



**UNIVERSIDAD DE CONCEPCIÓN
FACULTAD DE INGENIERÍA
DEPARTAMENTO DE INGENIERÍA ELÉCTRICA**



Superresolución en un Sistema Plenóptico Infrarrojo de Onda Corta Multiespectral

POR

Joaquín Alonso Lermenda Sandoval

Tesis de postgrado presentada a la Facultad de Ingeniería de la Universidad de Concepción para optar al grado académico de Magíster en Ciencias de la Ingeniería Mención Ingeniería Eléctrica

Profesor(es) Guía
Sergio Neftalí Torres Inostroza

Junio 2025
Concepción (Chile)
©2025 Joaquín Alonso Lermenda Sandoval

©2025 Joaquín Alonso Lermada Sandoval

Se autoriza la reproducción total o parcial, con fines académicos, por cualquier medio o procedimiento, incluyendo la cita bibliográfica del documento.

Para Matías, cuyo trabajo y esfuerzo demuestran que si se pueden lograr los sueños.

Agradecimientos

Quiero partir por agradecer al profesor Sergio Torres, por confiar en mis habilidades y otorgarme oportunidades de seguir investigando e incluso permitirme trabajar con el en su empresa de radiometría, donde aplique mi trabajo en la industria y pude conocer Australia, un país que nunca imaginé conocer. Por esto y por su constante apoyo y recomendación de mi nombre le estoy agradecido.

Deseo agradecer también a Guillermo Machuca, quien me ayudó en el trabajo presente y dedico incontables horas a reunirse conmigo y ayudarme con los resultados y conocimientos vitales del trabajo, sin el este trabajo hubiese sido mucho más difícil de realizar.

La gente de los laboratorios de procesamiento de imagen, en particular a Francisco Pérez quien siempre estuvo dispuesto a responderme dudas sobre óptica y otros conceptos.

Mi familia, que siempre me ha apoyado en cada decisión de vida que he tomado y me han permitido esforzarme y demostrar de lo que soy capaz.

Finalmente, al departamento de ingeniería eléctrica y todos sus profesores que me apoyaron durante el magíster, enseñándome conocimientos vitales en el campo de la investigación y formándome como un ingeniero más completo.

Sumario

El avance e invención de nuevas técnicas de captura de imágenes infrarrojas en los últimos años han permitido no solamente obtener imágenes más detalladas, sino que además obtener información de una escena cualesquiera que, sin las previamente mencionadas técnicas, no se podría conocer. Los ámbitos en los cuales se desempeña el infrarrojo (IR) son tan variados como únicos, implicando así que no existe una sola solución correcta para los problemas existentes. Sin embargo, el uso de esta tecnología posee desventajas notorias respecto a otras bandas espectrales, particularmente, la baja resolución de imagen producto de la tecnología del sensor IR, lo que provoca imágenes de baja calidad en una banda espectral que ya posee problemas como el Fixed Pattern Noise (FPN) y blur. La presente investigación se centra alrededor del desarrollo de un sistema plenóptico con las capacidades de resolver la baja resolución del espectro mediante superresolución (SR), en particular del short wave infrared (SWIR), utilizando además un método de eliminación de ruido con características multiespectrales, para generar un sistema óptico completo. Se aumenta la resolución del SWIR mediante iteraciones de un método Plug and Play basado en Alternating Direction Method of Multiplier (ADMM) para asemejarla a otras bandas espectrales como el near infrared (NIR). Posteriormente, se emplean métricas de calidad de imagen para comparar con los resultados obtenidos por otros métodos de SR, desarrollando pruebas de simulación y experimentales. Se eligen escenas donde el objeto de interés a medir posea detalle suficiente para poder realizar comparaciones y donde la componente espectral muestre diferencias, evaluando así el comportamiento de la resolución en varios espectros. Lo anterior resulta en imágenes con resolución aumentada, sin ruido y características multiespectrales variadas que de acuerdo a métricas de imágenes comunes como lo son PSNR, SSIM y una del IR llamada RLP. Comparando con otros métodos de SR, la mejora del método propuesto no solo supera en SSIM, sino que también en eliminación de ruido de manera considerable, generando así un sistema óptico completo.

Índice

1. Capítulo 1: Introducción	10
1.1. Hipótesis	13
1.2. Objetivos específicos	13
2. Estado del Arte	13
2.1. Eliminación de FPN y otros artefactos	13
2.2. Lightfield y plenóptica	15
2.3. Superresolución	17
3. Marco Teórico	18
3.1. El Infrarrojo	18
3.1.1. Bandas espectrales: NIR, SWIR, MWIR y LWIR	18
3.1.2. El cubo Multiespectral	19
3.1.3. El FPN en el IR	20
3.2. Lightfield y Fourier Slice Photography Theorem (FSPT)	21
3.2.1. Lightfield	21
3.2.2. FSPT	22
3.2.3. Eliminación de ruido mediante plenóptica	23
3.3. SR en el IR	25
4. Modelo propuesto y Setup	29
4.1. Modelo propuesto	29
4.2. Métricas	30
4.2.1. Roughness Laplacian Index	30
4.2.2. Structural Similarity Index	31
4.2.3. Peak Signal to Noise Ratio	31
4.3. Setup - Montaje Experimental	32
4.3.1. Simulación	33
4.3.2. Evaluación Experimental	33
5. Resultados	36
5.1. Simulación	36

5.1.1. Simulación - Tamaño de Matriz	36
5.1.2. Simulación - Tipo de Ruido	41
5.2. Experimental	43
5.2.1. Experimental - Resoluciones	43
5.2.2. Experimental - Longitudes de onda	44
6. Discusión	45
7. Conclusiones y trabajo a futuro	47
8. Anexo	49
8.1. Simulación - Tamaño Matriz - OBJ2	49
8.2. Simulación - Tipos de Ruido - OBJ2	50

Índice de cuadros

1. Comparación de métodos de captura plenóptica. Cada método posee ventajas y desventajas a considerar a la hora de implementar el sistema, convirtiéndolo en dependiente de la aplicación. . 16

Índice de figuras

3.1. Desglose del espectro IR, las secciones corresponden a NIR(<i>Near Infrared</i>), SWIR(<i>Short Wave Infrared</i>), MWIR(<i>Mid Wave Infrared</i>) y LWIR(<i>Long Wave Infrared</i>) [1].	18
3.2. Cubo multiespectral e Hiperespectral, cortesía de NIREO.	19
3.3. FPN presente en una cámara Xeva 320 SWIR. Notar como líneas verticales de intensidad aumentada plagan la imagen. Cada sensor y banda poseen un ruido característico distinto a los demás, por lo que no existe una solución única a este fenómeno.	21
3.4. Parametrización de Ren Ng del <i>light slab</i> . Se define un lente y un sensor y la intersección de una línea entre estos planos define un rayo de luz.	22
3.5. Re-parametrización del <i>lightfield</i> moviendo el sensor desde F a F'	23
3.6. Equipo de medición plenóptica consistente en una cámara y un desplazador mecánico, los cuales son utilizados para recopilar el LF en forma matricial del tamaño deseado.	24
3.7. Cubo multiespectral de 4 dimensiones generado por el método descrito. Notar como se puede desglosar el cubo de acuerdo a su longitud de onda λ o su profundidad α y revisar los píxeles específicos de la imagen (x,y)	25
3.8. Modelo de ruido ubicado en el infinito óptico desarrollado por Coelho. El cambio corresponde a uno de concepto que permite mirar desde otro punto de vista la problemática [2].	26
3.9. Aplicación de la eliminación de ruido mediante reenfoque propuesta por Coelho. El valor de α_n denota el plano de enfoque actual, A) $\alpha = 0$, B) $\alpha = 80$, C) $\alpha = 215$	26
3.10. Diferencias en resolución entre espectro SWIR (a la izquierda) y espectro visible (a la derecha). Al enfocarse en detalles como el asa del bidón, se pueden apreciar como los bordes y resolución del objeto están más definidos en el espectro visible.	27
4.1. De izquierda a derecha, el equipo de captura y modo de trabajo de este. A) La cámara Xeva montada en un desplazador mecánico con su respectiva montura de filtro para generar el cubo multiespectral. B) El movimiento del desplazador y la matriz $M \times N$ generada por este. C) El LF generado por el equipo, con cada imagen desplazada en el eje X e Y dependiendo de la fila y columna donde se ubique, la separación y cantidad de fotos son definidas por el usuario.	32
4.2. Escena IR a la cual se le aplicará la degradación de imagen y luego mejora mediante el método propuesto y los de Hardie y Meza. El OBJ1 corresponde al bidón de plástico más cercano relleno de agua, el OBJ2 corresponde al envase de plástico más lejano relleno de agua, el agua se puede apreciar como el color negro del envase.	33

4.3.	Esquema del procedimiento de captura para la fase de simulación. A) Se parte por capturar un LF de una escena SWIR calibrada sin ruido, luego se le añaden los artefactos de ruido, blur y decimación a la mitad, obteniéndose así un LF con imágenes de 160x128 y artefactos presentes. B) Se aplica el FSPT para generar Fourier Refocus Denoising (FRD) eliminando así el ruido en los objetos de interés OBJ1 y OBJ2. C) Se eligen ambos objetos enfocados y sin ruido por separado mediante el RLP para aplicarles el método de ADMM SR, con esto se obtiene cada imagen por separado con resolución aumentada, volviendo así a 320x256. D) Si bien se pueden comparar los objetos de interés por separado, se elige combinar ambos con un algoritmo de renfoque All in Focus (AiF) para tener una comparación visual más pareja con otros métodos. Esta fusión de imágenes no afecta la calidad de la imagen obtenida. E) Utilizando la misma entrada de LF con artefactos y decimación, se aplican los métodos de Hardie y Meza para comparar los resultados con el método propuesto.	34
4.4.	Escena IR con FPN inherente de la cámara SWIR Xeva 320x256. El OOI es una de las hojas de la planta, la cual presenta detalles elevados al utilizarse la banda IR correcta para detectar estrés hídrico.	35
4.5.	Esquema del procedimiento de captura para la fase experimental. De izquierda a derecha, se muestra como se introduce el LF con ruido de la cámara y sus sucesivos aumentos de resolución. Luego de aplicar el FRD, se aumenta la resolución a 640x512 y luego a 1280x1024, con el detalle de una de las hojas en la fila inferior.	35
5.1.	Variaciones en calidad de imagen obtenida dependiendo del tamaño de matriz empleado para cada método. Cada fila corresponde a un método distinto y las columnas separan los tamaños de matrices utilizados. Visualmente, uno puede percatarse de como los métodos varían sus resultados dependiendo del tamaño utilizado para la matriz.	36
5.2.	Planos focales para el OBJ1 y OBJ2 antes de aplicar SR y fusión de imagen. El tamaño de 2x2 genera planos focales con ruido para ambos objetos enfocados, mientras que el tamaño de 10x10 genera planos focales sin ruido, lo que provoca diferencias considerables en calidad al aplicar SR y fusión de imagen.	37
5.3.	Comportamiento del RLP a lo largo de distintos tamaños de matriz tanto para blur como para el detalle (NU). Es necesario recordar que se está buscando el menor blur y un peak de ruido que lo acompañe, el motivo de esto es que esto será el plano enfocado, ya que interpretará el ruido como detalle la métrica. A) Para la matriz de 10x10 el plano focal está marcado por el peak de FPN y declive en blur en el frame 198. B) Para la matriz de 8x8 el plano es el 170. C) Para la matriz de 4x4 el plano focal es el 80. D) Para la matriz de 2x2 el plano focal es el 50	38

