuánto tiempo se desea medir y si se quiere que se detenga la grabación tras un número determinado de fotogramas o de tiempo.

ANEXO III: Sistemas de iluminación y cabinas de iluminación

Durante las medidas experimentales realizadas en el laboratorio, se han utilizado sistemas de iluminación para mantener la luminosidad del objeto constante y así, reducir las variables a tener en cuenta. De esta forma, las diferencias detectadas en el sensor entre unos fotogramas y otros se deberán al movimiento del objeto y no a factores como la iluminación ambiente o los cambios de luminosidad producidos en largas medidas por la luz exterior que entra por las ventanas del laboratorio y que va cambiando a lo largo del día o incluso entre un día y otro.

Generalmente, las lámparas utilizadas en los laboratorios permiten modificar la cantidad de luz que emiten proporcionando el porcentaje de luz emitida para poder mantener las condiciones de iluminación siempre constantes. Además, algunas de ellas también permiten modificar la temperatura de color para tener más datos sobre la fuente de iluminación ya que el iluminante modifica el color percibido del objeto, aunque en nuestro caso, este dato no es imprescindible siempre y cuando se mantenga constante utilizando la misma lámpara en todas las medidas del mismo experimento. En la figura AIII.1 se muestra una lámpara iluminando un objeto con dos temperaturas de color diferentes, pero manteniendo la misma cantidad de luz emitida para que se pueda observar la diferencia en el color percibido del objeto. Si se cambiara el color percibido del objeto durante el proceso de medida, podría ser un problema a la hora de medir su posición con las técnicas subpíxel utilizadas en la presente tesis doctoral. No obstante, siempre que el iluminante se mantenga constante y el contraste del objeto con el fondo sea adecuado para poder realizar un buen procesado de imágenes a posteriori, los valores de cantidad de luz y temperatura de color no será importante conocerlos.

La temperatura de color de una fuente de luz se define a partir del color que emite dentro del espectro luminoso comparado con el que emitiría un cuerpo negro calentado a una determinada temperatura. Esa temperatura se expresa en grados kelvin, aunque no expresa una medida de temperatura, sino una medida relativa. Las luces se dividen en tres grandes grupos de temperatura de color: luz cálida con temperaturas de color entre 2800 °K y 3500 °K, luz neutra con temperaturas de color entre 3800 °K y 4500 °K y luz fría con temperaturas de color mayores de 5000 °K (“Temperatura de color,” 2006).

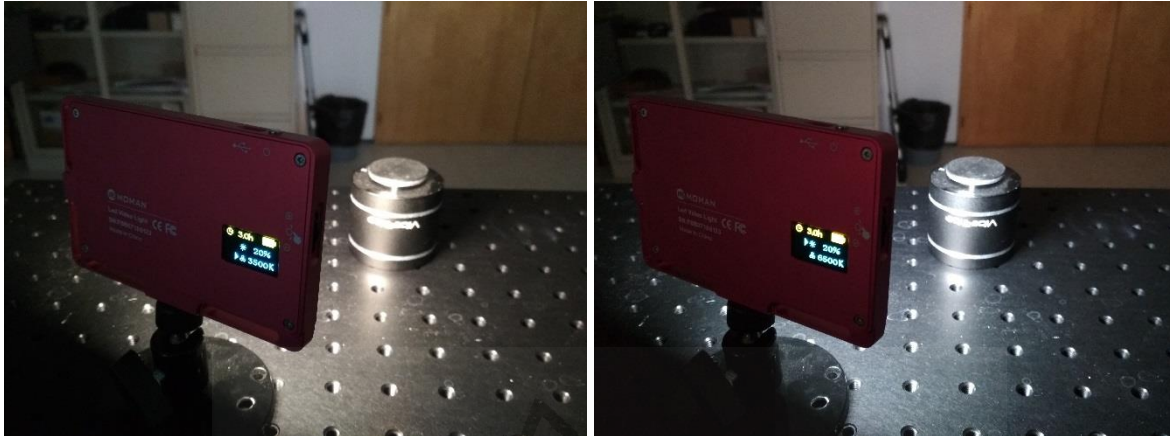


Figura AIII.1. Ejemplo de la misma fuente de iluminación emitiendo la misma cantidad de luz, pero con una temperatura de color diferente. A la izquierda se utilizan 3500 °K, mientras que a la derecha se utilizan 6500 °K.

Respecto a las cabinas de iluminación, se trata de una estructura con una iluminación constante en su interior de forma que el objeto de interés no tenga variaciones de iluminación o sombras producidas por la iluminación cambiante de la sala si el experimento tiene una duración relativamente larga. Pueden ser de diferentes tamaños y con una estructura más o menos rígida.

En el caso de la cabina de iluminación utilizada en algunos de los experimentos llevados a cabo en la presente tesis doctoral, la estructura no es demasiado rígida, por lo que se puede transportar fácilmente plegada y montar en el lugar que nos interese. Además, tiene las paredes de color gris reflectante para poder reflejar toda la luz interna de la cabina y generar así las mínimas sombras posibles. Las lámparas incluidas permiten variar la cantidad de luz emitida, pero manteniendo una temperatura de color fija. En la figura AIII.2 se puede observar la cabina de iluminación descrita con un montaje en su interior para realizar un experimento.

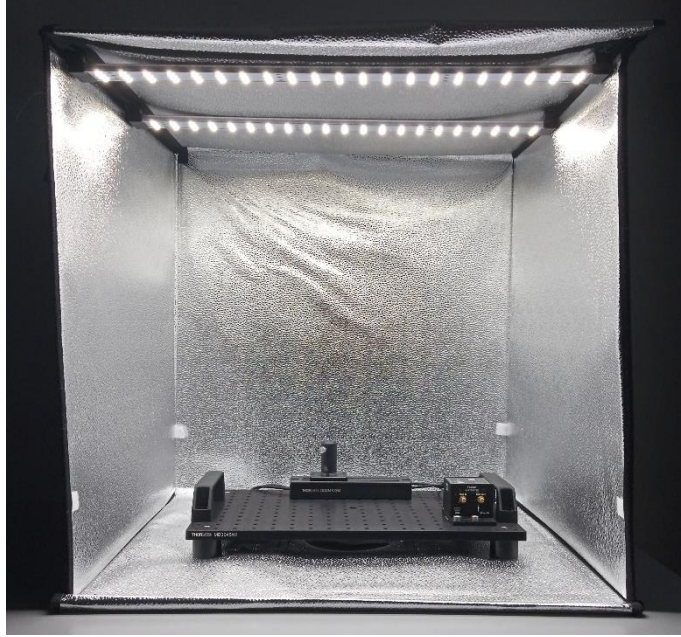


Figura AIII.2. Cabina de iluminación con un montaje en su interior.



Universitat d'Alacant
Universidad de Alicante

ANEXO IV: Publicaciones



Universitat d'Alacant
Universidad de Alicante

PUBLICACIÓN 1: Influence of neighborhood size and cross-correlation peak-fitting method on location accuracy

Tomás, M.-B.; Ferrer, B.; Mas, D. Influence of neighborhood size and cross-correlation peak-fitting method on location accuracy. *Sensors* **2020**, *20*, 6596, <https://doi.org/10.3390/s20226596>

Revista *Sensors* de MDPI (online)

ISSN: 1424-8220

Factor de Impacto: 3.275

JCR: Q1

Universitat d'Alacant
Universidad de Alicante



Article

Influence of Neighborhood Size and Cross-Correlation Peak-Fitting Method on Location Accuracy [†]

María-Baralida Tomás, Belén Ferrer and David Mas *

University Institute of Physics Applied to the Sciences and Technologies, University of Alicante, P.O. Box 99, 03080 Alicante, Spain; maria.baralida@ua.es (M.-B.T.); belen.ferrer@ua.es (B.F.)

* Correspondence: david.mas@ua.es; Tel.: +34-965-903-400

† This paper is an extended version of the conference paper: Tomás, M.B.; Mas, D.; Ferrer, B. Peak-locking minimization by three adjustment methods. In Proceedings of the Optics, Photonics and Digital Technologies for Imaging Applications VI; Schelkens, P., Kozacki, T., Eds.; SPIE: Jakarta, France, 2020; p. 53.

Received: 1 October 2020; Accepted: 16 November 2020; Published: 18 November 2020



Abstract: A known technique to obtain subpixel resolution by using object tracking through cross-correlation consists of interpolating the obtained correlation function and then refining peak location. Although the technique provides accurate results, peak location is usually biased toward the closest integer coordinate. This effect is known as the peak-locking error and it strongly limits this calculation technique's experimental accuracy. This error may differ depending on the scene and algorithm used to fit and interpolate the correlation peak, but in general, it may be attributed to a sampling problem and the presence of aliasing. Many studies in the literature analyze this effect in the Fourier domain. Here, we propose an alternative analysis on the spatial domain. According to our interpretation, the peak-locking error may be produced by a non-symmetrical sample distribution, thus provoking a bias in the result. According to this, the peak interpolant function, the size of the local domain and low-pass filters play a relevant role in diminishing the error. Our study explores these effects on different samples taken from the DIC Challenge database, and the results show that, in general, peak fitting with a Gaussian function on a relatively large domain provides the most accurate results.

Keywords: peak-locking; cross-correlation; subpixel; Gaussian fitting; thin-plate splines; polynomial fitting

1. Introduction

Cross-correlation is a useful technique for establishing similarity between two signals. As correlation can derive from minimizing the mean square error between two signals or images [1], it is a robust tool for comparing images corrupted by Gaussian noise, which is the normal case in most circumstances during image processing where illumination is good enough. Apart from giving a similarity metric, peak location describes the position where the reference image and template match present maximum coincidence and is, thus, often used for image aligning or object tracking in a scene.

Despite the many advantages and applications of the cross-correlation [2–6], its standard formulation presents two main drawbacks: the dependence of the correlation result on image and template amplitudes to, thus, produce high peaks when a dark template is compared to a bright object or vice versa, even though their similarity is minimum [7] and the limited resolution, which is set, by construction, to one pixel.

The first issue is solved by using a normalized cross-correlation algorithm, which is the common approach when dealing with images. Regarding the second issue, subpixel resolution can be achieved

by interpolation, which can be applied to either image before calculating their cross-correlation [8], the correlation function itself, to increase the accuracy in the location of its maximum [9]. The first approach is frequently followed to analyze deformations in solid materials as it allows for deformation mappings to be easily implemented [8]. The second one, i.e., peak interpolation, has faster and easier applications in non-deforming scenes and is, thus, adequate for aligning and tracking isolated objects. In our case, we pay attention to the second technique.

Briefly, the technique consists of interpolating the correlation function over a local area around the maximum peak and then refining the search by fitting the peak neighborhood to an analytical function [8]. This procedure may increase peak location accuracy by almost two orders of magnitude [10, 11]. Despite the evident improvement, the procedure also introduces a bias error, which limits its performance. The error, known as the peak-locking or pixel-locking effect, means that the peak location obtained through local fitting is always biased toward the closest integer coordinate [12].

The origin of peak-locking has been usually attributed to an aliasing effect due to a poor image texture [13] combined with an inadequate choice of the interpolant function [14]. The problem of the aliasing can be avoided with adequate sensors and lenses. In general, according to the Nyquist limit, a pseudospeckle scene will be well sampled when the dot unit is larger than 2 px [15].

Properly choosing the interpolating function is a more delicate issue and can be better explained in the spatial domain. Consider a scene with a low-noise object and a template containing a shifted version of the object. The finer the details of the object, the narrower the correlation peak will be, since a small displacement will degrade the correspondence between object and scene [15]. If one considers a narrow local domain around the maximum of the correlation function, there will be a reduced number of samples to fit to the interpolant function. Therefore, this maximum may have an excessive weight in the fitting, and it may pull the recalculated maximum to its position in the original grid. Consequently, a bias area towards the nearer integer corresponding to the original maximum position is introduced. Notice that when the maximum is exactly in the middle of two pixels the weight distribution is balanced, and the error is 0.

One can take a larger neighborhood and thus a larger number of samples to fit in order to compensate for the excessive weight of the correlation maximum, but this would eventually include information of non-correlated positions and thus distort the final result. Additionally, since the number of samples is still not very large, any secondary peak in the neighborhood will also unbalance the fitting and pull the fitted maximum towards it.

According to this, fitting functions that only consider the peak area, i.e., quadratic functions, may show a good fitting with a small neighborhood, but would be affected by the peak-locking effect. An extensive function that considers the peak and the region around, i.e., Gaussian fitting, would compensate for this effect but would need a larger neighborhood. Therefore, a proper selection of the interpolant function together with the interpolation domain is critical to decrease the peak-locking error.

In [16], the authors proposed different interpolation algorithms in an 8×8 neighborhood, and showed that the Gaussian function provided better results than the bilinear one, third-order polynomial or bicubic splines.

An additional strategy consists of filtering out the finer details of the sequence being analyzed, so that the correlation peak is softer and thus more suitable to a fitting by an analytical and derivable function. In [17], the authors introduced a defocus into the image-capturing process with an effect of reducing the peak-locking effect. This blurring was introduced experimentally by manipulating the objective. By doing so, and by adjusting the correlation peak through a Gaussian function, the results were slightly improved compared with the non-filtered results. This effect was thoroughly analyzed in [18]. Michaelis et al. [19] tested a different configuration for blurring an image using an optical diffuser. They also implemented two different interpolation functions (splines and bicubic), which gave good results for very small particle sizes.

Another strategy to diminish the peak error consists of maximizing the dynamic range of the image and template [20]. It has been shown that the accuracy in object-tracking tasks is directly related to the number of gray levels of the image [10,19]. Nevertheless, images and their luminance dynamic range are linked to the experimental setup. Hence, albeit important, it is not often a parameter that we can modify at will.