- 5.4. PSNR de OBJ1 y OBJ2 luego de aplicar el procedimiento descrito en la figura 4.3. A) el OBJ1 debido a su cercanía parte en valores más elevados de PSNR para los 3 métodos. Se observa que a medida que el tamaño de matriz aumenta, el método propuesto mejora en calidad de imagen, apreciándose la mejora en comparación a otros métodos desde 4x4 en adelante. B) La escena completa si bien sigue el mismo comportamiento, tiene valores de PSNR más cercanos entre los métodos, esto es producto de que los métodos de Hardie y Meza no lidian con el desenfoque y tienen la escena completa enfocada, mientras que el método propuesto realizó fusión de imagen para solo los dos objetos de interés, sin considerar el resto de la escena desenfocada. Pese a esto, el PSNR del método propuesto es superior debido a la eliminación de ruido. 39
- 5.5. SSIM de OBJ1 y escena completa luego de aplicar el procedimiento descrito en la figura 4.3. A) el OBJ1 muestra un comportamiento similar al PSNR, donde los valores van en aumento a medida que el tamaño de matriz aumenta para el método propuesto y disminuyen para los otros métodos. En esta ocasión, la mejora del método propuesto se nota a partir de la matriz de 8x8 en adelante B) La escena completa sigue el mismo comportamiento de ir aumentando conforme el tamaño de matriz aumenta, sin embargo, los valores son mayores en general, esto es producto de que al considerar más de la escena, los pequeños detalles como bordes tienen menos incidencia en el valor final del SSIM. 39
- 5.6. RLP de OBJ1 y escena completa luego de aplicar el procedimiento descrito en la figura 4.3, con la columna de la izquierda representando el OBJ1 y la de la derecha la escena completa. En el recuadro superior izquierdo, el OBJ1 exhibe un comportamiento de blur bajo para el método propuesto en comparación a los otros métodos, debido principalmente a la eliminación de ruido y mayor detalle. En el recuadro inferior izquierdo, el RLP_{NU} aumenta a medida que el tamaño de matriz aumenta para el método propuesto, mientras que disminuye para los otros métodos, señalando así un mayor detalle obtenido en la imagen para el propuesto. Resultados similares se pueden observar para la escena completa en la columna derecha. 40
- 5.7. De izquierda a derecha se tiene la imagen de referencia, la imagen con ruido gaussiano y la imagen con ruido de grilla. El ruido para ambos casos considero una desviación estándar de 0.01. 41
- 5.8. PSNR de OBJ1 y escena completa luego de aplicar el procedimiento descrito en la figura 4.3 para una matriz de 8x8. A) Para el OBJ1, si bien el valor de PSNR de gauss es superior al de grilla, ambos mantienen superioridad en el método propuesto sobre los métodos de Meza y Hardie. B) Para la escena completa, los métodos obtienen valores más cercanos entre ellos, manteniendo eso si la superioridad del método propuesto. 42

5.9. SSIM de OBJ1 y escena completa luego de aplicar el procedimiento descrito en la figura 4.3 para una matriz de 8x8. A) Para el OBJ1, los valores de SSIM siguen una tendencia similar a los valores de PSNR, con una clara superioridad del método propuesto sobre los otros dos B) Para la escena completa, los métodos obtienen valores más cercanos entre ellos, manteniendo eso si la superioridad del método propuesto.	42
5.10. Aumento en resolución de imagen aplicando la interpolación bicúbica y el método propuesto a una imagen A, con aumentos de resolución de x2 y x4. Para las imágenes B 1-3, la interpolación bicúbica no logra superar el FPN presente en la imagen y los resultados son pobres para el aumento de resolución. Para las imágenes C 1-3 Se logra eliminar el ruido y lograr un aumento de resolución exitoso, aumentando el detalle de la imagen original y por ende su calidad. . . .	43
5.11. RLP de OOI después de aplicar el método bicúbico y el método propuesto. A) El RLP_{NU} de la imagen con 640x512 muestra la superioridad del método propuesto, que se ve aumentada al pasar a la resolución de 1280x1024, cabe destacar que el contexto de porque se asemejan los valores en la resolución x2 es debido a la fuerte presencia de ruido en el método bicúbico que es interpretado como detalle por la métrica. B) El blur es notoriamente prevalente para el método bicúbico, con el método propuesto disminuyendo el valor a medida que aumenta la resolución. .	44
5.12. Diferentes longitudes de onda para la misma escena, mostrando la imagen inicial sin ruido, la imagen con el método propuesto y un aumento en el OOI. A) La imagen inicial con el ruido ya eliminado mediante el método propuesto con una resolución de 320x256 para 5 distintas longitudes de onda. B) Imágenes resultantes de aplicar la sección de aumento de resolución del método propuesto. Para este caso se llega a la resolución de 1280x1024. A 1-5) Acercamiento del objeto de interés OOI para las imágenes de la fila A, es decir, sin ruido tras la primera parte del método propuesto. B 1-5) Acercamiento del objeto de interés OOI para las imágenes de la fila B, es decir, con la parte de aumento de resolución incluida.	45
6.1. Escena de agua caliente capturada con dos cámaras distintas A) La cámara Xeva SWIR captura el nivel de agua en la botella pero con una resolución de 320x256. B) La cámara Gobi LWIR captura la temperatura del agua con una resolución mayor de 640x480 pero sin determinar el nivel del agua.	46
6.2. Escena capturada por distintas cámaras y su combinación en fusión de imagen. A) La botella en SWIR con el nivel de agua identificado por el color negro. B) La botella en LWIR con la temperatura irradiada en blanco. C) la fusión de ambas imágenes con el detalle del nivel y la emisión.	47

- 8.1. Comportamiento del RLP a lo largo de distintos tamaños de matriz tanto para blur como para el detalle (NU). Es necesario recordar que se está buscando el menor blur y un peak de ruido que lo acompañe, el motivo de esto es que esto será el plano enfocado, ya que interpretará el ruido como detalle la métrica. A) Para la matriz de 10x10 el plano focal está marcado por el peak de FPN y declive en blur en el frame 110. B) Para la matriz de 8x8 el plano es el 80. C) Para la matriz de 4x4 el plano focal es el 60. D) Para la matriz de 2x2 el plano focal es el 23. 49
- 8.2. SSIM y PSNR del OBJ2. A) El PSNR sigue el comportamiento de ir en aumento a medida que se aumenta el tamaño de matriz para el método propuesto, llegando a valores de 22 dB. B) El SSIM aumenta considerablemente para los tamaños de 8x8 y 10x10 para el método propuesto, alcanzando valores de 0.67 y 0.75 respectivamente. 49
- 8.3. Comportamiento del RLP para el OBJ2. A) El detalle de la imagen es consistente en ser mayor para el método propuesto para todos los tamaños de matriz, sin embargo, cabe recordar que el detalle con ruido pueden confundirse en estos tamaños pequeños. B) El blur se mantiene en niveles bajos para el método propuesto a lo largo de los distintos tamaños de matrices, demostrando la versatilidad en enfocar objetos a distancia. 50
- 8.4. SSIM y PSNR del OBJ2 para ruido gaussiano y de grilla. A) Los valores de PSNR son superiores para el método propuesto tanto para ruido gaussiano como de grilla. B) El SSIM mantiene valores mejores para el método propuesto en ambos tipos de ruido. 50
- 8.5. Comportamiento de RLP para ambos tipos de ruido. A) el detalle de la imagen NU es mayor para ambos tipos de ruido en el caso del método propuesto. B) El blur mantiene valores menores para el método propuesto para ambos tipos de ruido. 50

1. Capítulo 1: Introducción

La ciencia y la tecnología de imágenes IR son un campo en continuo crecimiento, que amplía las fronteras de la tecnología para aprovechar al máximo los fotones IR disponibles en la escena. Tiene numerosas aplicaciones que van desde diagnóstico médico, biología, astronomía, teledetección, detección de procesos industriales y vigilancia, por nombrar algunas[3, 4, 5, 6, 7, 8]. La banda espectral IR se divide en la banda espectral del infrarrojo cercano (NIR), que se extiende para longitudes de onda de aproximadamente $0,7 \mu m$ a $1 \mu m$; la banda espectral infrarroja de onda corta (SWIR), que se extiende para longitudes de onda de aproximadamente $1 \mu m$ a $1,7 \mu m$; la banda espectral infrarroja de onda media (MWIR), que se extiende para longitudes de onda de aproximadamente $3 \mu m$ a $5 \mu m$; y la banda espectral infrarroja de onda larga (LWIR), que se prolonga para longitudes de onda de aproximadamente $7 \mu m$ a $12 \mu m$. Esta última banda es de las más conocidas popularmente debido a su capacidad de capturar el calor humano y temperaturas similares. Las cámaras se posicionan dentro de una de estas bandas de acuerdo a las características de sus componentes electrónicas como los tipos de sensores empleados. Se elige, por tanto, la cámara dependiendo de la aplicación buscada, aprovechando la información adicional que entrega capturar una escena en IR.

Sin embargo, los sensores utilizados para el IR sufren de una serie de limitaciones que afectan enormemente la calidad de la imagen obtenida en varios aspectos. Lo anterior convierte la captura IR en una herramienta que, para ser utilizada eficientemente, debe resolver estos efectos detrimentales antes de analizar los resultados obtenidos. Debido a la tecnología empleada, las principales preocupaciones corresponden a la resolución espacial y angular de la imagen procesada por la cámara (es decir, que tan bien se puede separar un objeto de otro), artefactos como el ruido de patrón fijo (FPN) o lo borroso de la escena producto del efecto blur, degradan y alteran los valores de los píxeles. Lo anterior es altamente aleatorio y variable a lo largo del hardware empleado, es decir, cada cámara posee degradación de FPN y otros tipos de ruido distintos entre sensores incluso del mismo espectro. Los efectos negativos mencionados se complejizan aún mas cuando se busca un enfoque multiespectral, ya que por lo general las soluciones a artefactos se enfocan en la banda espectral entera de operación del sensor y no en una banda en específico. Esto convierte el análisis multiespectral en un problema de dificultad adicional cuando se desea resolver los problemas de calidad de imagen, sobretodo considerando que el análisis multiespectral es usado ampliamente en investigación de fenómenos visibles únicamente en esas bandas. Finalmente, los sensores IR (los sensores en general) no poseen por si solos la capacidad de capturar de la dirección de la luz y están limitados a obtener solamente la intensidad de esta, perdiendose así una dimensión completa de información relevante, lo que considerando las capacidades del IR de aportar información es una pérdida del potencial de las capacidades del sistema. Este ultimo concepto se denomina captura de lightfield (LF), que hace referencia a la información de dirección e intensidad de luz en una escena[9, 10].

La resolución de la imagen procesada es afectada enormemente por la tecnología del sensor y la longitud de onda que captura. Comunmente, existe una relación entre una serie de variables como lo son el tamaño del sensor, la composición de este y la implementación a la cámara, entre otros. Lo anterior introduce una serie de limitaciones físicas en el diseño y construcción de sensores, que a su vez limitan la calidad de la imagen obtenida por estos. Un sensor con una gran cantidad de detectores tendría una mejor resolución pero también aumentaría la interferencia de un detector sobre otro generando posiblemente ruido. Para capturar longitudes de onda mayor, un sensor utilizaría un material distinto en su construcción, que a su vez posee propiedades eléctricas diferentes a un sensor de menor longitud de onda. Si quisiera aumentar el tamaño del sensor porque no se pueden achicar mas los detectores, estaría la limitación física del espacio que el sensor usaría, volviendo poco práctico y caro para su uso. A lo anterior se le suma el hecho de que bandas como el LWIR detectan emisividad de la escena obteniendo resultados con bordes borrosos a diferencia de bandas como el SWIR y el NIR que detectan reflexiones, provocando diferencias de la calidad de imagen aceptable entre bandas y la imposibilidad de comparar bandas entre sí al ser de resoluciones distintas.

La resolución espacial no es el único parámetro que se ve afectado por cambios en la longitud de onda. El FPN y otros artefactos similares aumentan su fuerza y modifican su patrón espacial al alterar las longitudes de onda de NIR a LWIR, generando problemas de fidelidad de datos y claridad de la escena únicos para cada banda espectral, obteniéndose bajos valores de señal a ruido (SNR). Un caso ejemplar de esto es la baja resolución espacial obtenida por bandas como el LWIR al generar una borrosidad o blur óptico en los bordes de los objetos, al detectar su emisión, problema que el SWIR y NIR no poseen comunmente en sus aplicaciones.

La capacidad de capturar de manera multiespectral (MS) no se considera en las cámaras IR convencionales, esto debido a que los detectores engloban la banda completa en las imágenes capturadas sin permitir revisar una longitud de onda en específico. Para poder realizar una revisión de las características espectrales de una escena cualesquiera, se vuelve necesario poseer la capacidad de determinar las longitudes de onda de interés para el fenómeno observado, ya sea mediante filtros ópticos u otras tecnologías como cámaras hiperespectrales [6, 7, 11]. En el mismo ámbito de obtener más información de una escena, la captura del LF (también llamada captura plenóptica) es un problema novedoso que permite adquirir información de la dirección de la luz mediante diversas técnicas de captura, sin embargo, cada una de estas induce limitaciones que requieren ser solucionadas. Se puede decir entonces, que con el objetivo de capturar una escena existen diversas metodologías que se enfocan en obtener información espectral, dirección de la luz y además poseen diferentes resoluciones. Esto permite dividir los desafíos que la óptica posee en categorías, las que son atacadas por separado normalmente.

Actualmente, existen soluciones para cada una de estas problemáticas que atacan uno de los problemas, sin considerar los otros o utilizando métodos estándar de corrección para lidiar con ellos cuando no son el foco de la investigación. El FPN ha sido solucionado por métodos que se categorizan en calibración de cuerpo negro,

basados en escena y métodos estadísticos, cada uno de estos posee ventajas y desventajas como lo son el tiempo, costo, asumir valores, entre otros. Con estas limitaciones en mente, se vuelve complejo utilizar la imagen ya procesada para otros fines como lo es análisis MS. Además, la mayoría de estos métodos se aplican a una única cámara en específico o en el mejor de los casos, una porción del IR. Actualmente, se está prefiriendo utilizar Deep Learning e inteligencia artificial para corregir estos defectos, sin embargo, cuenta con la desventaja de necesitar una red entrenada, bases de datos extensas y un poder de procesamiento no menor [12, 13]. En cuanto a la captura del LF, existe el problema de costo del equipo utilizado. Una medición de LF requiere de un arreglo de cámaras, un arreglo de microlentes o una cámara con un desplazador para capturar no solo la intensidad de la luz sino también su dirección. En IR, un arreglo de cámara es operacionalmente muy costoso aún. Un arreglo de microlentes reduce considerablemente la resolución, además de que necesita ser construido específicamente para una cámara. La cámara con el desplazador requiere de una escena estática y del costo del desplazador mecánico. Actualmente [14, 15, 16], se están utilizando los 3 métodos dependiendo de la aplicación y espectro. Cuando la escena es estática se prioriza el desplazador. Cuando no es necesaria una resolución alta o identificar bordes como en el LWIR, se utilizan los microlentes. Si el costo no es un problema, entonces se usa el arreglo de cámaras. Independiente del método usado, aún persiste el problema de la baja resolución espacial de las imágenes IR, que es natural para la tecnología empleada, si bien una solución lógica podría ser la de hacer más pequeño el tamaño del pixel, esto tiene un límite físico que ya ha sido alcanzado por las tecnologías actuales. Hacer sensores más grandes para que tengan más píxeles es otra solución sugerida, pero cuenta con el problema de ser poco práctico tener sensores de enormes tamaños. Es por esto que se ha avanzado en el desarrollo de superresolución (SR), que consiste en aumentar la resolución de una imagen basándose en una o varias imágenes de menor resolución. Actualmente, para el uso de la técnica de SR [17, 18], se desarrollan complejos modelos matemáticos de una o varias imágenes que consideran factores como el movimiento sub-pixel, distintas Point Spread Function (PSF) y bases de datos para machine learning. Este último ha cobrado fuerza y se han desarrollado complejos y variados algoritmos, sin embargo, la necesidad de contar con redes de entrenamiento y bases de datos extensas disminuyen sus capacidades y adaptabilidad a otras bandas.

Existen trabajos [2] que solucionan los problemas de FPN y captura de LF mediante la construcción de un sistema de captura plenóptica de reenfoque, considerando el FPN como un objeto en el infinito óptico y utilizando el Fourier Slice Photography Theorem (FSPT) [19]. Este tipo de soluciones escapan de las desventajas de los métodos de eliminación de FPN existentes y proporcionan una solución elegante al trabajar con la óptica y la física del campo de luz. No obstante, no se considera un análisis MS ni una solución a la baja resolución espacial. Por tanto, aún existen áreas donde mejorar y la necesidad de construir un sistema de adquisición completo que solucione los problemas mencionados anteriormente.

Se propone, entonces, basándose en lo de mostrado en el trabajo de Coelho en [2], construir un sistema de

imágenes con la capacidad de generar cubos MS en la banda SWIR. Agregando además la capacidad de lograr superresolución (SR) de imágenes digitales dentro del cubo de datos MS IR para cada plano generado por el algoritmo de reenfoque y eliminación de FPN. Es decir, aumentar la resolución espacial más baja de la banda SWIR a los estándares más altos como lo es por ejemplo la banda NIR. Esto implicaría integrar la capacidad de eliminación de ruido usando el FSPT en el algoritmo de SR.

1.1. Hipótesis

La resolución espacial de una cámara SWIR de sensor 320x256 se puede aumentar a una resolución espacial mayor como lo es una cámara NIR de sensor 1280x1024 mediante la implementación de SR.

1.2. Objetivos específicos

- Construir un sistema SWIR de captura de campo de luz con capacidad multiespectral.
- Aumentar la baja resolución de una cámara SWIR mediante el uso de una técnica de SR modular basada en optimización (ADMM) .
- Generar cubos MS High Resolution (HR) de una escena de interés aprovechando las características de eliminación de FPN, MS y la mejora mediante SR.