In this manuscript, we propose a combined analysis of three factors that may help to compensate for the peak-locking error and increase the accuracy of the proposed methods. Therefore, we will analyze the interaction between the two mentioned fitting functions, Gaussian and quadratic, in order to analyze their dependence with the neighborhood and its capability to reduce the error. The analysis will be complemented with the analysis of the results obtained using spline fitting. These fitting functions are more adaptable than the other two, so they may be useful in a wide range of neighborhood sizes. Additionally, the effect of a Gaussian defocus on the scene and the template is also analyzed together with the interaction with the fitting function.

The final aim of this paper is to determine which is the best method, which includes fitting function, application domain and amount of defocus, for obtaining object displacement with reduced peak-locking error. Due to the large amount of variability in images, we took a set of speckle sequences from the 2D-DIC public images bank from the Society for Experimental Mechanics (SEM) [21]. We selected the first image as a reference and the texture on it was tracked through the sequence.

Preliminary results of this analysis were presented at the SPIE Photonics Meeting 2020 [22], concluding that quadratic functions are more suitable for small neighborhoods, while Gaussian functions stand for large neighborhoods without further analysis. In what follows, we will explain the reasons for that and a final rule of thumb.

2. Methods

The purpose of the simulation is to analyze the relationship between three different peak-fitting methods and the neighborhood size together with the influence of the low-pass filters on the peak-locking method. In order to make our conclusions more general and facilitate the reproducibility of our results, we have checked our method with synthetical images taken from an image bank. Therefore, the variability due to the setup, or noise in the image, is excluded.

We tested five different sequences taken from the 2D-DIC image bank provided by the Society for Experimental Dynamics [21]. These sequences, which were synthetically generated, come with a full description of the movement and subpixel displacement of the texture, and are often used as testing benches for tracking algorithms. The sequences from 2D-DIC are images of random dots whose contrast, noise and shift differ in the distinct sequences. Figure 1 depicts the selected samples and their properties. Sequences contain horizontally and vertically shifted versions of the first image according to the specified steps. The number of frames in each sequence is the amount needed to accomplish a one-pixel accumulated displacement.

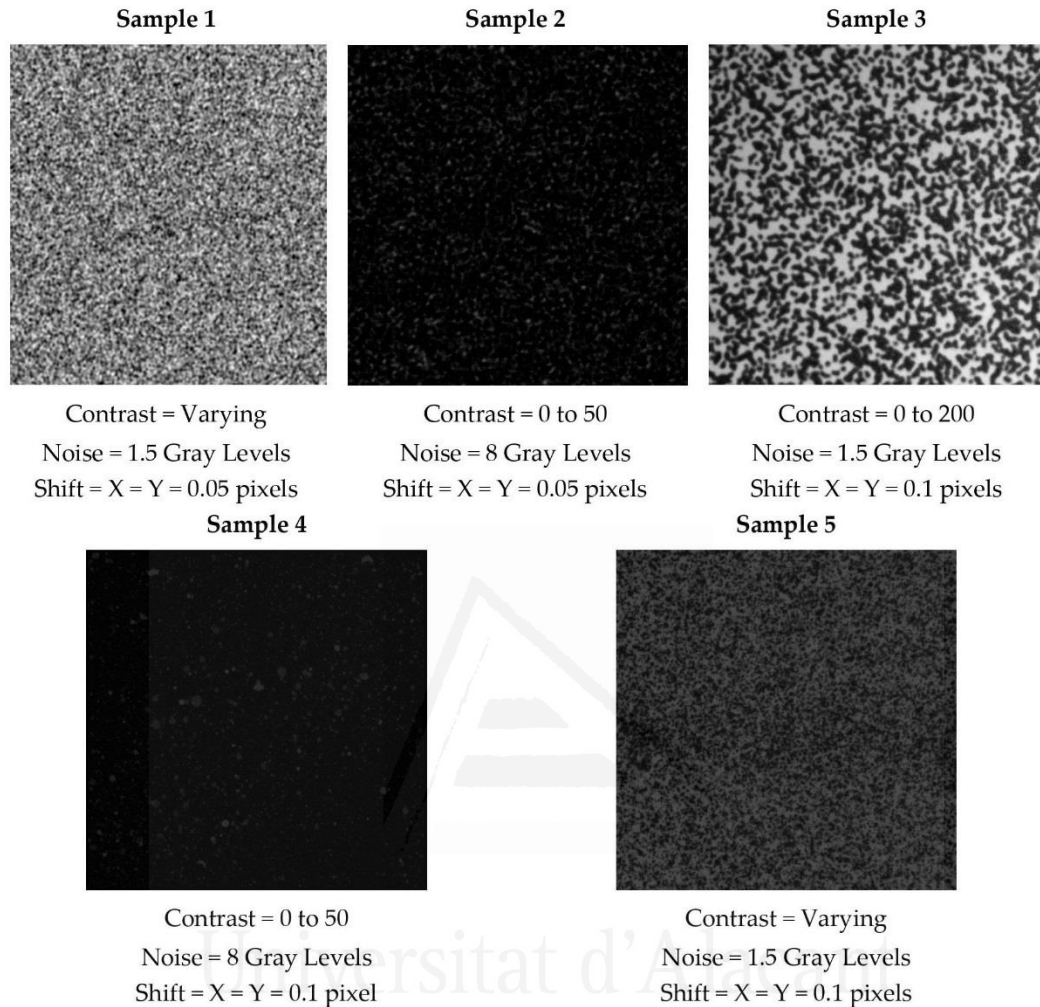


Figure 1. Images of each sample selected from 2-D DIC from the Society for Experimental Mechanics (SEM) with random dots and their properties [22].

From each sequence, the provided reference was selected as the template and its position was tracked throughout the sequence by using the cross-correlation operation which is implemented here through the normalized cross-correlation algorithm, *normxcorr2*, in Matlab [23]:

$$\gamma(u, v) = \frac{\sum_{x,y} [f(x, y) - \bar{f}_{u,v}] [t(x - u, y - v) - \bar{t}]}{\sqrt{\sum_{x,y} [f(x, y) - \bar{f}_{u,v}]^2 \sum_{x,y} [t(x - u, y - v) - \bar{t}]^2}}, \quad (1)$$

where f is the image taken as a reference and t the template, \bar{t} is the mean value of the template and \bar{f} is the mean value of $f(x, y)$ in the region under the template.

The test was carried out at full field, i.e., taking all the image as the template. Nevertheless, a frame of 8 pixels on all sides was imposed on the template in order to prevent the shifted image moving outside the boundaries of the reference image [24].

After obtaining the correlation function, a small region around the peak is selected and fitted to a soft function. These operations eventually relocate the peak inside a pixel region so that its maximum can be recalculated with incremented accuracy. As stated in the Introduction, we use a

Gaussian function, a second-order polynomial and cubic spline [19,25] as fitting functions on different neighborhood sizes around the correlation peak. Cubic splines have been implemented through the thin-plate spline algorithm provided by Matlab, which provides a smoother fitting function [26,27].

The fitting was calculated on different neighborhood areas around the maximum of the correlation peak (*Nbd*). This size was taken from 3×3 to 11×11 pixels, with the peak centered in the region, so only odd sizes were considered. Average speckle size in the samples used was estimated through autocorrelation, showing an average radius larger than 5 pixels. Therefore, areas larger than 11×11 would include unmatched results, which may distort the error estimation. We also tested the influence of defocusing on the accuracy of the tracking results. In mathematical terms, blurring can be described by a convolution (3) of the image with a Gaussian function (2):

$$G_{r_b}(x, y) = \exp\left(-\frac{x^2 + y^2}{2 \cdot r_b^2}\right), \quad (2)$$

$$f_{r_b}(x, y) = f \odot G_{r_b}(x, y) = \sum_{u,v} f(u, v) G_{r_b}(x - u, y - v), \quad (3)$$

where r_b is the blur radius.

In principle, if the camera is defocused throughout the capturing process, both the reference image and template will be blurred, so the cross-correlation between the blurred image and template can be written as (4):

$$C_{r_b} = f_{r_b} \otimes t_{r_b} = \sum_{u,v} f(u, v) G_{r_b}(x + u, y + v), \quad (4)$$

where, for simplicity's sake, we used the general definition of correlation instead of the normalized one. In any case, generalization is straightforward.

According to the basic properties of both correlation and convolution, we can rewrite Equations (4) as (5):

$$C_{r_b} = f_{r_b} \otimes t_{r_b} = f_{r_b} \otimes (t \odot G_{r_b}) = (f_{r_b} \otimes t) \odot G_{r_b} = [(f \odot G_{r_b}) \otimes t] \odot G_{r_b}, \quad (5)$$

As the Gaussian function is symmetric, the above-written expression can be finally expressed as (6):

$$C_{r_b} = f_{r_b} \otimes t_{r_b} = (f_{r_b} \otimes t) \odot G_{r_b} = [(f \otimes t) \odot G_{r_b}] \odot G_{r_b}, \quad (6)$$

Thus, we can see that the effect of the blurred reference and template is a double blurring of the correlation peak. As blurring was symmetrical, the main effect was to soften the peak to, thus, make a more adequate profile for accurate fitting. Unfortunately, a double defocus can introduce excess blurring, and can also degrade the function and mask the peak, which would cancel out the obtained advantages. Therefore, it is worth analyzing the amount of blurring that provides the best possible results. Accordingly, an analysis to compare a sharp reference with a blurred template was done and is presented in the Discussion in order to check whether the double-blurring filter was redundant or not.

The calculation process started by introducing a Gaussian filter with radius r_b to both the image and template before calculating the normalized cross-correlation. The value of the radius was varied from $r_b = 0$ (delta function, no blur) to $r_b = 5$. Figure 2 shows a flow chart with the algorithm implemented in Matlab. The background of the program is depicted in gray, whereas the specific parts of the fitting algorithms are represented in green (Gaussian fit), yellow (thin-plate splines) and pink (second-order polynomial fit) as they are shown in the Results. The depicted sequence was repeated for each sequence with a different r_b .

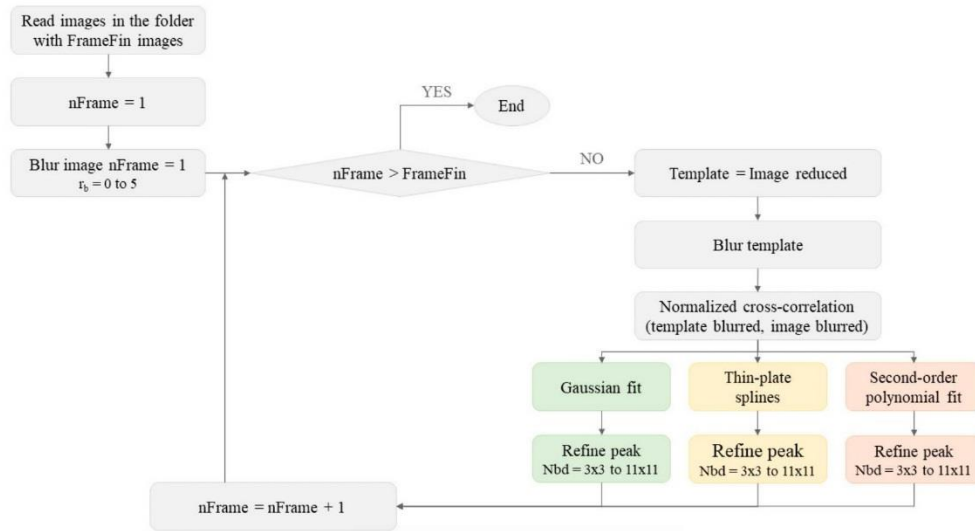


Figure 2. Flow chart with the algorithm implemented in Matlab to reduce the peak-locking error.

Briefly, each frame was compared with the first one in the sequence and the correlation function was obtained. Then a region around the correlation peak was fitted to three different functions over distinct neighborhood areas around the peak. The new peak position of the fitted function was then obtained. The new maximum was obtained through a minimum search of the inverse of the fitted functions, following the algorithm developed in [28]. Although it can be analytically calculated for the Gaussian and the quadratic case, we preferred to use the same method for all the functions in order to avoid distortions introduced by the calculation algorithms.

As the real movement was provided by the 2D-DIC image bank, our results were compared to the theoretical displacement to evaluate the errors and to obtain the best combination to measure target movement. The error was calculated through the mean error (μ) with its standard deviation (STD) and the maximum error (MaxErr) of each sequence:

$$\mu = \frac{\sum_{i=1}^N (x_{i_{calc}} - x_{i_{teor}})}{N}, \quad (7)$$

$$STD = \sqrt{\frac{\sum_{i=1}^N ((x_{i_{calc}} - x_{i_{teor}}) - \mu)^2}{N}}, \quad (8)$$

$$MaxErr = \max(|x_{i_{calc}} - x_{i_{teor}}|). \quad (9)$$

where $x_{i_{teor}}$ is the reference value provided by the DIC Challenge site, $x_{i_{calc}}$ is the value calculated through the different methods and N is the number of samples. In our case, as the measurement extended throughout the sequence, N refers to the number of frames in the evaluated sequence.

The main programs, subroutines for calculation of the maxima and obtained results can be downloaded from [29] as Supplementary Materials.

3. Results

We obtained the parameters in Equations (7)–(9) for all five samples with the three peak interpolation methods applied to neighborhoods of all odd sizes ranging from 3×3 to 11×11 , and with six different Gaussian filters with radii ranging from 0 (no blur) to 5, which totaled 90 tests per sample. The obtained displacement values were compared to the data provided by the DIC database and the error was evaluated. As presenting all the results would be extensive, we selected the results

according to the maximum error values by simply selecting the best and worst cases for each sample and interpolation method, which corresponded to the minimum and maximum MaxError, respectively. The results are summarized in Table 1. Note that, for Table 1, for remaining calculations and graphs presented in the manuscript, errors only refer to vertical shifts. The errors obtained from the horizontal displacements were similar, but their analysis was omitted to avoid a redundant analysis.

Table 1. Minimum (best adjustment) and maximum (worst adjustment) values in pixels of μ , STD and maximum error (MaxErr) in each adjustment type with the values of Nbd and r_b that provided the results. Below the sample number, the shift between successive frames is specified. The best and worst adjustments in the table are highlighted in red, and both cases were produced with the Gaussian function, with the best in Sample 3 and the worst in Sample 2.

| | | Gaussian Fit | | Thin-Plate Splines | | 2 nd -Order Polynomial Fit | |
|----------|------------------|--|---------------|---|-------------------|--|-------------------|
| | | Best | Worst | Best | Worst | Best | Worst |
| Sample 1 | $\mu \pm \sigma$ | $4 \times 10^{-4} \pm 6 \times 10^{-4}$ | 0.7 ± 0.7 | $-0.001 \pm 7 \times 10^{-4}$ | 0.004 ± 0.09 | $-9 \times 10^{-4} \pm 9 \times 10^{-4}$ | 0.009 ± 0.1 |
| | MaxErr | 0.0012 | 2.3436 | 0.0021 | 0.1105 | 0.0033 | 0.1638 |
| | Nbd | 11×11 | 3×3 | 9×9 | 3×3 | 3×3 | 11×11 |
| | r_b | 0 | 3 | 5 | 0 | 5 | 0 |
| Sample 2 | $\mu \pm \sigma$ | $9 \times 10^{-4} \pm 0.005$ | 1 ± 0.7 | 0.001 ± 0.004 | 0.006 ± 0.09 | $8 \times 10^{-4} \pm 0.005$ | 0.006 ± 0.07 |
| | MaxErr | 0.0122 | 1.9149 | 0.0096 | 0.1287 | 0.0111 | 0.1152 |
| | Nbd | 11×11 | 3×3 | 9×9 | 3×3 | 3×3 | 11×11 |
| | r_b | 3 | 4 | 5 | 0 | 5 | 0 |
| Sample 3 | $\mu \pm \sigma$ | $-3 \times 10^{-4} \pm 3 \times 10^{-4}$ | 0.8 ± 0.7 | $3 \times 10^{-4} \pm 4 \times 10^{-4}$ | -0.008 ± 0.09 | $2 \times 10^{-4} \pm 7 \times 10^{-4}$ | -0.006 ± 0.04 |
| | MaxErr | 7.67×10^{-4} | 2.3761 | 0.0010 | 0.1155 | 0.0012 | 0.0596 |
| | Nbd | 7×7 | 3×3 | 9×9 | 3×3 | 3×3 | 11×11 |
| | r_b | 4 | 2 | 5 | 0 | 5 | 0 |
| Sample 4 | $\mu \pm \sigma$ | $4 \times 10^{-4} \pm 0.003$ | 0.7 ± 0.6 | -0.001 ± 0.006 | -0.006 ± 0.1 | -0.003 ± 0.004 | -0.02 ± 0.09 |
| | MaxErr | 0.0063 | 2.0000 | 0.0125 | 0.1537 | 0.0094 | 0.1436 |
| | Nbd | 7×7 | 3×3 | 9×9 | 3×3 | 3×3 | 11×11 |
| | r_b | 0 | 5 | 2 | 0 | 2 | 0 |
| Sample 4 | $\mu \pm \sigma$ | $9 \times 10^{-4} \pm 0.001$ | 0.7 ± 0.7 | 0.002 ± 0.001 | 0.008 ± 0.1 | 0.002 ± 0.001 | 0.02 ± 0.1 |
| | MaxErr | 0.0026 | 1.9000 | 0.0033 | 0.1273 | 0.0034 | 0.1651 |
| | Nbd | 11×11 | 3×3 | 9×9 | 3×3 | 3×3 | 11×11 |
| | r_b | 0 | 4 | 4 | 0 | 5 | 0 |

The results showed that the best results for all the methods had mean errors below 0.005 px and standard deviations below 0.006 px. Additionally, note that the maximum error (MaxErr) for the best result in each fitting function, which can be taken as a measure of the peak-locking effect, was at least one order of magnitude smaller than the sample shift. This means that the tracking of samples was very good, provided that the parameters were well selected.

The results presented in Table 1 can serve to set the extreme results that were obtained through the different methods herein presented but did not indicate the influence of the different parameters. In order to better understand the influence of the different parameters, we depicted the error for the three fitting functions for a fixed defocus parameter $r_b = 2$ with all the possible neighborhood sizes in Figure 3. As the 3×3 neighborhood provided such bad results by the Gaussian and spline-fitting methods (see Table 1), the corresponding results were deleted from the graph. The lines for samples 3 to 5 corresponding to the 5×5 area of the Gaussian fit were also deleted for the same reason. Note that a complete figure was added in the Matlab format as a Supplemental File.

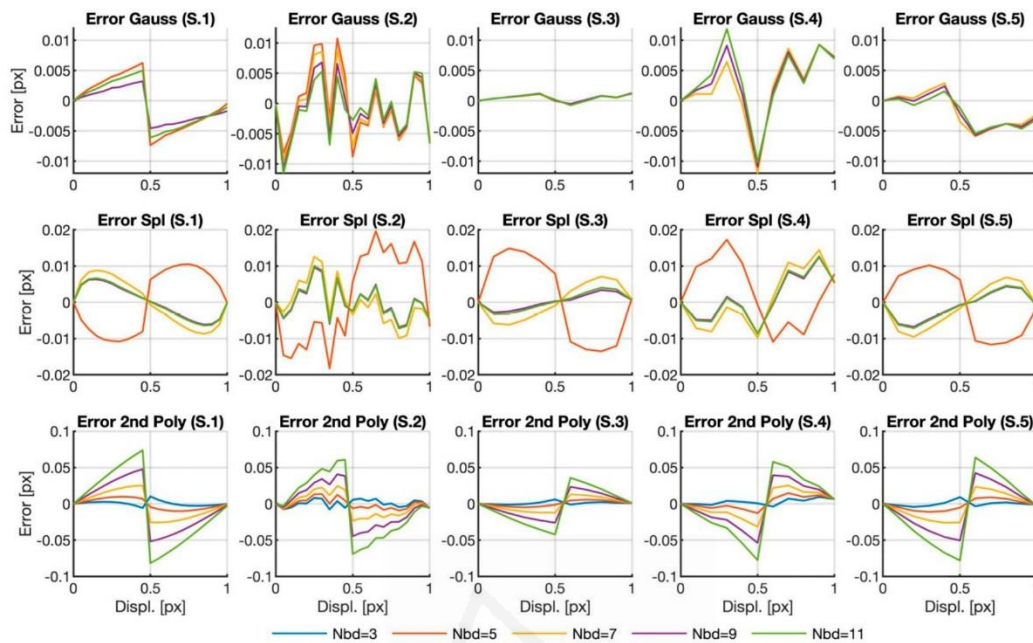


Figure 3. Location error curves obtained for $r^b = 2$. Curves for the parameter $Nbd = 3 \times 3$ in Gaussian and spline cases have not been represented for better visualization purposes. Curves for the parameter $Nbd = 5 \times 5$ have also not been represented for Samples 3 to 5 in the Gaussian case. The complete graph in Matlab format can be downloaded as a Supplemental Materials from [29].

The first noticeable fact in the graphs is that not all the error curves presented the typical sigmoid shape with symmetry around the 0.5 pixel shift value due to the pixel-locking effect (see the errors for Sample 1 in Figure 3). This is especially noticeable in the Gaussian case, where three of the five samples do not even show any clear trend. This unusual behavior does not imply large errors since the graphs depicted for the Gaussian case are of the same order or lower than the error obtained by the other fitting functions. Moreover, for this fitting function, and except for the non-depicted cases, the error does not strongly depend on neighborhood size, provided that it is large enough.

In the splines case, the typical sigmoid shape only appears in Samples 1, 3 and 5. The curve for $Nbd = 5 \times 5$ presents a very different behavior to that of the other curves. Once again, if we do not consider the anomalous cases, i.e., $Nbd = 3 \times 3$ and $Nbd = 5 \times 5$, there seems no systematic error dependence on neighborhood size: although we can see some differences in each individual sample, it is relatively small and dependence on size is not the same in all cases.

Finally, the curves representing the errors calculated with the second-order polynomials present the typical shape due to the peak-locking effect. In this case, the error is bigger than that obtained with the other two functions and is of the same order in all the samples. Unlike what happened with the other two fitting functions, here, we notice a marked dependence of neighborhood size, where the bigger the error, the larger the interpolation area.

Figure 4 depicts our analysis of the influence of defocusing on the error, along with the error for fixed neighborhood $Nbd = 7 \times 7$ and changing defocus parameter r_b . At first glance, the curves share some similarities with the curves in Figure 3, i.e., lack of the typical sigmoid shape in the same cases as before. It is also noticeable that, in all those cases, there was no significant dependence on the defocusing parameter. Notwithstanding, and as before, the error obtained by fitting the peak with a Gaussian function is equal to or lower than the error obtained by other fitting functions.

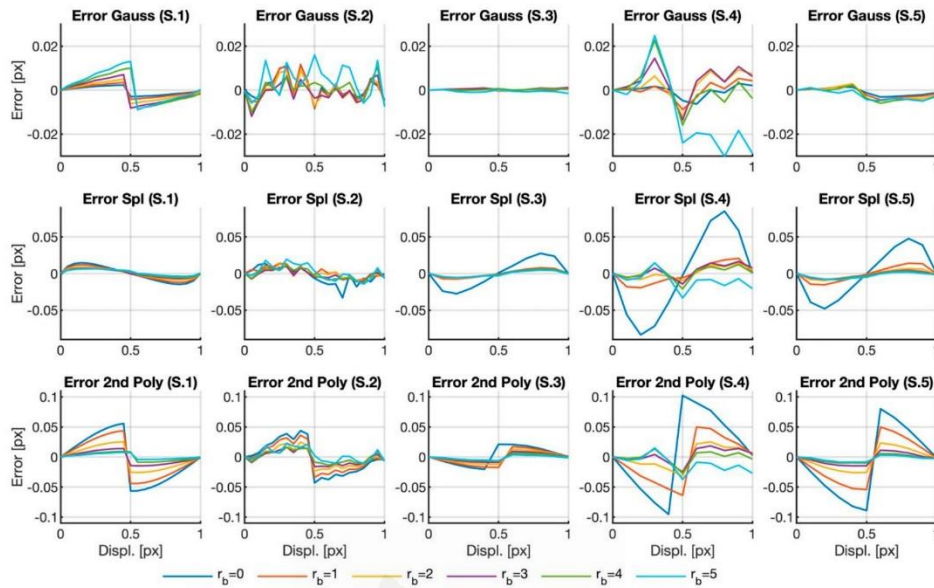


Figure 4. Location error curves obtained for $Nbd = 7 \times 7$. The complete graph in Matlab format can be downloaded as a Supplemental Materials from [29].

Regarding the influence of the defocus, note that the more marked the defocus, the lower the error for all the cases calculated by the quadratic polynomial. This can also be stated for the spline method, although in this case, the larger difference lies between $r_b = 0$ and all the other cases.

According to the depicted figures, it would seem that the results obtained using the Gaussian function for fitting the peak were independent of neighborhood size and image blurring, provided that the calculation area was large enough. Moreover, the results obtained by this method did not present the typical peak-locking shape, and errors were similar to or lower than in the other methods.

Despite the obtained results, we have only analyzed the error due to the calculation area for one fixed defocusing filter and the effect of the defocus for a fixed neighborhood, respectively. Therefore, in order to gain more insight into the error dependence on the analyzed parameters, Figures 5–7 reveal the plots of the variation of the three error parameters for the row shifting expressed in Equations (7)–(9) (mean error, standard deviation and maximum error, respectively) with the defocus parameter for all the neighborhood sizes and for the three fitting functions herein analyzed.

Figure 5 illustrates the mean value of the error obtained for each case according to the radius of the blurring Gaussian filter. As we were interested only in the error magnitude, the absolute value of the error is represented. A different line is depicted for each neighborhood size. As we can see in Table 1, the Gaussian function does not provide good results with small interpolation areas. Therefore, the 3×3 neighborhoods graphs were deleted to facilitate the visualization of the other curves because of their large errors, with values higher than 0.5 px (see Table 1). For the same reason, the 5×5 neighborhood error curve obtained for the Gaussian function was also deleted for Sample 5, with a peak close to 0.1 px. Apart from the reported cases, the mean value of the error was below 0.05 px in most cases, which went below the imposed shift between frames. The complete figure was added as Supplemental Materials to better allow the interpretation of the results.

At first glance, it would seem that the defocus increased the error when Gaussian or spline functions were used. In both cases, a defocus with radius $r_b = 4$ seemed to produce an error reduction in some samples, but we hypothesize that this happened because the particular texture of this sequence and is not a general rule. Nevertheless, we can see that the error change due to the blurring filter was less than 1% of the shift (0.05 px in the first two samples and 0.1 in the other three). So the effect on the

mean value could not be considered very strong, but could be important when using the quadratic fitting function with large interpolation areas.

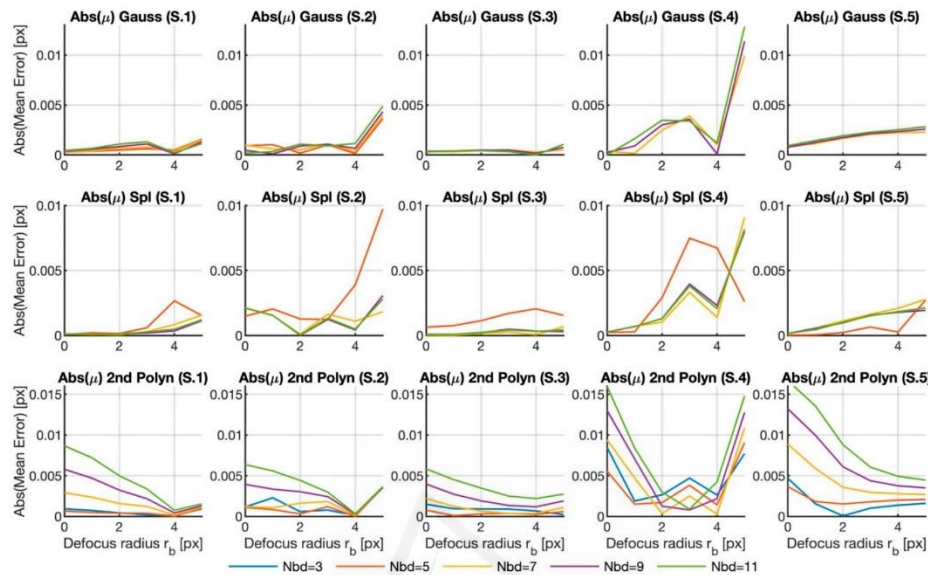


Figure 5. Mean errors calculated for all the samples, fitting functions and different neighborhood sizes versus the Gaussian filter radius. The graphs for the 3×3 region in the Gaussian and spline fitting surfaces were deleted for better visualization purposes. The graphs for the 5×5 region were also deleted for Samples 3 to 5 in the Gaussian case. The complete graph in Matlab format can be downloaded as a Supplemental Materials from [29].