2. Estado del Arte

2.1. Eliminación de FPN y otros artefactos

La corrección de FPN se puede dividir en 3 grupos, los dependientes de calibraciones de cuerpo negro (BB), los basados en escena y los basados en comportamiento estadístico de variables. El método tradicional para corregir el FPN es utilizar la denominada calibración de BB de uno o dos puntos[20]. Este método entrega imágenes con un alto grado de limpieza de FPN. Técnicas basadas en la electrónica de la cámara se han empleado. M. Kłosowski[21] propone una nueva arquitectura para el conversor análogo-digital utilizado en una cámara. Otros métodos más tradicionales han desarrollado correcciones bidimensionales [22], donde se emplea no solo cambios de temperatura sino que también en tiempo de integración, permitiendo así eliminar artefactos mediante sucesivas calibraciones. Pese a los beneficios que trae, utilizar BB implica un costo elevado tanto en obtención, mantención y tiempo. Un cuerpo negro posee rangos de operación limitados de temperatura, con valores rondando entre los 100°C hasta los 400°C para los de gran superficie y de más de 800°C los de superficie menor. Mantener una temperatura elevada en una superficie extensa es costoso y por ende se utilizan superficies menores para rangos más elevados, a lo anterior se le añade requerir un tiempo de estabilización tanto para

subir como bajar la temperatura y además requiere un ambiente controlado sin interferencias de temperaturas variables entre el sensor y el cuerpo negro como lo son viento, temperaturas variables o suciedad en general. Hay que considerar además que la cámara debe de extraerse de donde se esté utilizando para realizar la calibración, obteniéndose así calibraciones offline. Dependiendo de la tecnología de la cámara, recalibraciones pueden ser necesarias cada cierto tiempo[23], además, esto debe hacerse para cada cámara por separado en distintos rangos de temperatura del BB. Los métodos basados en escena utilizan la imagen e información previa que se tenga de la escena para eliminar el FPN. Muchos de estos métodos incluyen bases de datos y machine learning con el objetivo de entrenar redes convolucionales e inteligencia artificial para limpiar la imagen. J. Guan[24] utilizó redes convolucionales para separar el ruido de una escena con buenos resultados, comparando después con métodos ya existentes. S. Kim[25] incorporó un circuito de calibración y redujo el FPN utilizando píxeles de referencia y píxeles activos de las escenas específicas. Se utilizan también, por otra parte, algoritmos de procesamiento de imágenes que logran eliminar el ruido. Los métodos basados en el comportamiento estadístico de las variables se enfocan en determinar las características estadísticas de cada sensor. El histogram matching [26] se basa en utilizar el histograma de una imagen de referencia que tenga radiación incidente constante en cada sensor. De esta manera se puede ajustar y hacer calzar el histograma de referencia con el de una imagen. Sin embargo, asegurar una radiación incidente constante en los sensores es extremadamente difícil y las aproximaciones utilizadas entregan resultados deficientes en muchos casos. Moment matching es otro método que utiliza la varianza y el promedio de los píxeles en una sucesión de frames, para luego compararlo con una referencia y así eliminar el ruido. I. Kudinov[27] propuso utilizar videos en IR para realizar una eliminación de FPN más efectiva en comparación a una sola imagen. S. Amalan[28] utiliza una versión modificada de sustracción de frame oscuro en sensores CMOS. Sin embargo, los resultados son dependientes de la escena, en particular del detalle de esta. Entre mayor detalle tenga la escena, peor será la eliminación de ruido. Generalmente, los algoritmos empleados son probados para una sola banda, ya sea SWIR, MWIR o LWIR. Además, existen casos donde la solución aplica para una estructura específica del FPN y presenta problemas para aplicarse en otros espectros [29]. Por lo que no se puede asumir que una técnica de eliminación de ruido funcione para todo el espectro. El enfoque de preferir algoritmos tiene la ventaja de poder comprobarse su funcionalidad en otros espectros sin costos altos, siendo así más versátil y menos costoso que incorporar nuevo hardware o crear sensores propios para cada nuevo espectro, con el fin de probar la factibilidad de una técnica en específica de eliminación de FPN. Entre las obras más citadas, Narendra, Harris et al [30, 31, 32] propone un algoritmo NUC que estima continuamente el parámetro para el NU siguiendo un supuesto estadístico constante. Ratliff et al[33, 34, 35] proponen un algoritmo algebraico basado en estimaciones de cambios de sub píxeles en una secuencia de imágenes, extrayendo el parámetro NU usando un modelo de interpolación lineal. Torres y Hayat[36] introdujeron una técnica de filtrado de Kalman que captura la deriva temporal estocástica de cada detector. El grupo de la Universidad de

Concepción propuso en [2, 37] un nuevo enfoque para reducir FPN basado en la escena capturada IR LF y reenfoándolo digitalmente a cualquier objeto de escena cercano de interés. En dicho trabajo, inicialmente se probó esta nueva idea de filtrar FPN mediante un reenfoque digital. Logrando resultados iniciales muy prometedores. Si bien el FPN es el artefacto de mayor impacto en los sensores IR, no es el único presente en una medición. El blur es otro artefacto importante que consiste en la visualización de imágenes de manera borrosa. Esto puede ser producto del fuerte procesamiento de imágenes aplicado que puede distorsionar una imagen, o bien debido a los factores ambientales como agua[38]. Para solucionar esto, en la actualidad las redes convolucionales y Deep Learning son utilizadas. Tratar con datos reales y en gran cantidad ha permitido utilizar como referencia para estas redes blur real y no sintético, entrenando así modelos más eficientes [12, 13]. Otro acercamiento ha sido descomponer imágenes de largo tiempo de exposición en intervalos más cortos [39]. Donde al tener tiempos de captura reducidos, se puede extraer imágenes con una mayor nitidez del objeto de interés. Como es lógico, escenas donde se aprecie movimiento es susceptible a presentar blur. Luego, las soluciones empleadas se enfocan en escenas donde exista variaciones de la posición de los objetos en la escena. Sin embargo, si se incluye el reenfoque e imágenes que naturalmente son borrosas, la eliminación del blur abarca un área mucho más vasta y significativa en la obtención de resultados de diversos métodos.

2.2. Lightfield y plenóptica

En cuanto a la plenóptica y la captura del campo de luz, diversas han sido las aplicaciones y mejoras del método. La capacidad de reenfocar una imagen ha sido empleado y optimizado para cámaras submarinas, como lo demuestra G. Jiang[40] en su revisión de factores que afectan la captura submarina del campo de luz. La estimación de profundidad ha sido revisada por investigadores como M. Labussière[14] desarrollando métricas propias. Mejoras a la detección de profundidad se han hecho basados en otras técnicas de captura modernas. Un ejemplo de esto corresponde a la inclusión de técnicas de realidad virtual para mejorar la velocidad y capacidad de reconstrucción de imágenes mediante captura del campo de luz[16]. Aplicaciones menos convencionales incluyen su uso para determinar la trayectoria en el espacio de elementos, donde un ejemplo actual corresponde a la investigación de R. Fischer[41], donde se identifican las partículas metálicas que salpican producto de la aplicación de un láser.

Los métodos de captura de LF pueden resumirse en tres principales ramas, cada una con sus ventajas y desventajas resumidas en la tabla 1.

Donde dependiendo de la aplicación y recursos disponibles, se prioriza un método sobre otro. Cada uno de estos métodos genera conjuntos de imágenes en forma matricial, donde el análisis del campo de luz recopilado por esta matriz permite generar algoritmos que aprovechen la información proporcionada.

La información obtenida ha sido utilizada para reconstrucción 3D en distintos campos de la ciencia. Al-

Método	Costo	Resolución	Tiempo
Arreglo de Cámaras	Alto	Alta	Instantáneo
Cámara + Microlentes	Bajo	Baja	Instantáneo
Cámara + Desplazador	Bajo	Alta	Minutos

Cuadro 1: Comparación de métodos de captura plenóptica. Cada método posee ventajas y desventajas a considerar a la hora de implementar el sistema, convirtiéndolo en dependiente de la aplicación.

gunos de estos incluyen reconstrucción de rostros[42], identificar objetos con volumen en una escena con una sola cámara plenóptica[43] e incluso se han visto aplicaciones en procesos naturales como la evolución de superestructuras de hielo[44]. Se han logrado obtener resultados en otras aplicaciones como lo es poder realizar predicción de objetos fuera del campo de visión con la información del campo de luz[45]. Una extensión del enfoque ha sido generar imágenes *all in focus* donde combinando los distintos planos se logra obtener imágenes con todos los elementos enfocados en la misma captura[46].

La industria en general también ha visto aplicaciones. Se ha logrado reconstruir llamas irregulares mediante captura plenóptica[47]. Logrando identificar fases de fundido y temperaturas. Debido a la maquinaria pesada utilizada en muchos de estos procesos, el correcto funcionamiento de esta ha sido revisado detectando partículas de combustibles reutilizados en la industria[48]. La industria de los semiconductores también ha empleado la captura plenóptica en la revisión de placas y componentes[49]. Si bien tiende a asociarse a la plenóptica con una cámara grande, en el mundo de la microscopía también han existido aplicaciones de esta tecnología. Se han generado nuevos protocolos y alternativas manteniendo la resolución espacial[50], factor que se degrada considerablemente a la hora de emplear esta técnica. En la medicina ha sido posible emplear la captura de campo de luz como herramienta en detección de irregularidades[51], mediante el uso de una sola captura de rayos X.

Muchos de estos estudios se basan principalmente en cámaras del espectro visible, el IR no ha sido revisado en profundidad en la captura del campo de luz. Entre las razones están el elevado costo de la tecnología y la dificultad de obtener una. Dentro de los pocos que ha realizado captura de campo de luz en IR cabe destacar a Coelho[2], que ha sido de los pocos investigadores que ha desarrollado un sistema plenóptico en el IR mediante un desplazador.

El procesamiento del campo de luz es un proceso computacionalmente pesado. Esto implica que para aplicaciones donde se necesite obtener la información instantáneamente, se necesita contar con el hardware lo suficientemente capaz. Por este motivo, a fin de convertirlo en una herramienta más versátil, se ha trabajado en optimizar los tiempos de cálculo, almacenamiento de imágenes y trata de datos. Algunos métodos se han enfocado en como desarrollar un *pipeline* del procesamiento de datos de manera de acortar tiempos sin perder calidad en la imagen[52]. El desarrollo de redes neuronales ha permitido generar algoritmos de reconstrucción

veloces que operan eficientemente y con mejor compresión de imágenes[53]. La construcción de hardware especializado como lo son las FPGA ha mejorado considerablemente la capacidad de cómputo y poder del análisis del campo de luz[54].

2.3. Superresolución

La SR es un proceso para obtener una o más imágenes de alta resolución a partir de una o más observaciones de baja resolución (LR)[55]. Se ha utilizado para variadas aplicaciones diferentes, como imágenes aéreas y satelitales, procesamiento de imágenes médicas, imágenes por ultrasonido [56], imágenes infrarrojas, mejora de imágenes faciales, imágenes de texto[57] y mejora de imágenes huellas[58]. El SR es un algoritmo que tiene como objetivo proporcionar detalles más finos que el muestreo de un dispositivo de imágenes determinado al aumentar el número de píxeles por unidad de área en una imagen. Comúnmente, los algoritmos de SR se dividen en distintas categorías de acuerdo a diversos factores. Los dos grupos principales son los métodos basados en una imagen (SI) y los basados en múltiples imágenes(MI), cada uno de estos se divide en otras categorías dependiendo del método empleado. Los métodos SI, como su nombre lo indica, requieren de una sola captura de la escena y trabajan basándose en esta. Inicialmente, solucionar el problema con una sola imagen utilizaba suavizado de bordes, deconvolución, basado en ejemplos, entre otros[59]. Se conseguía tener un mejoramiento de imagen, pero se sufría de problemas en exactitud y fidelidad de la información. Autores como K. Kim[60]propusieron utilizar métodos basados en ejemplos (también llamado alucinación de imagen) repitiendo varias veces su aplicación a distintos parches de una imagen. Sin embargo, como el nombre indica, la nueva información era alucinada por el algoritmo, por tanto, no era fiel a los detalles reales de la imagen. Los métodos basados en mejoramiento de bordes lograban combatir los bordes toscos presentes en imágenes. Sin embargo, sufren de problemas en imágenes con muchos detalles al suavizar gran parte de estos[61]. Rápidamente y a medida que las capacidades computacionales lo fueron permitiendo, se pasó a utilizar machine learning y convolutional networks (CNN). SRCNN [62] fue el primer modelo basado en CNN. En el método, la interpolación bicúbica se utiliza como preprocesamiento y la convolución para extraer los parches de información. Otros modelos como Deep Projection CNN[63] utilizan las bases de datos para detectar estructuras repetitivas en imágenes LR. Recientemente, las generative adversarial networks (GAN) [64] han surgido como una alternativa a los modelos tradicionales de mejoramiento de imagen. Los cuales utilizan una serie de capas, discriminadores y combina la percepción de SNR con la visual. Si bien los CNN y otros métodos de machine learning han entregado resultados satisfactorios, estos no están exentos de problemas y complejidades. La principal adversidad corresponde a las enormes bases de datos que se requieren para entrenar una red, esto también implica una capacidad de procesamiento de información alta. A esto se le añade la complejidad de algunas de estas redes, donde por ejemplo GAN ha sido criticada debido a su complejidad de implementación y de estruc-

tura. Por otra parte, los métodos basados en múltiples imágenes surgen como una alternativa. un método basado en múltiples imágenes hace referencia a obtener información de la escena mediante el movimiento sub-píxel de esta. Es decir, obtener distintos puntos de vista y construir, basándose en la información al desplazar la escena, una imagen HR. El área de aplicación donde más se ha reconocido su aporte ha sido en el ambiente aeroespacial. Ejemplos de esto corresponden al satélite francés SPOT-5 donde se instalaron 2 arreglos de sensores CCD de 12000 píxeles desfasados en medio píxel[65]. Las cámaras aéreas Leica ADS40 también han incluido una estructura similar[66]. Esto ha permitido obtener imágenes de alta resolución en ambientes extremos donde la capacidad computacional muchas veces está limitada. Sin embargo, tanto la implementación como los resultados al tener menos tiempo siendo investigados han quedado algo rezagados en comparación a otros métodos de SR.

3. Marco Teórico

3.1. El Infrarrojo

3.1.1 Bandas espectrales: NIR, SWIR, MWIR y LWIR

En el espectro electromagnético se puede encontrar entre los 700 *nm* y 1 *mm* la banda infrarroja, conocida popularmente como la banda de detección de temperatura humana y detección de personas [11, 67]. Sin embargo, estas aplicaciones corresponden a una subdivisión del IR solamente y las aplicaciones reales de esta tecnología van más allá de solamente detección de temperatura.

Ejemplos de escenas en las cuales el IR ha aportado de manera sustancial incluyen, reconocimiento facial [68], agricultura e industria de alimentos [69], entre otros. La figura 3.1 provee una visión organizada de como se estructura el IR de acuerdo a sus subdivisiones.



Figura 3.1: Desglose del espectro IR, las secciones corresponden a NIR(*Near Infrared*), SWIR(*Short Wave Infrared*), MWIR(*Mid Wave Infrared*) y LWIR(*Long Wave Infrared*) [1].

La banda NIR se utiliza comúnmente en aplicaciones de agricultura y control de alimentos[70]. El SWIR se

puede aplicar al análisis del estrés hídrico de plantas o detección de metales preciosos, entre otros[71]. el MWIR se ha usado en identificación de materiales de pintura y reconocimiento facial[11]. Finalmente, el LWIR se ha empleado en rescate de personas por calor corporal en catástrofes naturales[72]. Una diferencia importante en el funcionamiento de cámaras IR es el tipo de radiación que captan. Por un lado, el NIR y SWIR captan reflexiones en una escena, mientras que la tecnología sobre la base de MWIR y LWIR detecta emisividad de los objetos.

Estas cámaras captan una sección larga de longitudes de onda, por lo que capturan una vasta cantidad de información que puede ser difícil a veces de aislar, es por esto que para un análisis más detallado, se requiere de una revisión de las longitudes de onda por separado ya sea mediante filtros ópticos u otras técnicas.

3.1.2 El cubo Multiespectral

Al momento de separar longitudes de onda, se genera el denominado análisis multiespectral, que consiste en la captura de una escena utilizando distintas longitudes de onda, esto generalmente se logra mediante el uso de filtros ópticos o bien haciendo uso de una cámara hiperespectral obteniéndose de este modo una tercera dimensión que consiste valores de longitud de onda. La inclusión de una tercera variable genera de esta manera un cubo multiespectral o hiperespectral dependiendo del método empleado. Se utilizará el término cubo multiespectral en esta investigación, considerando que el *modus operandi* fue utilizar filtros ópticos y reemplazarlos de manera manual. La figura 3.2 compara como se separa comúnmente cubo multiespectral e hiperespectral.

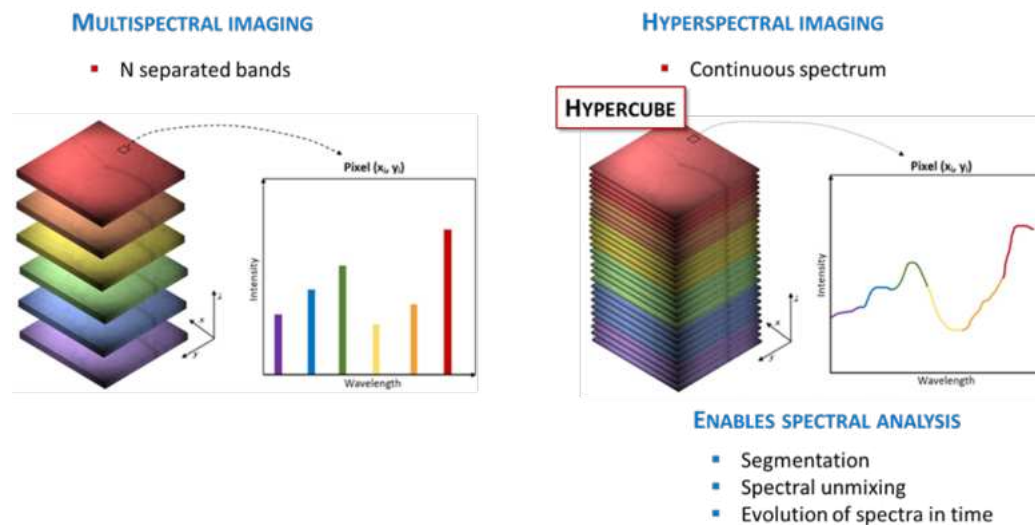


Figura 3.2: Cubo multiespectral e Hiperespectral, cortesía de NIREO.

El muestreo multiespectral se ha utilizado en distintas secciones del espectro y ha demostrado ser una herramienta eficiente de obtención de información del comportamiento de un objeto de interés, como lo son control de calidad de alimentos [73] y criminología [74]. Es justamente esta versatilidad de escenarios y de espectros donde puede utilizarse lo que la posiciona como una herramienta ideal para la revisión de fenómenos

ópticos, permitiendo construir un histograma completo e identificar la frecuencia de interés donde se distinga claramente el fenómeno buscado.

3.1.3 El FPN en el IR

la óptica IR posee una serie de problemas asociados a la tecnología empleada, principalmente el FPN, que consiste en la degradación de la imagen obtenida mediante un ruido inherente del sensor utilizado. El principal problema de las mediciones con cámaras IR y revisión de las imágenes obtenidas es el ruido de una imagen, esta interferencia es producida por diversos factores que corresponden a los componentes electrónicos de la cámara y a las condiciones bajo las cuales opera. Esto puede generar un desgaste de la imagen substancial, por lo que es necesario preocuparse de como eliminarlo y de identificar correctamente cuál es el tipo de ruido presente. Existen dos tipos de ruido principalmente, los de no uniformidad (NU) y los temporales. El primero ocurre debido a la configuración que tienen los sensores en las cámaras IR el cual se denomina *Infrared Focal Plane Array* (IRFPA) donde al ser múltiples detectores dentro del sensor y además al ser utilizada tecnología extremadamente sensible a la radiación IR se produce el fenómeno de diferentes valores detectados entre cada sensor pese a ver una imagen uniforme. Los temporales, por otra parte, son los que tienen relación con el ruido de disparo y el *flicker noise*. El FPN entra en la categoría de no-uniformidad y se caracteriza en que es estático al ser independiente del tiempo, es por esto que recibe la categoría de patrón fijo.

Dentro de la categoría de no-uniformidad (NU) se tienen dos parámetros, los cuales son los que alteran la imagen, el offset y la ganancia de cada detector. El primero corresponde a un valor independiente de la radiación incidente y variable con la temperatura del sensor, mientras que el segundo tiene relación con la diferencia en la conversión de radiación recibida a señal eléctrica. Ambos son diferentes para cada detector del arreglo y por ende para cada pixel, provocando así el FPN. Comúnmente se define el modelo de FPN en imágenes como se muestra en la ecuación 1.

$$X'_i = X_i G_i + O_i \quad (1)$$

Donde para cada pixel i , X' es la escena a ser capturada, X es la imagen capturada por la cámara, G es la ganancia y O es el offset. Notar como cada uno de estos parámetros es único para cada pixel de la imagen.