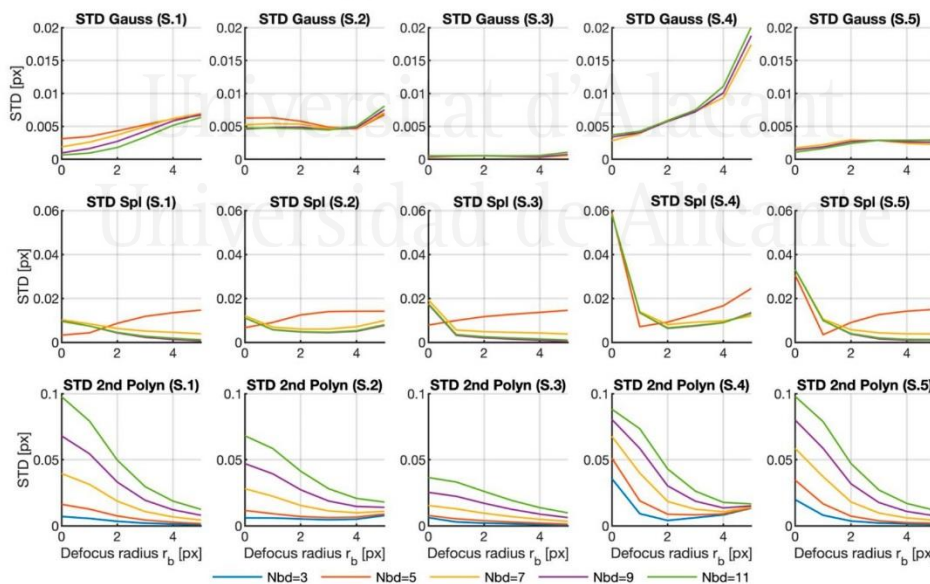


Figure 6. Standard deviations calculated for all the samples, fitting functions and different neighborhood sizes versus the Gaussian filter radius. The graphs for the 3×3 region in the Gaussian and spline fitting surfaces were deleted for better visualization purposes. The graphs for the 5×5 region were also deleted for Samples 3 to 5 in the Gaussian case. The complete graph in Matlab format can be downloaded as a Supplemental Materials from [29].

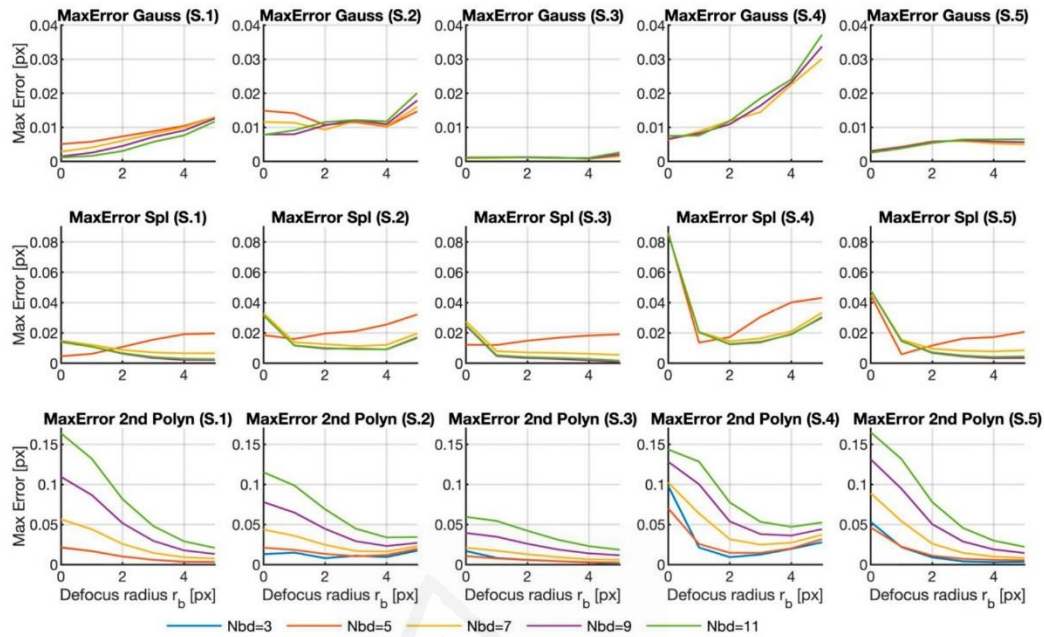


Figure 7. Maximum error calculated for all the samples, fitting functions and different neighborhood sizes versus the Gaussian filter radius. The graphs for the 3×3 region in the Gaussian and spline fitting surfaces were deleted for better visualization purposes. The graphs for the 5×5 region were also deleted for Samples 3 to 5 in the Gaussian case. The complete graph in Matlab format can be downloaded as a Supplemental Materials from [29].

Regarding the Nbd parameter, we can see that, for the depicted Gaussian and polynomial cases, and for $Nbd = 7 \times 7, 9 \times 9$ and 11×11 , the larger the interpolation area, the bigger the error, but not in all cases. Note also that the $Nbd = 3 \times 3$ case was somewhat anomalous because in the Gaussian and spline functions, the errors obtained for the smaller case were huge in all the samples. The 5×5 domain also produced large errors in Samples 3 to 5 with the Gaussian function and erratic behavior with splines. Finally, when quadratic functions are used, these two particular domains seemed to provide opposite results according to the general trend described for this case.

Despite this analysis, the mean value was a poor parameter for measuring the error. It described trends in the result but, as these sequences were artificially generated, no a strong bias was expected here, as previously seen in Figures 3 and 4. In any case, we discovered that using very small interpolation areas may be problematic in the majority of cases.

Figure 6 offers the graph for the standard deviation (STD) for all the discussed cases. This parameter indicated the variability of the results. As in the previous case, the graphs with the highest values were deleted. This happened for all the curves corresponding to $Nbd = 3 \times 3$ in the Gaussian and spline cases, and to the 5×5 curves for the Gaussian case and Samples 3 to 5. Once again, the complete graph is included as Supplemental Materials.

In the graphs, we can see that blurring may help to narrow the variability in the results. The strength of this effect depends very much on the fitting function. In the Gaussian case, the effect strongly depends on the sample as the benefit is observable only in Sample 1, while the error increases with blurring for the other samples. In the polynomial case, the effect is general, with better results for large r_b . For the spline method, the improvement is also noticeable, albeit very weak. Note that the effect on some samples is nonexistent, or even negative.

Regarding the influence of the neighborhood size, once again, $Nbd = 3 \times 3$ combined with Gaussian or spline functions resulted in wide variability and a large error, as mentioned above. For the

Gaussian case, this dependence strongly depended on the sample as we observed the opposite behavior in different samples. With the polynomial function, the results were less dispersed the smaller the area was, which agrees with what is deduced from the table, but is the opposite to what happened with the mean error. However, the relation was very weak in that case. When spline fitting was applied, neighborhood size displays no clear dependence, except for the 3×3 case.

When considering the absolute value of the standard deviation, we find that, accordingly with the previous results, the Gaussian function generally gives lower values than the other two methods.

Another useful parameter for determining the performance of each fitting function is the maximum error, which may correspond to the peak of the peak-locking error. Figure 7 displays the graphs with the maximum error for each sequence, fitting method and neighborhood in front of the blurring radius. As in previous cases, the curves corresponding to the Gaussian and spline fitting methods in a 3×3 region were deleted from all the samples. The 5×5 regions in Samples 3 to 5 was also deleted for the Gaussian method as they gave values around 1 px which would not, thus, allow the other cases to be visualized.

We can see that the curves depicted for the maximum error are similar to those with the standard deviation, which implies that the maximum value of the peak-locking error is probably the main source of the errors in the calculation. Hence, the depicted results confirmed the conclusions drawn from the other graphs and no further comments will be added.

4. Discussion

The results depicted above confirm the hypotheses posed in the Introduction: quadratic functions provide good results for peak fitting, but are prone to peak-locking error, while Gaussian functions give more robust results, provided that the fitting area is large enough (see Figures 4 and 7). As we said, narrow peaks are supported by a few samples, so the maximum weight determines the fitting result, pulling the recalculated maximum location towards the location of the sample where the original maximum is located. Including more samples would compensate the result but would also include areas outside the peak. Eventually, the fitting neighborhood would include the peak skirt and curvature changes which would no longer fit to a paraboloidal surface, and thus the error will increase.

Gaussian functions are capable of reproducing both the peak and the planar surrounding area, although more samples are required. Because of this, the weight of the maximum is compensated, and the location error is less prone to being affected by peak-locking error. Among both situations, spline functions provide reasonable results in all neighborhood sizes, but because of their adaptability, the fitting is also biased towards the peak maximum and, thus, it is also affected by the peak-locking effect.

In order to illustrate the adaptability of each fitting function to the peak surface, we have represented in Figure 8 the correlation peak corresponding to a shift of 0.4 px in the two extreme cases calculated for Sample 1; i.e., with 3×3 and 11×11 neighborhoods and no blur. Although the curves may have wide variability from frame to frame and for different samples, they serve to illustrate the point herein explained. We notice there that the Gaussian function cannot adapt to the skewed shape of the local area around the maximum, while the polynomial function correctly fits to it. In the case of the larger neighborhood, the Gaussian function can reproduce the curvature change while the quadratic function just reproduces a paraboloidal dome. As we said, the spline function can adapt to both situations.

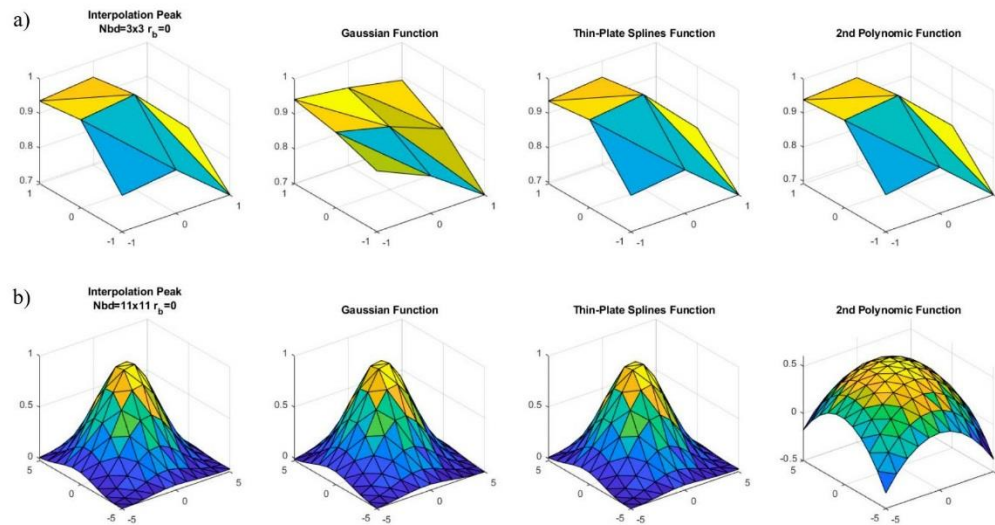


Figure 8. Peak adjustment for Sample 1 with a 0.4 shift using no blur and (a) 3×3 and (b) 11×11 neighborhoods. For each neighborhood, we show the correlation peak and the reconstructed surface by employing the fitting functions specified above each plot.

The results here shown also explain the reason why the Gaussian function seems to be more insensitive to the neighborhood size once it is large enough (see Figure 3). Because of its particular shape, the weight of the samples farthest from the center is very low and therefore will not affect the location results. In the case of the spline function, there is not such distance compensation and the result may be affected. Finally, in the quadratic case, it is clear that larger neighborhoods produce larger errors.

Regarding the blurring filter, it has a double effect in the correlation surface. On the one hand, the peak gets softer, so the fitting by smooth functions is more accurate. This affects more noticeably the results obtained with the quadratic function which are very dependent on the central samples (see Figures 4 and 7). On the other hand, blurring also decreases the weight of the eccentric samples, increasing the relative weight of the central sample and thus may slightly increment the location error obtained through Gaussian fitting. In any case, the effect introduced by the Gaussian filter is similar to that produced by the Gaussian fitting function, hence explaining why the error obtained through this fitting method is not very much affected by blurring or even increases in some cases. At this point, we wish to recall Equation (6) where the mathematical formulation of the blurring filter is explained. By comparing a blurred reference to a blurred template, we obtained a double-blurring of the correlation function. As this double application of the blurring filter is redundant, we hypothesized that it is possible to compare a sharp reference to a blurred template (or vice versa) without increasing errors. Thus, we tested the results when a sharp reference was compared to a blurred template. In Figure 9, we plotted the curves with the maximum errors.

As we can see, the results are similar to those obtained using a double defocus. This means that blurring is not a decisive parameter in the peak-locking error. Although it can help to improve the results, adding a Gaussian filter (or an experimental defocus) in both the images used for the normalized cross-correlation was redundant.

This “invariability” in the template defocus proved most convenient for some experimental implementations: long-time experiments may cause small mechanical drifts in the camera or the sample and, therefore, some frames may appear slightly blurred. A defocus may take place during experiments using short depth-of-field lenses and samples whose size may change due to heat dilation, tension or swelling [30]. The herein shown results demonstrate that these changes applied to one of

the images being compared had no marked effect on the final result, which remained valid and, thus, indicated that the subpixel tracking through local interpolation was a robust method.