La sensibilidad del SWIR lo hace susceptible al FPN y entrega imágenes como la de la figura 3.3, donde se aprecian líneas verticales que muestran píxeles con una receptividad distinta y potenciada. Los sensores utilizados ya sean de microbolómetros o de ferroeléctrico basándose en fotodetectores en los distintos espectros en IR fuera del SWIR presentan el ruido de distinta manera, pero aun así la procedencia de estos sigue siendo el hardware utilizado y sus defectos e imperfecciones visibles y diferentes para cada pixel.



Figura 3.3: FPN presente en una cámara Xeva 320 SWIR. Notar como líneas verticales de intensidad aumentada plagan la imagen. Cada sensor y banda poseen un ruido característico distinto a los demás, por lo que no existe una solución única a este fenómeno.

3.2. Lightfield y Fourier Slice Photography Theorem (FSPT)

3.2.1 *Lightfield*

El *lightfield* se puede definir mediante la función plenóptica [10] como el conjunto de rayos de luz que viaja en todas las direcciones a través de cada punto en el espacio 3D desde una perspectiva óptica geométrica. Como es una función que describe un concepto tan general, esta se compone de numerosas variables, dentro de las cuales están las coordenadas (x, y, z) , los ángulos posibles (θ, ϕ) , la longitud de onda (λ) y el tiempo (t) .

$$L(x, y, z, \theta, \phi, \lambda, t) \quad (2)$$

Esta función, si bien posee una descripción completa de una escena, es a su vez difícil de operar y en muchas ocasiones posee variables redundantes que no son necesarias considerar en mediciones donde se cumplan las condiciones. Por ende, se desarrollaron una serie de simplificaciones a este modelo bajo las cuales se pudo obtener una función de carácter sencillo en comparación a la función original. La representación usada actualmente corresponde a una desarrollada por M. Levoy [9] quien lo redefinió como la intersección de 2 planos. Esto con el objetivo de generar un renderizado de imagen de punto de vista variable a diferencia de los métodos preexistentes, Levoy denominó a este constructo como *light slab*.

Posteriormente, el trabajo de Ren Ng[19] define una configuración de los planos tal que se considere $L(x, y, u, v)$, donde un plano se ubica en el sensor y el otro en el lente de la cámara, provocando así que la función describa el interior de la cámara. Esta representación se muestra en la figura 3.4 y es la base para su investigación y creación del FSPT, el cual se usará en el trabajo presente.

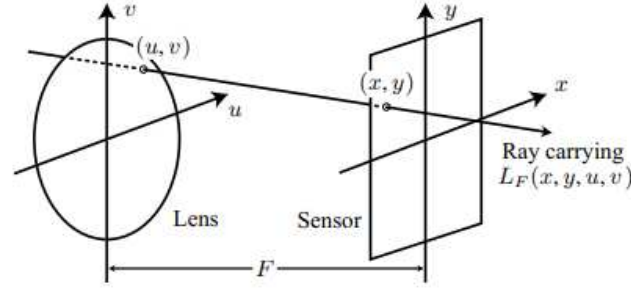


Figura 3.4: Parametrización de Ren Ng del *light slab*. Se define un lente y un sensor y la intersección de una línea entre estos planos define un rayo de luz.

3.2.2 FSPT

Las parametrizaciones realizadas por Ren Ng tienen como objetivo llegar al FSPT, para esto primero se define la generación de imágenes basándose en el campo de luz de las cámaras convencionales y las plenópticas donde las primeras se definen de acuerdo a:

$$E_f(x, y) = \frac{1}{F^2} \int \int L_F(x, y, u, v) du dv \quad (3)$$

Donde E_f corresponde a la irradiancia en (x, y) y F es la separación lente-sensor. La dependencia angular de la fórmula se puede ignorar al ser absorbida por la ecuación del *lightfield*.

Para una cámara plenóptica la captura consta de dos partes: medición y procesamiento. Este tipo de cámara utiliza un sensor plano que provee muestreo direccional de la irradiancia pasando por cada sensor, por lo que si el sensor está a una profundidad F del lente, entonces el muestreo será de L_F . El procesamiento de imagen permite utilizar el campo de luz descrito de la cámara convencional para computar la imagen obtenida a cualquier profundidad E'_F , donde F' no tiene que ser igual a F . Para lograr esto se re-parametriza L_F para obtener L'_F de acuerdo a la figura 3.5 y se aplica la ecuación 4

$$\begin{aligned} L'_F(x, y, u, v) &= L_{(\alpha F)}(x, y, u, v) \\ &= L_F(u + (x - u)/\alpha, v + (y - v)/\alpha, u, v) \\ &= L_F(u(1 - 1/\alpha) + x/\alpha, v(1 - 1/\alpha) + y/\alpha, u, v) \end{aligned} \quad (4)$$

De esta manera se puede definir un operador de fotografía, el cual se denomina $\mathcal{P}_\alpha$, que transforma un campo de luz a una profundidad F en una fotografía formada a una profundidad (αF) . Considerando que $\mathcal{P}_\alpha[L_F]$ representa la aplicación de $\mathcal{P}_\alpha$ a un campo de luz L_F , entonces en la ecuación 5 se tiene:

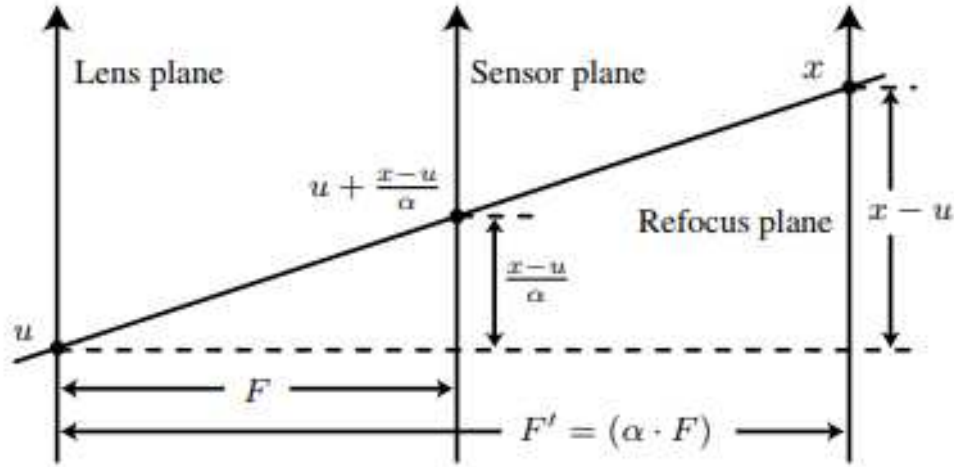


Figura 3.5: Re-parametrización del *lightfield* moviendo el sensor desde F a F'

$$\begin{aligned}\mathcal{P}_\alpha[L_F](x,y) &= E_{\alpha F}(x,y) \\ &= 1/\alpha^2 F^2 \int \int L_F(u(1-1/\alpha) + x/\alpha, v(1-1/\alpha) + y/\alpha, u, v) du, dv\end{aligned}\quad (5)$$

Esta operación se puede apreciar como cortar un espacio 4D y luego proyectarlo en un plano 2D y permite a su vez calcular fotografías a diferentes profundidades con la medición de un solo campo de luz, convirtiéndolo así en la base de la fotografía plenóptica. Una vez ya definido el operador de fotografía plenóptica, se lleva al plano de Fourier con el objetivo de llegar al FSPT definido en la ecuación 6.

$$\mathcal{P}_\alpha = \mathcal{F}^{-2} \circ \mathcal{B}_\alpha^{-T} \circ \mathcal{F}^4 \quad (6)$$

La ecuación 6 nos dice que una fotografía es la transformada 2-dimensional inversa de Fourier, de un corte 2-dimensional de la transformada 4-dimensional de Fourier del *light field*. De esta manera podemos obtener el reenfoque de una escena mediante el uso del plano de Fourier, otorgándonos una mejora en procesamiento de imagen y optimizando los tiempos de cómputo.

3.2.3 Eliminación de ruido mediante plenóptica

La plenóptica es la captura del LF mediante la toma de imágenes en forma matricial en los ejes X e Y, donde cada imagen de la matriz lleva parte de la información óptica del escenario y se puede procesar para obtener distintos resultados. En la figura 3.6 se muestra visualmente como es el equipo de medición plenóptica del método elegido en este trabajo.

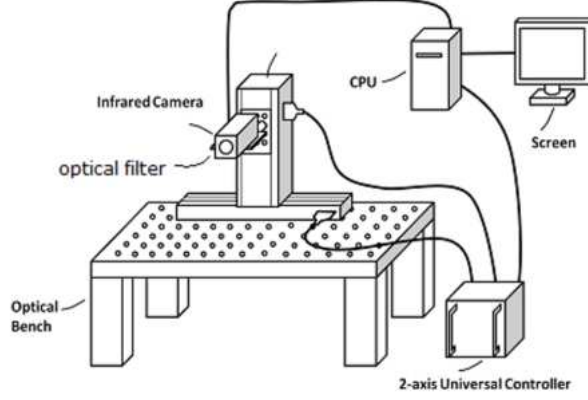


Figura 3.6: Equipo de medición plenóptica consistente en una cámara y un desplazador mecánico, los cuales son utilizados para recopilar el LF en forma matricial del tamaño deseado.