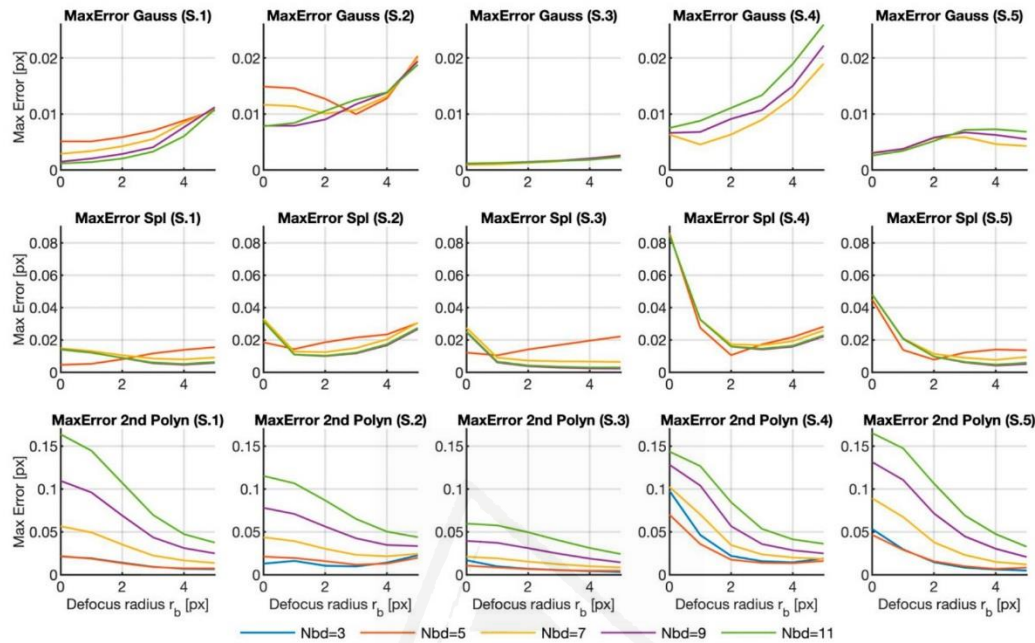


Figure 9. Maximum error calculated when comparing sharp references with blurred templates for all the samples, fitting methods and different neighborhood sizes versus the Gaussian filter radius. The graphs for the 3×3 region in the Gaussian and spline fitting surfaces were deleted for better visualization purposes. The graphs for the 5×5 region were also deleted for Samples 3 to 5 in the Gaussian case. The complete graph in Matlab format can be downloaded as a Supplemental Materials from [29].

According to the results herein presented, we can reach several conclusions. Regarding the fitting function, we found that, except for the smallest fitting region, the Gaussian fit gave the smallest errors that are, in some cases, almost one order of magnitude smaller than the errors introduced by the other methods. However, the best results in all the functions were similar, which means that all the fitting functions would display similar performance under optimal conditions.

When Gaussian or spline functions are used, small areas around the peak should be avoided. With these two functions, and except for areas 3×3 and 5×5 , the errors did not seem to depend on either neighborhood size or the blurring filter radius. This result agrees with the partial results represented in Figures 3 and 4. On the contrary, when employing quadratic polynomials as fitting functions, both the neighborhood size and blurring strongly impacted the results, and the smaller the errors, the smaller the fitting area and the larger the defocus parameter.

The effect of blurring to improve the results was noticeable in many cases, but there are many exceptions. Hence, we cannot state that the use of blurring to diminish the error is a general benefit because it depends on both the fitting function and the sample. In any case, except for the Gaussian fitting functions, a minor defocus could help to narrow the variability of the results (taken as the standard error) and to slightly reduce the peak error. So, introducing it could be advisable.

Thus, in summary, we found the smallest error with a Gaussian fitting applied in a large neighborhood around the peak and with no blurring.

The results here obtained agree with those that appear in the literature [16], where the authors obtained better results for the Gaussian fitting than for the other methods by using an 8×8 neighborhood. In [17,19], the authors reported a minor improvement when using a slight defocus and a Gaussian fitting.

5. Conclusions

In this manuscript, we tested the accuracy of the commonest subpixel tracking methods based on cross-correlation. We focused on the methods that employ local interpolation in a small area around the correlation peak to refine the maximum location. To this end, we tested the influence of neighborhood size around the peak with three different fitting functions: Gaussian, thin-plate splines and second-order polynomials. We also checked the use of defocus as a strategy for diminishing the peak-locking error and how it was affected when that defocus changed along the image sequence. All the tests were carried out in five sequences taken from the DIC Challenge site [21].

We generally noticed that the three functions provided good accuracy, and slight blurring helped to increase accuracy, despite us finding slight variation among samples. The fitting functions provided different results depending on neighborhood size. Therefore, the Gaussian function provided the best results with large neighborhoods (11×11), while the second-order polynomial seemed to work better with small areas (3×3). The thin-plate function apparently worked correctly with any neighborhood size. Our tests reveal that the best result was obtained for the Gaussian function with a neighborhood of 11×11 and no defocus. However, as the worst adjustment was also achieved with the Gaussian function, it is important to correctly select the values of both the focus and the neighborhood depending on the fitting function. Blurring significantly improved the error for a second-order polynomial fitting, but had no clear trend for the other two studied functions.

Additionally, in the Methods section, we show that comparing two blurred images by a cross-correlation operation is the equivalent to comparing two sharp images and then introducing double deblurring into the correlation function. Accordingly, we recalculated the errors by comparing a sharp reference with a blurred template. The results showed that the performance of the three methods with all the different parameters were similar to the double deblurring case. So we can conclude that the methods presented herein are blur invariant, which means that the subpixel technique based on the interpolation of the correlation peak is robust for experimental implementations in which the focus may change due to drifts in the optical system or to the sample's position change.

As a consequence of this, the rule of thumb that can be derived from our tests is that Gaussian fitting applied on large neighborhoods around the maximum of the correlation function may provide the most accurate results in pseudospeckle images without the need for blurring filters.

Supplementary Materials: The following are available online at <http://rua.ua.es/dspace/handle/10045/110141>, Matlab code: Calculate_subpixel_errors.m; Results: all numerical results in Matlab binary format.

Author Contributions: D.M. and B.F. conceived the presented idea and developed the theory. M.-B.T. implemented the programs, ran the tests and proposed some interesting conclusions. D.M. and M.-B.T. presented the final data. All authors have supervised the findings of this work. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been supported by the Generalitat Valenciana and the European Social Fund (FSE) through the Recruitment of Predoctoral Research Staff ACIF/2018/211 included in the FSE Operational Program 2014–2020 of the Valencian Community. Belén Ferrer and María-Baralida Tomás acknowledge the support of the Generalitat Valenciana through Project GV/2020/077.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Maragos, P. Morphological correlation and mean absolute error criteria. In Proceedings of the International Conference on Acoustics Speech, and Signal Processing, Glasgow, UK, 23–26 May 1989; pp. 1568–1571.
2. Stanier, S.A.; Blaber, J.; Take, W.A.; White, D.J. Improved image-based deformation measurement for geotechnical applications. *Can. Geotech. J.* **2016**, *53*, 727–739. [CrossRef]

3. Dias-da-Costa, D.; Valença, J.; Júlio, E.; Araújo, H. Crack propagation monitoring using an image deformation approach. *Struct. Control Health Monit.* **2017**, *24*, e1973. [CrossRef]
4. Vora, S.R.; Bogner, B.; Patanwala, H.S.; Young, C.D.; Chang, S.-Y.; Daux, V.; Ma, A.W.K. Global strain field mapping of a particle-laden interface using digital image correlation. *J. Colloid Interface Sci.* **2018**, *509*, 94–101. [CrossRef]
5. Bai, R.; Wei, Y.; Lei, Z.; Jiang, H.; Tao, W.; Yan, C.; Li, X. Local zone-wise elastic-plastic constitutive parameters of Laser-welded aluminium alloy 6061 using digital image correlation. *Opt. Lasers Eng.* **2018**, *101*, 28–34. [CrossRef]
6. Peña, J.A.; Corral, V.; Martínez, M.A.; Peña, E. Over length quantification of the multiaxial mechanical properties of the ascending, descending and abdominal aorta using Digital Image Correlation. *J. Mech. Behav. Biomed. Mater.* **2018**, *77*, 434–445. [CrossRef]
7. Lewis, J.P. Fast Normalized Cross-Correlation. *Industrial Light & Magic* **1995**, *10*, 7.
8. Schreier, H.; Orteu, J.-J.; Sutton, M.A. *Image Correlation for Shape, Motion and Deformation Measurements*; Springer: Boston, MA, USA, 2009; ISBN 978-0-387-78746-6.
9. Lei, X.; Jin, Y.; Guo, J.; Zhu, C. Vibration extraction based on fast NCC algorithm and high-speed camera. *Appl. Opt.* **2015**, *54*, 8198. [CrossRef]
10. Ferrer, B.; Mas, D. Parametric evaluation of errors using isolated dots for movement measurement by image cross-correlation. *Sensors* **2018**, *18*, 525. [CrossRef]
11. Ferrer, B.; Espinosa, J.; Mas, D. A method to measure small local strains in concrete surfaces using its natural texture and image cross-correlation. *Struct. Control Health Monit.* **2019**, *26*. [CrossRef]
12. Nogueira, J.; Lecuona, A.; Nauri, S.; Legrand, M.; Rodríguez, P.A. Quantitative evaluation of PIV peak locking through a multiple Δt strategy: Relevance of the rms component. *Exp. Fluids* **2011**, *51*, 785–793. [CrossRef]
13. Sjö Dahl, M.; Benckert, L.R. Systematic and random errors in electronic speckle photography. *Appl. Opt.* **1994**, *33*, 7461. [CrossRef]
14. Murray, C.A.; Hoult, N.A.; Take, W.A. Dynamic measurements using digital image correlation. *Int. J. Phys. Model. Geotech.* **2017**, *17*, 41–52. [CrossRef]
15. Stanier, S.; Dijkstra, J.; Leśniewska, D.; Hambleton, J.; White, D.; Muir Wood, D. Vermiculate artefacts in image analysis of granular materials. *Comput. Geotech.* **2016**, *72*, 100–113. [CrossRef]
16. Nobach, H.; Damaschke, N.; Tropea, C. High-precision sub-pixel interpolation in particle image velocimetry image processing. *Exp. Fluids* **2005**, *39*, 299–304. [CrossRef]
17. Overmars, E.F.J.; Warncke, N.G.W.; Poelma, C.; Westerweel, J. *Bias Errors in PIV: The Pixel Locking Effect Revisited 15th Int Symp on Applications of Laser Techniques to Fluid Mechanics*; ITCES: Lisbon, Portugal, 2010; p. 10.
18. Zhou, Y.; Sun, C.; Song, Y.; Chen, J. Image pre-filtering for measurement error reduction in digital image correlation. *Opt. Lasers Eng.* **2015**, *65*, 46–56. [CrossRef]
19. Michaelis, D.; Neal, D.R.; Wieneke, B. Peak-locking reduction for particle image velocimetry. *Meas. Sci. Technol.* **2016**, *27*, 104005. [CrossRef]
20. Mas, D.; Perez, J.; Ferrer, B.; Espinosa, J. Realistic limits for subpixel movement detection. *Appl. Opt.* **2016**, *55*, 4974. [CrossRef] [PubMed]
21. Challenge Dataset 1.0:2D-DIC. Available online: <https://sem.org/2ddic> (accessed on 10 July 2020).
22. Tomás López, M.B.; Mas, D.; Ferrer, B. Peak-locking minimization by three adjustment methods. In *Proceedings of the Optics, Photonics and Digital Technologies for Imaging Applications VI*; Schelkens, P., Kozacki, T., Eds.; SPIE: Jakarta, France, 2020; p. 53.
23. *Matlab*, version R2020a; MathWorks: Natick, MA, USA, 2020.
24. Sjö Dahl, M. Accuracy in electronic speckle photography. *Appl. Opt.* **1997**, *36*, 2875. [CrossRef] [PubMed]
25. Roesgen, T. Optimal subpixel interpolation in particle image velocimetry. *Exp. Fluids* **2003**, *35*, 252–256. [CrossRef]
26. Thin Plate Spline. Wikipedia. Available online: https://en.wikipedia.org/wiki/Thin_plate_spline (accessed on 16 October 2020).
27. Thin-Plate Smoothing Spline—MATLAB Tpsps—MathWorks España. Available online: <https://es.mathworks.com/help/curvefit/tpsps.html> (accessed on 16 October 2020).
28. D'Errico, J. *fminsearchbnd*, *fminsearchcon*. Available online: <https://es.mathworks.com/matlabcentral/fileexchange/8277-fminsearchbnd-fminsearchcon> (accessed on 16 October 2020).

29. Tomás, M.B.; Ferrer, B.; Mas, D. Supplementary Materials on Influence of Neighborhood Size and Cross-Correlation Peak-Fitting Method on Location Accuracy. Available online: <http://rua.ua.es/dspace/handle/10045/110141> (accessed on 5 November 2020).
30. Yang, D.S.; Bornert, M.; Chanchole, S.; Gharbi, H.; Valli, P.; Gatzmiri, B. Dependence of elastic properties of argillaceous rocks on moisture content investigated with optical full-field strain measurement techniques. *Int. J. Rock Mech. Min. Sci.* **2012**, *53*, 45–55. [[CrossRef](#)]

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).