Para este trabajo nos vamos a enfocar en el reenfoque digital, que permite generar planos focales donde se puede enfocar a cada objeto de la escena por separado dependiendo de su profundidad, utilizando el LF para construir dichos planos focales. Lo anterior implica que el resultado de una medición plenóptica, en nuestro caso, es un arreglo de imágenes donde cada una posee un distinto plano focal, además, a cada uno de estos arreglos le añadiremos un filtro óptico, sumándole así una cuarta dimensión al cubo. Como punto de partida generaremos nuestro propio modelo de medición plenóptica en 7.

$$LF(x, y, M, N, \lambda) \quad (7)$$

Donde x, y representan la resolución de la imagen. M, N la dimensión en el eje X e Y respectivamente de la matriz de imágenes y λ es la longitud de onda particular seleccionada del cubo multiespectral utilizando distintos filtros ópticos. Aplicando el FSPT descrito en 6 al modelo de medición plenóptica recién descrito en 7, se obtiene un stack de imágenes con distinto plano focal, obteniéndose así la ecuación 8.

$$\mathcal{P}_\alpha[LF(x, y, M, N, \lambda)] = E_\alpha(x, y, \lambda), \quad 0 < \alpha < n \quad (8)$$

Con E_α representando el stack de imágenes obtenido con un tamaño definido por n donde se puede elegir cada plano focal de acuerdo al valor de α . Entre más cercano sea a 0 el valor de α , más cercano estará el plano focal al infinito óptico, mientras que entre mayor sea su valor, más cercano se estará del sensor. Con lo anterior, se obtiene un cubo multiespectral de 4 dimensiones expresado en la ecuación 8, con (x, y) siendo las 2 dimensiones espaciales básicas de una imagen, λ la tercera dimensión espectral y α la cuarta dimensión de profundidad espacial. La figura 3.7 expresa el análisis de las dimensiones posibles del resultado.

Utilizando el teorema de FSPT, Coelho [2] desarrolló una teoría sobre como el reenfoque digital podía

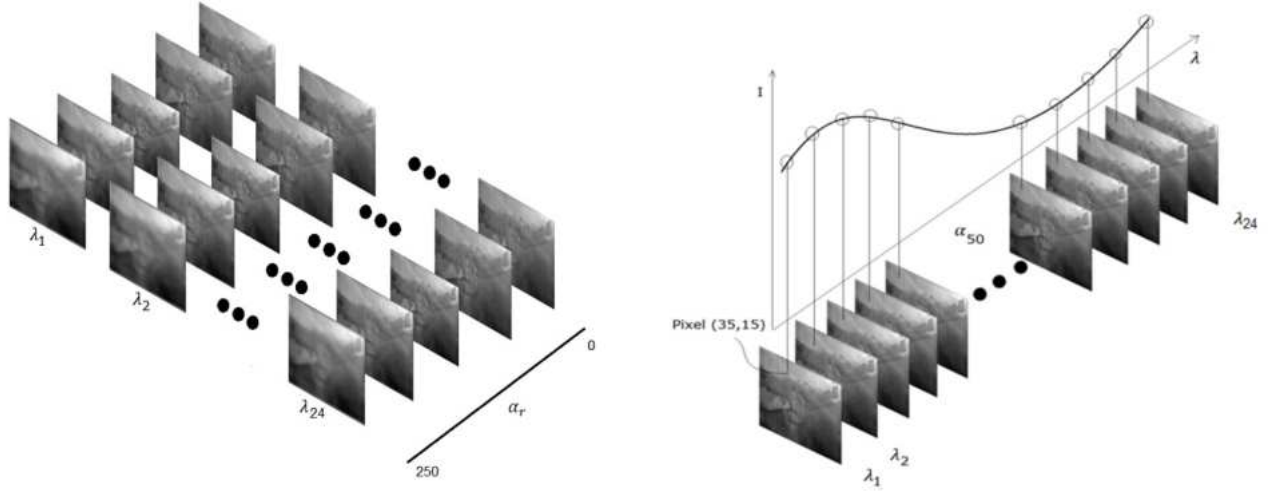


Figura 3.7: Cubo multiespectral de 4 dimensiones generado por el método descrito. Notar como se puede desglosar el cubo de acuerdo a su longitud de onda λ o su profundidad α y revisar los píxeles específicos de la imagen (x, y) .

combatir el FPN presente en el IR, para esto, definió el siguiente modelo de imagen detallado en la figura 3.8.

Ubicando el FPN en el infinito óptico ($\alpha = 0$) en vez de al principio ($\alpha = n$), $n \in \mathbb{N}$, como lo hace el modelo clásico, se puede aplicar el reenfoque mediante plenóptica para eliminar el ruido y enfocar los objetos en la escena. Los resultados obtenidos muestran imágenes sin ruido y con distintos planos focales, donde a medida que se acerca el plano focal al sensor, es decir, α aumenta, el FPN va desapareciendo. La figura 3.9 muestra como opera el algoritmo de reenfoque digital para una escena con 2 objetos con $0 < \alpha < 250$.

Como se aprecia en la figura 3.9, a medida que el valor de α va en aumento, el plano focal se va acercando y el ruido desaparece. Entre más cerca se esté de $\alpha = 0$, más prevalente será el ruido y se estará más cerca del infinito óptico. Por otra parte, para valores de α mayores, los objetos se van enfocando y el FPN va desapareciendo.

3.3. SR en el IR

La resolución de la imagen procesada es afectada enormemente por los sensores y la banda elegida, siendo el límite de difracción el que determina la resolución tanto espacial como angular de un sistema de imagen y por ende el tamaño de pixel utilizado. La resolución espacial hace referencia a la distancia física mínima entre dos puntos que un sistema de imagen puede determinar, está definida por la ecuación 9. La resolución angular, que hace referencia al ángulo de separación mínimo entre dos puntos que el sistema de imagen captura, está definida en la ecuación 10

$$SpatialResolution = \frac{\lambda}{2NA} \quad (9)$$

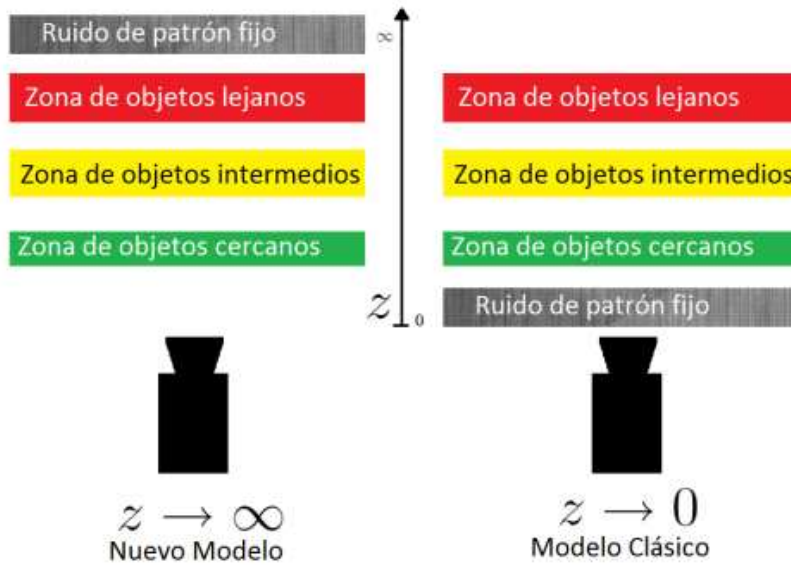


Figura 3.8: Modelo de ruido ubicado en el infinito óptico desarrollado por Coelho. El cambio corresponde a uno de concepto que permite mirar desde otro punto de vista la problemática [2].



Figura 3.9: Aplicación de la eliminación de ruido mediante reenfoque propuesta por Coelho. El valor de α_n denota el plano de enfoque actual, A) $\alpha = 0$, B) $\alpha = 80$, C) $\alpha = 215$.

$$\theta = 1,22 \frac{\lambda}{D} \quad (10)$$

donde λ es la longitud de onda, NA es la apertura numérica del sistema y D es el diámetro de la apertura. Para mayores valores de λ el límite de difracción de ambas aumenta, implicando que a mayor longitud de onda el tamaño de pixel debe de incrementar proporcionalmente para una captura correcta y satisfacer el criterio de muestreo de Nyquist. Un mayor tamaño de pixel significa que la imagen tendrá menor resolución y detalle, pero será más sensible a la radiación emitida por cuerpos, como para LWIR, donde se captura la emisividad de los objetos.

Bandas como SWIR poseen una mayor resolución en comparación a LWIR por lo general, de forma similar,

el espectro visible posee una mayor resolución espacial en comparación al SWIR. La figura 3.10 muestra la comparación entre las resoluciones del espectro visible y el SWIR.

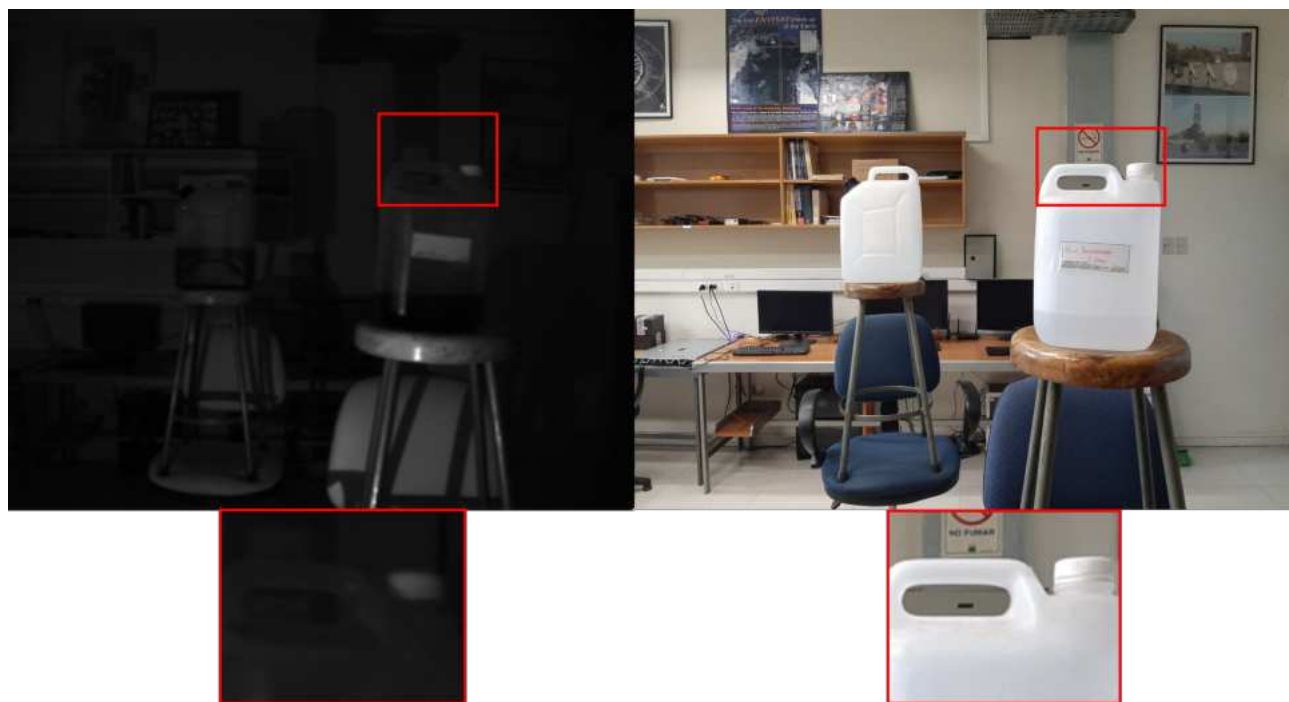


Figura 3.10: Diferencias en resolución entre espectro SWIR (a la izquierda) y espectro visible (a la derecha). Al enfocarse en detalles como el asa del bidón, se pueden apreciar como los bordes y resolución del objeto están más definidos en el espectro visible.