Universitat d'Alacant
Universidad de Alicante

PUBLICACIÓN 2: Use of image correlation to measure macroscopic strains by hygric swelling in sandstone rocks

Ferrer, B; Tomás, M.-B.; Mas, D. Use of image correlation to measure macroscopic strains by hygric swelling in sandstone rocks. *Applied Sciences* **2021**, *11*, 2495, <https://doi.org/10.3390/app11062495>

Revista *Applied Sciences* de MDPI (online)

ISSN: 2076-3417

Factor de Impacto: 2.474

JCR: Q2

Article

Use of Image Correlation to Measure Macroscopic Strains by Hygric Swelling in Sandstone Rocks

Belén Ferrer ^{*}, María-Baralida Tomás  and David Mas 

University Institute of Physics Applied to the Sciences and Technologies, University of Alicante, P.O. Box 99, 03080 Alicante, Spain; maria.barialida@ua.es (M.-B.T.); david.mas@ua.es (D.M.)

* Correspondence: belen.ferrer@ua.es; Tel.: +34-965-903-400

Abstract: Some materials undergo hygric expansion when soaked. In porous rocks, this effect is enhanced by the pore space, because it allows water to reach every part of its volume and to hydrate most swelling parts. In the vicinity, this enlargement has negative structural consequences as adjacent elements support some compressions or displacements. In this work, we propose a normalized cross-correlation between rock surface texture images to determine the hygric expansion of such materials. We used small porous sandstone samples ($11 \times 11 \times 30 \text{ mm}^3$) to measure hygric swelling. The experimental setup comprised an industrial digital camera and a telecentric objective. We took one image every 5 min for 3 h to characterize the whole swelling process. An error analysis of both the mathematical and experimental methods was performed. The results showed that the proposed methodology provided, despite some limitations, reliable hygric swelling information by a non-contact methodology with an accuracy of 1 micron and permitted the deformation in both the vertical and horizontal directions to be explored, which is an advantage over traditional linear variable displacement transformers.

Keywords: rock swelling; image processing; DIC; macroscopic strain measurement



Citation: Ferrer, B.; Tomás, M.-B.; Mas, D. Use of Image Correlation to Measure Macroscopic Strains by Hygric Swelling in Sandstone Rocks. *Appl. Sci.* **2021**, *11*, 2495. <https://doi.org/10.3390/app11062495>

Academic Editor:
Giuseppe Lacidogna

Received: 15 February 2021
Accepted: 8 March 2021
Published: 11 March 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Hygric swelling is a process that some sandstone rocks undergo when humidity rises. It implies the volume increment of some rock particles, which causes strains, stresses and, depending on the rock composition, a global volume increment in rock. This process can produce small cracks, which is quite a common phenomenon in Europe [1,2]. Consequently, strain measurement due to hygric swelling is an important parameter to assess the suitability of a particular sandstone rock in the construction and restoration of historic buildings.

The most straightforward way to measure hygric swelling is to partially or completely submerge a rock probe in water and measure the vertical displacement of the upper probe boundary with a linear variable displacement transformer (LVDT) [3]. Given the total vertical probe length, and by assuming a zero displacement of its basis, a global vertical strain can be easily calculated as the mean value for all the points on the rock. However, on the one hand, not all points undergo the same strain, and on the other hand, only vertical strains can be measured by this simple procedure. Horizontal swelling measurements by LVDT need more complex devices that are not always suitable for the rock probe under study [4].

Those inconveniences can be overcome by methods based on imaging techniques. Measuring strains by image processing with the proper setup and calculation methods allows the strain to be found for each point in the image according to time [5]. However, image procedures followed to find strains also have some drawbacks. Some are related to the calculation method for tracking a specific detail with time. To this end, the most widely used operation is normalized digital image cross-correlation (DIC), which is implemented herein by the normalized cross-correlation algorithm, *normxcorr2*, in MATLAB [6].

Although DIC has a nominal accuracy of one pixel, accuracy can increase by subpixel techniques. These techniques consist of interpolating the image or interpolation peak. In both cases, the result is biased toward the nearer integer to, thus, introduce a symmetric error with a sigmoid shape to determine the subpixel position. Some errors may also appear during image recording, such as unexpected camera movements due to mechanical drifts, overheating, changes in ambient light or image distortion due to the camera lens. Therefore, a thorough error study that considers all these factors should be conducted as part of any image procedure.

Image methods for measuring rock swelling have been implemented in very different ways in the literature: generally, probe size determines the most convenient imaging device for acquiring images, ranging from microscope to commercial cameras, and even different image setups can be used for the same probe size. In [7], scanning electron microscope images were used to compare the rock microstructure to the strains measured on a sample of centimeters in size. To measure strains, a 5 Mpx camera with a telecentric lens was used on the same probe side. The resolution for DIC calculation images was $0.5 \text{ px}/\mu\text{m}$. Those images were analyzed by a DIC method developed by [8], which considers the possible cracks that were expected in the analyzed sample. Strains were calculated by displacement derivation using the correlation results. The results showed a very heterogeneous strain distribution during desiccation, probably due to the presence of heterogeneous water and microstructural non-homogenous distribution.

Images directly taken by electron microscopy have also been used to perform DIC strain calculations [9]. In [10], a sample slice (1 mm thick and few millimeters on plane) of an argillaceous rock was recorded during swelling by electron microscopy with a 13.8 Mpx size and a resolution of $0.06 \text{ px}/\mu\text{m}$. The results showed that the macroscopic swelling strain was the combined result of the local free swelling strains and the additional mechanical strains induced by particle interactions. Argillaceous rocks are inhomogeneous materials with different hygric properties that lead to incompatibilities of free-swelling deformations for all its different particles. Additionally, the moisture gradient in the transient state of moisture transport makes some parts swell before others, which confers deformation additional incompatibility. These incompatibilities result in mechanical internal stresses that affect macroscopic local movement during swelling. In summary, previous works have demonstrated that the argillaceous swelling process is substantially affected by non-controllable factors such as particle distribution in the sample or moisture transport distribution, which mean that any swelling experiment is hardly reproducible.

From the marked uncertainty point of view that comes with measuring rock swelling, our approach involved simplifying the procedure to obtain a comparable measure of the macroscopic strain to those obtained by LVDT, but without using complex setups and with the possibility of obtaining strains in both directions.

This paper analyzes the swelling process of an argillaceous rock, with a sample of one centimeter in size and a similar image resolution to the papers herein cited, but by using a simpler setup and image processing methods. Rock expansion was analyzed by changes in the rock surface texture. The movement of texture imposed by swelling was tracked by DIC, applied to six different regions of interest (ROIs) located on the edge of the rock surface. The errors of both the experimental setup and numerical methods were carefully analyzed. Finally, it was possible to obtain the relative deformation in the vertical and horizontal directions, and to observe non-uniform stone deformation depending on the proximity of the ROIs to the wet surface.

Some preliminary results obtained with this study have been presented in [11]. Based on these results, the setup and the calculation method very much improved. Mathematical methods have been discussed in [12], where full access to the code is available.

This manuscript is structured as follows. First, we describe the experimental setup and numerical methods, including the error analysis. The results of the dry and wet experiments, and their discussion, are included in Section 3. Finally, the main results are summarized in the last section.

2. Materials and Methods

2.1. Experimental Setup

In this manuscript, we tested porous sandstone used as a construction and building material. It is composed of quartz and feldspar with a clay-rich matrix with known expansive properties. It presents a connected porosity within the 10–12% range. The rock was cut in rectangular parallelepiped samples (size of $11 \times 11 \times 30 \text{ mm}^3$). In all, four rock samples were tested. In Figure 1, we present a picture of a sample used in our experiments.



Figure 1. Picture of one rock samples. The image has been obtained through the optical setup described in the text.

Samples were placed in a small container and covered with water up to 1/3 of their height. Water moved by capillarity forces to the top of the samples as with a piston-like imbibition process. As water rose through the sample, it hydrated clay, increased its size and forced rock swelling.

Rock surface images were taken by a color camera Basler acA4600-10uc with a CMOS sensor of $6.5 \times 4.5 \text{ mm}$ with a Bayer color filter and a spatial resolution of $4608 \times 3288 \text{ px}$, being pixel size of $1.4 \times 1.4 \mu\text{m}$. Since the stone is greyish (see Figure 1), no relevant differences were found between channels, and provided that the green channel is better sampled due to the Bayer filter, only this channel was analyzed herein. However, the other channels were also tested, resulting in similar but noisier results.

A telecentric objective Myutron VTL0513, with the additional VTL05FC lens was used. This objective, together with the lens, provides a magnification of $0.5\times$. One of the main advantages of using telecentric lenses is that they have a constant non-angular field of view, so they present a significant smaller distortion and field of curvature than conventional lenses. Additionally, the depth of focus value was very low and helped to replace samples in the same exact place. We would like to underline that, since the illumination was controlled by exposition time, the diaphragm was set at its maximum aperture in order to further minimize depth of focus.

The optical system was centered in the upper part of the probe, thus avoiding the presence of the container border in the image but including both the sample's lateral borders. The employed magnification, at a working distance of 176 mm (Figure 2), provided a clear image of the rock's components (Figure 1) with a $0.36 \text{ px}/\mu\text{m}$ ratio. Both the sample and image system were placed inside a photo studio light box (Figure 3) to obtain uniform light during the whole experiment and to avoid the influence of air movement on the free water movement through the sample. As it can be seen in the picture, the whole setup was mounted on an antivibratory table. The image brightness, regulated through

the artificial light and the exposition time, was adjusted in order to have the maximum contrast, measured through the image histogram.

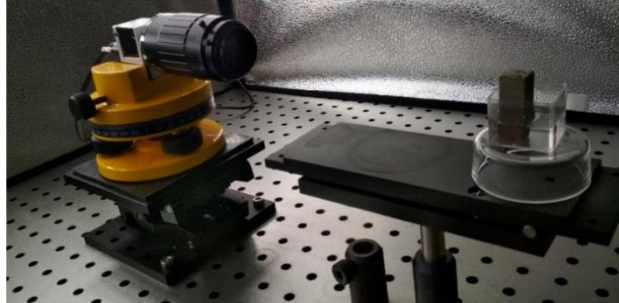


Figure 2. Setup for the imaging system and sample.

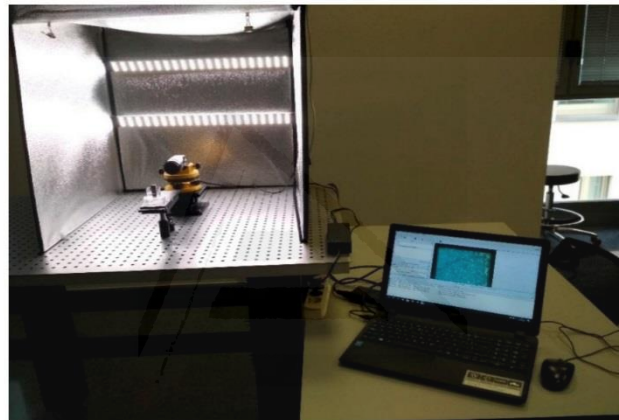


Figure 3. Photo studio light box, with one of its sides open.

Before any measurements were taken, the sample was dried in an oven at 50 °C for 1 h. Then, the sample was placed in its position, and the recipient was filled with water using a pipette. After checking that the image was focused, a picture was taken every 2 min for 3 h, so each sequence was composed of a total of 90 images.

2.2. Image Processing Methods

Rock deformation was assessed by surface texture displacements, achieved by calculating the normalized digital image cross-correlation (DIC) between the first frame in the sequence and all the other frames. To increase the accuracy of this result, the correlation peak was interpolated as detailed below [13]. This is a preferred procedure to the other strategy, which consists of interpolating the image itself, as this procedure can induce systematic DIC errors due to inaccurate subpixel image reconstruction [14].

The DIC calculation was not applied to the whole specimen but to some specific ROIs. The size of these ROIs was determined after a texture size analysis by autocorrelation. Different ROIs of fixed sizes were randomly located in 50 different positions on the sample's surface. Autocorrelation was calculated for each ROI, as was the full width at half maximum of the autocorrelation peak to, thus, estimate the average size of the particles in ROIs. This procedure was repeated for the different ROI sizes, ranging from 5×5 px to 200×200 px. Figure 4 represents the mean values of the correlation peak of each ROI size for one particular sample and the green channel. Standard deviations are represented

as error bars. This figure depicts that small ROIs presented wide variability in detail size, depending on their position, and merely showed sample non-homogeneity. Note that for ROIs larger than 80×80 , peak width is stable and the variation with ROI location is relatively narrow. A similar graph was obtained for all the samples, with the stability region ranging between 80 and 120 px. Thus, our calculations were performed on ROIs measuring 100×100 px for all the samples. Note that the larger ROIs were, the longer the calculation time required, with no clear benefits obtained in quality results terms. Finally, we stress that after some preliminary tests [11], we decided to select the 20 px template smaller than the ROI on both sides to allow texture displacement without going beyond interrogation area limits.

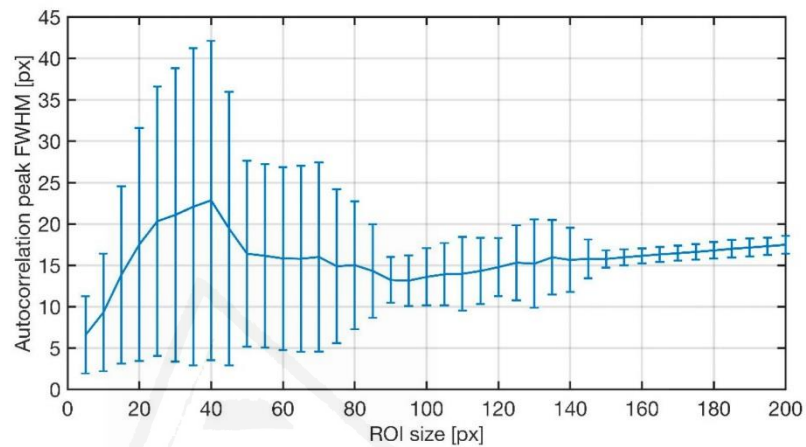


Figure 4. Size of the texture details estimated as the autocorrelation FWHM for the different region of interest (ROI) sizes for sample 1 in the green channel at 50 random positions. The results of other channels and samples showed similar outcomes.

The aim of the method is to determine the macroscopic strains due to hygric swelling. According to this, the deformation has been only calculated in six ROIs of the calculated size at specific positions close to the upper and lateral edges (see Figure 5). Every ROI from each frame was compared by DIC to the corresponding ROI in the same position from the first frame. This distribution allowed us to find the material strain in both the vertical (on both sides of the probe) and the horizontal (at three different heights) directions by using (1).

$$\varepsilon_{ij}(t) = \frac{Pos_{ROI_i}(t) - Pos_{ROI_i}(0)}{Pos_{ROI_i}(0) - Pos_{ROI_i}(0)} \quad (1)$$

where $Pos_{ROI_i}(t)$ is the location of ROI i for time t .

As previously mentioned, the DIC results were refined by interpolating the correlation peak following the procedure explained in [12]. According to the literature, both the interpolation area around the peak and the used fitting function are critical for obtaining accurate results. Two different strategies can be followed: taking a small interpolation area and using a quadratic function as a fitting function, or employing a large area with a Gaussian function. In order to select the best method, we implemented both strategies on the particular rock texture by imposing a synthetic displacement and comparing the results obtained by both methods.

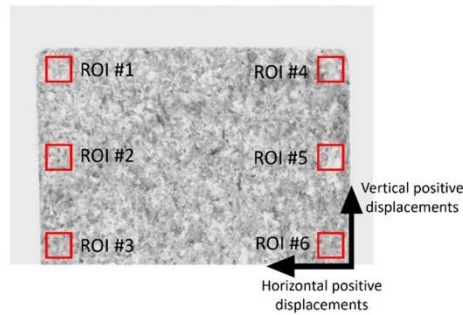


Figure 5. Number and location of ROIs for which the image processing was done and the direction of the positive displacements.

This test was prepared by taking the six ROIs represented in Figure 5 from the first picture of a sequence and numerically displacing by a total of 10 px in the vertical direction in steps of 0.1 px. No horizontal displacement was imposed. The displaced versions were calculated by the Fourier transform shifting property. The synthetic displacement was then compared to the shift obtained by DIC using a quadratic interpolation function in a 3×3 neighborhood and a Gaussian fitting function in an 11×11 neighborhood around the correlation maximum, according to the recommendations in [12]. The errors found for each ROI are represented in Figure 6, which clearly shows that quadratic fitting of a small neighborhood around the correlation peak was the best strategy.

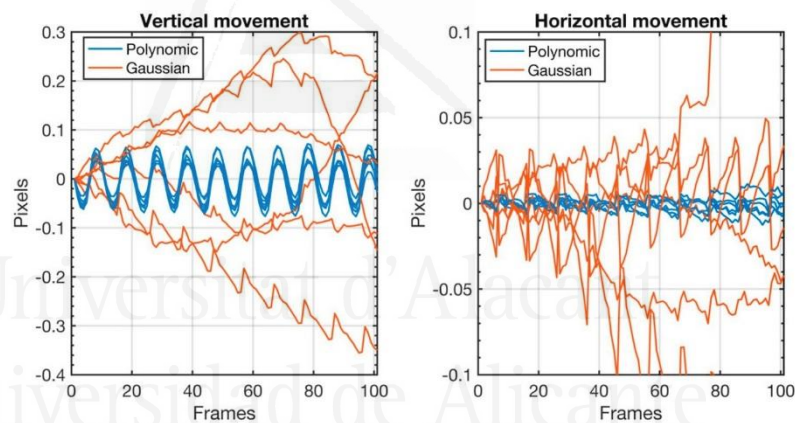


Figure 6. Location errors obtained for all the ROIs using polynomic and Gaussian fits in vertical (left) and horizontal (right) directions.

For the polynomic function, errors were bound, while errors showed no clear trend for Gaussian fits. Note that the errors obtained with Gaussian fitting functions were much larger than polynomic ones. The standard deviation values for polynomic fits were one order of magnitude higher in the direction of displacements than in the other direction, and the maximum value was 0.045 px. As a result of this analysis, a polynomic peak fit with a 3×3 neighborhood size was selected for the refinement peak. Finally, Figure 7 summarizes the calculation algorithm that was applied to analyze the samples.

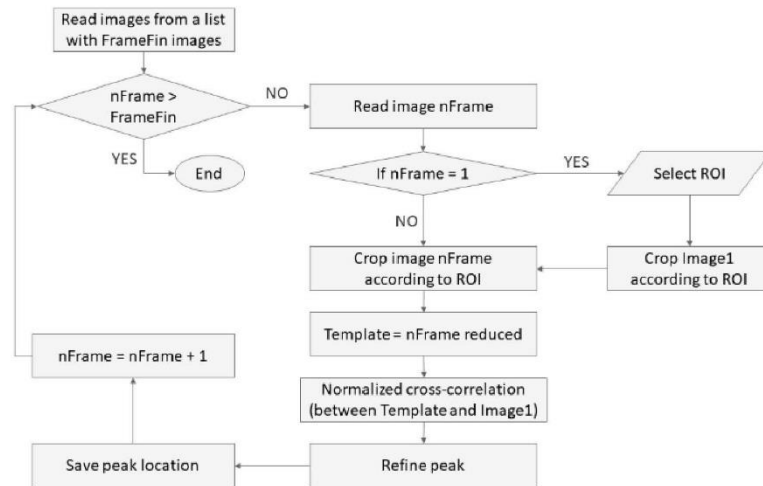


Figure 7. Flow chart for the image processing used in this work.

3. Results and Discussion

A final check was performed before continuing with the experiments. It is known that long image acquisition times can imply image distortions due to slight movements of the camera or supporting systems, or sensor deformation due to heat. Therefore, we implemented a “dry” experiment: i.e., the probe was filmed and measured without adding water. From the ROI location for each time, strains were obtained by (1). As the system is supposed to be static, all the obtained shifts may be assigned to experimental errors.

Figure 8 represents the absolute displacement of each ROI from initial position, where we can see a clear drift in the setup, mainly in the vertical direction, which can be up to 15 μm . This behavior was repeated in all the experiments similarly, so this was probably an effect of camera heating, which caused the support to loosen and made the camera move down due to objective weight. Notice also that the results present small instabilities of the order of 0.25 μm or, equivalently, 0.1 px, which is approximately twice the error obtained with the synthetic sequence in the movement direction (see Figure 6). These errors may be due to Gaussian noise in the image and fitting errors in the subpixel calculation algorithms. These fluctuations can be cancelled out by applying smoothing filters in the signal, but since they do not distort the main trend of the result, we preferred to show the data as obtained.

According to these data, our method is not advisable for absolute measurements. However, according to Equation (1), as we were interested in relative positions and strains, it was possible that overall drifting did not affect our results. Figure 9 represents the horizontal and vertical strains obtained for each pair of corresponding ROIs from the images obtained in this dry experiment. Note that in the vertical case, in the same color, we show the strains obtained for each pair of ROIs that are located at the same height, but on different sides of the samples.

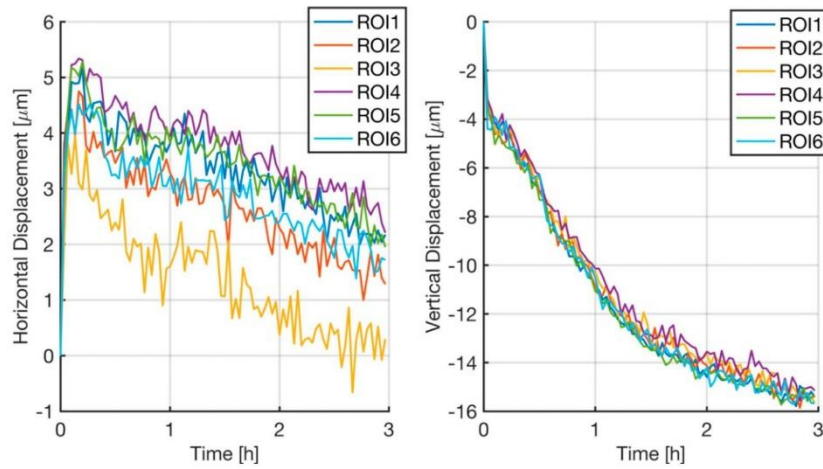


Figure 8. The measured horizontal (left) and vertical (right) shifts for a static dry experiment for each ROI. The results are shown for sample 1, but similar graphs were obtained for all the measured samples.

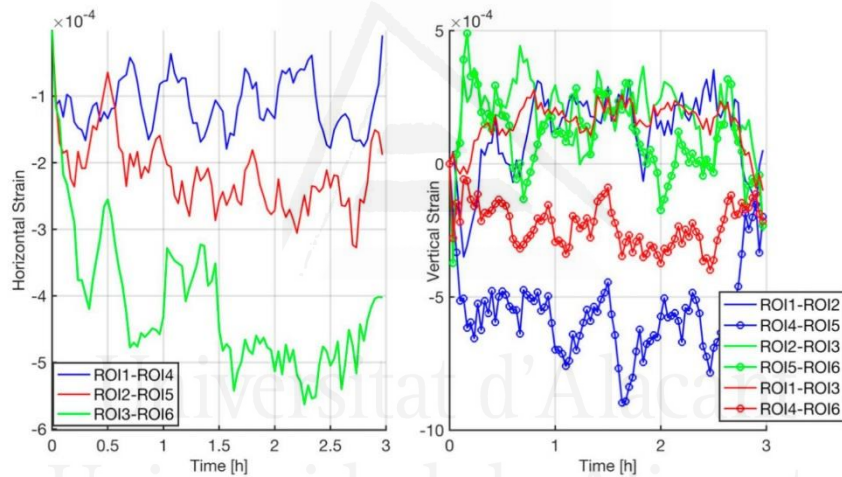


Figure 9. Measured horizontal (left) and vertical (right) strains for a static dry experiment for each ROI. The results are shown for sample 1, but similar graphs were obtained for all the measured samples.

Despite experimental instabilities, the method’s accuracy was acceptable with error peak values below 1×10^{-3} . Mean errors and their standard deviations are presented in Tables 1 and 2. From them, we take the worst obtained case as the error of this method, which was $(0.56 \pm 0.16) \times 10^{-3}$.

Table 1. Horizontal strain error for a long-term still sequence of the recorded images.

| ROIs | Horizontal Strain Error ($\times 10^{-3}$) | |
|------|--|--------|
| | MEAN | STD |
| 1–4 | −0.1110 | 0.0411 |
| 2–5 | −0.2163 | 0.0560 |
| 3–6 | −0.4175 | 0.1013 |

Table 2. Vertical strain error for a long-term still sequence of the recorded images.

| ROIs | Vertical Strain Error ($\times 10^{-3}$) | |
|------|--|--------|
| | MEAN | STD |
| 1–2 | 0.1167 | 0.1467 |
| 2–3 | 0.1841 | 0.1172 |
| 1–3 | 0.1504 | 0.0833 |
| 4–5 | −0.5586 | 0.1580 |
| 5–6 | 0.0938 | 0.1461 |
| 4–6 | 0.2313 | 0.0823 |

It is worth noting that errors in the vertical deformations are slightly larger and more disperse than in the horizontal direction. On the one hand, this may be due to the camera moving, but also to the shorter distance between ROIs. In fact, distance between the lateral ROIs was 3-fold longer than the distance between two consecutive vertical ROIs. As the strain was inversely proportional to the initial position of ROIs, according to Equation (1), the expected error was 3-fold bigger.

Figure 10 depicts the relative displacement measured for all the ROIs in sample 1 during the experiment. Compared to the error graphs in Figure 8, we can see that the displacement measured in the wet experiment was much larger than in the dry experiment. This means that the influence of the error must be considered to be minimum. For the vertical displacement, we see a clear movement due to hydration in the vertical direction. The curve is the typical one that has been observed in other experiments [15], with rapid swelling and a slow stabilization phase. Horizontal displacements were more difficult to interpret as the ROIs on the left and right sides were expected to move in opposite directions. However, the obtained results showed that the movement of all the ROIs went in the same direction, although the displacement of the three ROIs on the left was always less than that on the right. This effect could be due to sample rotation from the irregular expansion of the base or a composition camera drift effect and rock expansion.

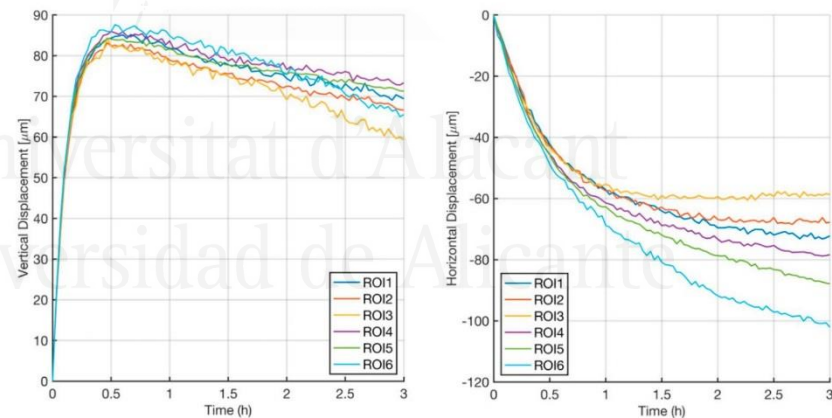


Figure 10. Measured vertical (left) and horizontal (right) shifts for each ROI in a wet experiment. The results are shown for sample 1, but similar graphs were obtained for all the measured samples.

In order to better understand this situation, Figure 11 represents the horizontal and vertical strains obtained for all four samples. As strains represent relative displacement between ROIs, all the global effects on the sample (rigid body movements) can be cancelled out, and all the observable effects may be due to internal forces.

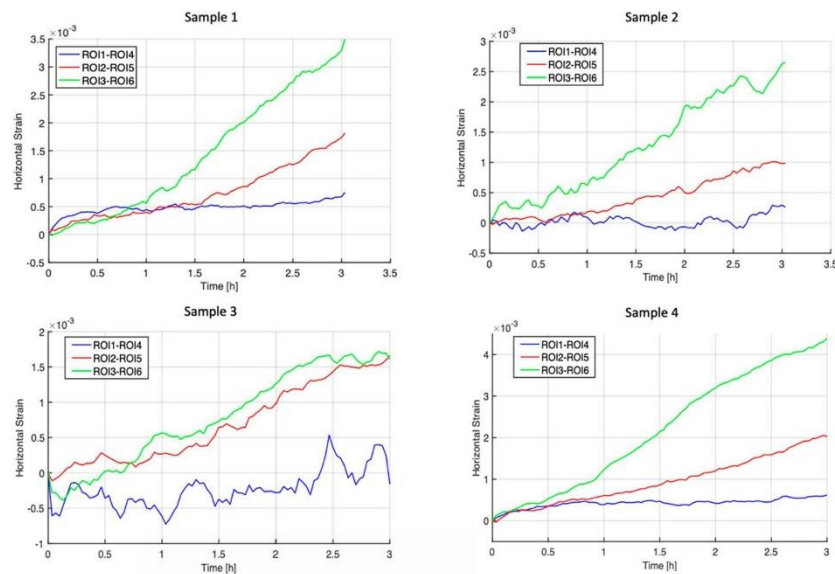


Figure 11. Measured horizontal strains for the four hydrated samples.

As the sample was submerged at the bottom, horizontal swelling was expected to be greater at the bottom than at the upper top. This was observed in all the horizontal displacement graphs, where each pair of ROIs presented minor deformations, because they were further away from water. The observed behavior was similar for all the samples, and the observed variability can be explained by rock non-homogeneity, which can produce samples with slightly different compositions and structures. Note that the upper part of all the samples, which corresponded to ROIs #1 and #4, was much less affected by hygric swelling.

The results for the vertical strains were not so clear. Once again, it would seem that the upper part of samples, which corresponded to the strains between ROIs 1–2 and 4–5 (depicted in blue in Figure 12), were less affected by swelling. The green line, which represents the strains for the ROIs closer to the water level, presented a positive strain during the first 2 h, which means that the rock expanded. After this time, however, the strain decreased, possibly because an early capillary pressure effect was followed by relaxation and water circulation through pores to cause slight shrinkage in the vertical direction but maintained expansion in the horizontal direction. We underline that the strains observed in the vertical directions were of the same order as the peak error limits observed in the dry test (see Figure 9), hence the possibility of the oscillations herein shown not being so marked. However, the trend was very clear, and this behavior was repeated in all four samples with the same shape approximately 2 h after the experiment starts. At that time, as the drift was not as important, the effect cannot be fully explained by experimental errors, and the total displacement in Figure 10 is several orders of magnitude larger than the total displacement observed in the dry experiment (see Figure 8).

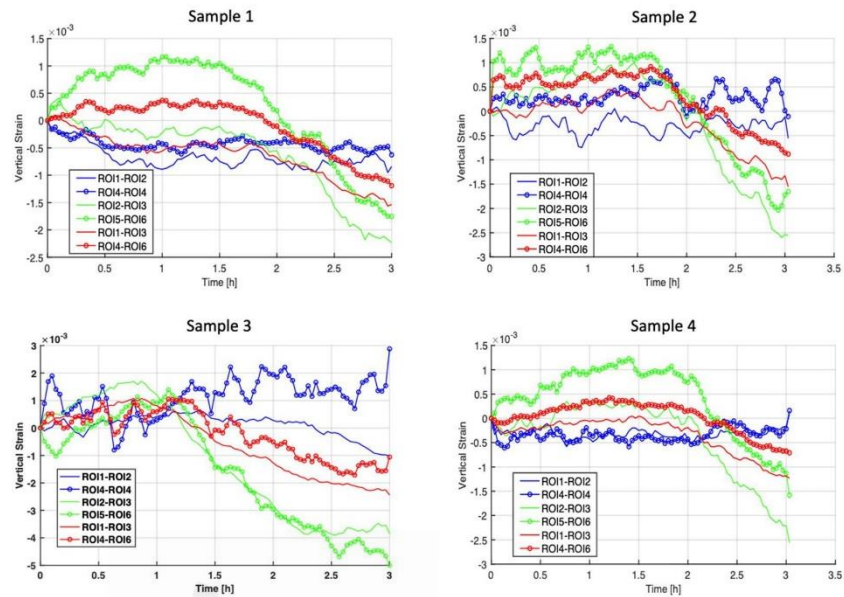


Figure 12. Measured vertical strains for the four hydrated samples.

In order to get more insight about the rock behavior, we have calculated different deformation maps of sample 1. In Figure 13, we present the results obtained for the horizontal and vertical deformations at 1, 2 and 3 h after the water addition. To calculate the deformation, the image of the rock has been tessellated in ROIs of 100×100 px with an overlapping of 50 px with the adjacent ROIs. Displacement through correlation has been calculated between the corresponding ROIs of the initial picture and the subsequent images. The relative deformation has been obtained between every two alternate ROIs in the considered direction. Finally, in order to avoid outliers and obtain a softer appearance, a 3×3 median filter has been applied to the result.

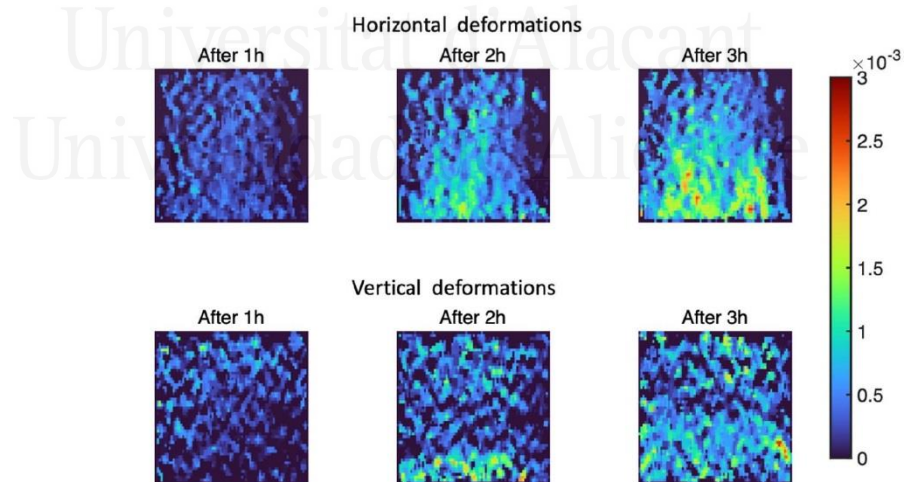


Figure 13. Deformation maps obtained for sample 1 in horizontal and vertical directions at different moments after the rock immersion.

The maps confirm the results in Figures 11 and 12. As it appeared there, the horizontal deformation is stronger than the vertical one. In the horizontal direction, one can see that the lower part of the rock, which is closer to the water surface, suffers a larger dilation than the upper part. In the vertical direction, the effect is not so strong. One can see a clear deformation in the lower part of the sample after 2 h that seems to vanish and spread through all the sample at the end of the experiment.

Note also that the maps show clear inhomogeneities in the distribution of the deformations, which reflects the distribution of particles with different swelling properties through the rock volume.

4. Conclusions

This manuscript discusses a method to measure the deformation of a sandstone rock partially submerged in water due to hygric swelling. This method uses a simple camera and a telecentric objective, which allows macroscopic deformations to be measured. The method uses digital image cross-correlation on small rock areas so that changes in both the vertical and horizontal directions can take place.

Both the numerical and experimental procedures are described, and possible errors were analyzed. According to our calculations, the calculation error method was below 0.1 px or the equivalent to 0.04 μm . Experimental implementation requires long time sequences, which means that marked mechanical and thermal stability conditions are required. In our case, the required stability was not achieved, and an image drift was observed. However, the key parameter in our measurement was strain, which is a relative magnitude, and therefore, rigid body movements did not affect measurements.

Four sandstone rock samples were measured by the herein presented methods. The horizontal strain demonstrated that rock deformation was not uniform, which became larger the closer it was to the submerged part. Our measurements showed a clear dilation in the horizontal direction, which is expected for a partially submerged porous body. The strain measurements in the vertical direction displayed unexpected behavior, so deformation maps at different moments during the experiment were calculated. The maps showed clearly the distribution of the deformations and revealed important inhomogeneities in their distribution, reflecting the composition of the rock.

The method showed good capabilities in measuring local displacement and, from this, the strain of a small rock sample in both the vertical and horizontal directions. The methods herein followed, despite the limitations described, are simpler to implement than those that involve microscopic analyses and may be of much interest for analyzing macroscopic rock dynamics as construction material.

Note that, with the exception of the telecentric lens, which can be replaced by a standard objective with proper distortion calibration, the setup is relatively cheap, which makes this procedure affordable for students or for preliminary tests in research labs. The main drawback of the method is the thermal drift observed. Future works should pay attention to compensate this issue without increasing the price and the complexity of the setup.

Finally, we would like to add that the proposal demonstrates that natural textures can be used as reliable targets for DIC techniques. Further analysis and characterization of textures is needed in order to optimize the subpixel correlation results.

Author Contributions: Conceptualization, D.M. and B.F.; methodology, B.F.; software, D.M.; validation, M.-B.T., D.M. and B.F.; formal analysis, B.F.; investigation, M.-B.T.; resources, B.F.; data curation, D.M.; writing—original draft preparation, B.F.; writing—review and editing, D.M.; visualization, M.-B.T.; supervision, D.M.; project administration, B.F.; funding acquisition, M.-B.T. and B.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Generalitat Valenciana and the European Social Fund (FSE) through the Recruitment of Predoctoral Research Staff ACIF/2018/211 included in the FSE Operational Program 2014–2020 of the Valencian Community (Spain). Belén Ferrer and María-Baralida Tomás acknowledge the support from the Generalitat Valenciana through Project GV/2020/077.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Software for correlation is available at <http://rua.ua.es/dspace/handle/10045/110141> (accessed on 15 February 2021). Images or videos of the samples are disposable upon demand.

Acknowledgments: The authors acknowledge David Benavente for providing the rock samples used in the experiments included in this work.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Di Benedetto, C.; Cappelletti, P.; Favaro, M.; Graziano, S.F.; Langella, A.; Calcaterra, D.; Colella, A. Porosity as Key Factor in the Durability of Two Historical Building Stones: Neapolitan Yellow Tuff and Vicenza Stone. *Eng. Geol.* **2015**, *193*, 310–319. [CrossRef]
2. Colas, E.; Mertz, J.D.; Thomachot-Schneider, C.; Barbin, V.; Rassinoux, F. Influence of the Clay Coating Properties on the Dilation Behavior of Sandstones. *Appl. Clay Sci.* **2011**, *52*, 245–252. [CrossRef]
3. Siegesmund, S.; Dürrast, H. Physical and mechanical properties of rocks. In *Stone in Architecture*; Siegesmund, S., Snethlage, R., Eds.; Springer: Berlin/Heidelberg, Germany, 2011; pp. 97–225. ISBN 978-3-642-14474-5.
4. Delage, P. On the Thermal Impact on the Excavation Damaged Zone around Deep Radioactive Waste Disposal. *J. Rock Mech. Geotech. Eng.* **2013**, *5*, 179–190. [CrossRef]
5. Sutton, M.A.; Orteu, J.-J.; Schreier, H.W. *Image Correlation for Shape, Motion and Deformation Measurements*; Springer: Boston, MA, USA, 2009; ISBN 978-0-387-78746-6.
6. *MATLAB*; The MathWorks, Inc.: Natick, MA, USA, 2019.
7. Fauchille, A.-L.; Hedan, S.; Valle, V.; Prêt, D.; Cabrera, J.; Cosenza, P. Effect of Microstructure on Hydric Strain in Clay Rock: A Quantitative Comparison. *Appl. Clay Sci.* **2019**, *182*, 105244. [CrossRef]
8. Valle, V.; Hedan, S.; Cosenza, P.; Fauchille, A.L.; Berdjane, M. Digital Image Correlation Development for the Study of Materials Including Multiple Crossing Cracks. *Exp. Mech.* **2015**, *55*, 379–391. [CrossRef]
9. Carrier, B.; Wang, L.; Vandamme, M.; Pellenq, R.J.-M.; Bornert, M.; Tanguy, A.; Van Damme, H. ESEM Study of the Humidity-Induced Swelling of Clay Film. *Langmuir* **2013**, *29*, 12823–12833. [CrossRef] [PubMed]
10. Wang, L.L.; Bornert, M.; Yang, D.S.; Héripré, E.; Chanchole, S.; Halphen, B.; Pouya, A.; Caldemaison, D. Microstructural Insight into the Nonlinear Swelling of Argillaceous Rocks. *Eng. Geol.* **2015**, *193*, 435–444. [CrossRef]
11. Ferrer, B.; Tomás, M.B.; Benavente, D.; Mas, D. Use of Image Correlation to Measure Hygric Swelling in Rocks. In *Optics, Photonics and Digital Technologies for Imaging Applications VI*; Schelkens, P., Kozacki, T., Eds.; International Society for Optics and Photonics: Paris, France, 2020; p. 64.
12. Tomás, M.-B.; Ferrer, B.; Mas, D. Influence of Neighborhood Size and Cross-Correlation Peak-Fitting Method on Location Accuracy. *Sensors* **2020**, *20*, 6596. [CrossRef] [PubMed]
13. Lei, X.; Jin, Y.; Guo, J.; Zhu, C. Vibration Extraction Based on Fast NCC Algorithm and High-Speed Camera. *Appl. Opt.* **2015**, *54*, 8198–8206. [CrossRef] [PubMed]
14. Schreier, H.W.; Braasch, J.R.; Sutton, M.A. Systematic Errors in Digital Image Correlation Caused by Intensity Interpolation. *Opt. Eng.* **2000**, *39*, 2915–2921. [CrossRef]
15. Benavente, D.; Lock, P.; García del Cura, M.Á.; Ordóñez, S. Predicting the Capillary Imbibition of Porous Rocks from Microstructure. *Transp. Porous Media* **2002**, *49*, 59–76. [CrossRef]