Lo anterior genera una serie de diferencias en la calidad de las imágenes obtenidas por cada sistema de captura para cada banda, donde tanto la resolución de la imagen obtenida como la definición de los objetos en la escena varían enormemente. Es por esto que se desarrolló el concepto de SR, que involucra aumentar la resolución de una imagen determinada mediante métodos de procesamiento de imágenes. El enfoque que se le dio en este trabajo fue el de un método basado en optimización, es decir, se utiliza un modelo consistente en una función de costo y regularizadores, asegurando así una fidelidad de datos alta, minimizando su función de costo mediante sucesivas iteraciones. Un modelo de optimización básico se estructura normalmente de acuerdo a la ecuación 11.

$$\arg \min_X \|S H X - Y\|^2 \quad (11)$$

Considerando una función de costo $\|S H X - Y\|$ que se compone de un downsampling S , un blur H , una imagen de alta resolución X y la imagen de baja resolución Y . La función anterior busca minimizar las diferencias entre la imagen baja resolución y la imagen alta resolución utilizando iteraciones para disminuir las diferencia entre ambas, es decir, minimizar la función de costo alterando el valor de X . Si se le agregara un regularizador, este sería de la forma de suma de un término $reg[X]$, el cual se encarga de imponer restricciones siguiendo un parámetro reg , con el objetivo de priorizar imágenes sin ruido, con mayor detalle o de bordes suaves, según el usuario requiera. Mediante este proceso se puede obtener una imagen de alta fidelidad siguiendo las iteraciones, alterando el valor de X y eligiendo el regularizador adecuado.

El método de optimización utilizado en el presente trabajo es el desarrollado por H. Chan [75] que utiliza un algoritmo Plug and Play ADMM (ADMM) para restaurar la calidad de imágenes, donde resuelve el problema de optimización aprovechando la naturaleza modular de ADMM obteniendo así SR.

ADMM es un algoritmo de uso común en recuperación de imagen cuya ventaja es la naturaleza modular que provee, es decir, separa el problema en múltiples ecuaciones donde los distintos problemas pueden resolverse por separado en cada ecuación. Esto permite simplificar el procesamiento de la imagen y atacar específicamente cada problema, para nuestro caso, nos interesa la primera de estas ecuaciones, que involucra el modelo de observación y podemos utilizar para generar SR mediante optimizaciones. La ecuación 12 es nuestro punto de partida para SR ADMM.

$$x^{k+1} = \arg \min_X f(x) + \frac{\rho}{2} \|x - \tilde{x}^k\|^2 \quad (12)$$

La ecuación 12 utiliza un modelo de optimización general en $f(x)$, como el que se detalla en la ecuación 11. Con k representando la cantidad de iteraciones, $\frac{\rho}{2}$ la variable del regularizador de penalización y x la imagen. Posterior a esto y para lograr una solución cerrada se reordena la ecuación y se convierte al plano de Fourier,

el detalle matemático puede verse en [75] donde aplica identidades como la de Sherman-Morrison-Woodbury y Polyphase Implementation, el resultado de esto es la ecuación 13.

$$x = \rho^{-1}b - \rho^{-1}G^T(\mathcal{F}^{-1}(\frac{\mathcal{F}(Gb)}{|\mathcal{F}(h_0)|^2 + \rho})) \quad (13)$$

La explicación de cada término puede revisarse en [75], donde los términos nuevos son la combinación de las variables de 12, lo relevante es incorporar nuestro modelo de observación en la ecuación recién descrita. Sin embargo, se vuelve necesario redefinir nuestro modelo para que pueda utilizarse en la ecuación 13 esto se detalla en el siguiente capítulo y junta las componentes de eliminación de ruido con SR.

4. Modelo propuesto y Setup

4.1. Modelo propuesto

Antes de implementar nuestro modelo al modelo de SR ADMM descrita en la ecuación 13, debemos partir por establecer un modelo descrito por la ecuación 7 para el caso que se trabajará, por lo que generamos la ecuación 14.

$$LF(x, y, M, N, \lambda) = CDB_\lambda X_\lambda + R_\lambda \quad (14)$$

Definiendo C como el movimiento sub pixel, D la decimación de la imagen, B el blur, X como la escena y R como el ruido presente, al aplicar el FSPT descrito en la ecuación 8, el modelo cambia al descrito en la ecuación 15, donde el stack de imágenes $E_\alpha(x, y, \lambda)$ (para un plano focal α) se define a los artefactos presentes en la escena, con α al plano focal elegido.

$$E_\alpha(x, y, \lambda) = DB_{\alpha, \lambda} X_{\alpha, \lambda} \quad (15)$$

Notar que las componentes B, X y R son dependientes de la longitud de onda λ y del plano de enfoque α , además, los componentes R y C desaparecen debido a que el stack de imágenes ya no posee movimiento sub pixel y tampoco ruido si acortamos el rango de α para que incluya planos focales sin FPN. La dependencia del blur a la longitud de onda viene de la Point Spread Function (PSF) utilizada en el modelo, para nuestro modelo la PSF se define mediante los coeficientes de Zernike[76], polinomios que permiten modelar la PSF de acuerdo a la longitud de onda de la imagen, por lo que se obtiene una PSF adaptativa de acuerdo a la longitud de onda empleada, lo que es útil cuando se trabaja con cubos multiespectrales. Con esta ecuación tenemos un punto de partida nuevo para utilizar SR en las imágenes generadas por un LF, revisando las principales variables que

afectan a un sistema de imágenes. El primer paso es describir la solución de la ecuación 15 como un problema de optimización, de la forma descrita en la ecuación 11, tal como se muestra en la ecuación 16.

$$\arg \min_X \|DB_{\alpha,\lambda}X_{\alpha,\lambda} - E_{\alpha}(x,y,\lambda)\|^2 = 0 \quad (16)$$

Ahora que hemos incorporado la imagen enfocada en el plano α a una longitud de onda λ a una ecuación de optimización, resolvemos la optimización como se describe en la ecuación 12. Esto genera la ecuación en 17, donde se introduce el plano focal elegido a una longitud de onda específica como la imagen que se desea optimizar bajo el esquema de SR propuesto por el modulo de ADMM encargado.

$$X^{k+1} = \arg \min_X \|DB_{\alpha,\lambda}X_{\alpha,\lambda} - E_{\alpha}(x,y,\lambda)\| + \frac{\rho}{2} \|X - \tilde{X}^k\|^2 \quad (17)$$

Finalmente, la resolución de la ecuación 17 se detalla en la ecuación descrita en 18, con el reemplazo correspondiente de las variables $b = DB_{\alpha,\lambda}E_{\alpha}(x,y,\lambda) + \rho\tilde{X}$ y de $G = DB_{\alpha,\lambda}$, de acuerdo al proceso detallado en [75]. Por temas de espacio y comprensión matemática, se optó por mantener la misma forma de la ecuación 13 pero con las variables reemplazadas con los valores de la ecuación 17 que ya fueron detallados.

$$x = \rho^{-1}b - \rho^{-1}G^T(\mathcal{F}^{-1}(\frac{\mathcal{F}(Gb)}{|\mathcal{F}(h_0)|^2 + \rho})) \quad (18)$$

El proceso consistió en primera instancia en mediante el FSPT y el reenfoque digital, obtener una imagen libre de ruido y enfocada en un plano α bajo una cierta longitud de onda elegida, luego, se introdujo el conjunto de planos focales y longitudes de onda como la imagen de baja resolución Y de un problema de optimización. Una vez incorporado, se emplea el algoritmo ADMM para resolver la optimización de la imagen enfocada en α en una longitud de onda λ . Esto permite obtener una imagen con resolución aumentada, enfocada y para una longitud de onda en específico.

4.2. Métricas

Con el objetivo de determinar si los resultados en cada parte son los deseados, se emplearon una serie de métricas para revisión de ruido, blur y calidad de imagen.

4.2.1 Roughness Laplacian Index

El Roughness Laplacian Index (RLP) [77] es un índice de imagen libre de imagen de referencia que cuantifica simultáneamente los artefactos de patrón fijo (FPN) y de desenfoque en cualquier imagen de infrarrojo (IR), para ello, combina dos operadores laplacianos. La detección de FPN se realiza mediante el operador laplaciano

discreto, mientras que el desenfoque se determina utilizando el Laplaciano de Gauss (LoG). Se definen tres parámetros para controlar la detección de artefactos de cada operador laplaciano.

El parámetro σ se utiliza para cuantificar los componentes de frecuencia espacial alta asociados con los artefactos FPN, o los componentes de frecuencia espacial baja relacionados con los artefactos de desenfoque. El parámetro w representa el tamaño de la ventana del filtro para el componente de desenfoque. Y el parámetro l determina la forma del operador laplaciano en el componente FPN. El RLP se expresa como un vector de dos componentes, tal como se muestra en la Ecuación 19.

$$\overrightarrow{RLP}(E_\alpha(x, y, \lambda_s)) = \langle 1 - LoG_\sigma(E_\alpha(x, y, \lambda_s)), \nabla^2(E_\alpha(x, y, \lambda_s)) \rangle \quad (19)$$

Donde $1 - LoG_\sigma$ evalúa el desenfoque y ∇^2 evalúa el FPN. El índice RLP se utiliza para cuantificar la presencia de desenfoque y FPN en todos los planos de imagen digitales de reenfoque α , de modo que se pueda determinar en qué α un objeto de la imagen está enfocado. Esto corresponde al plano de profundidad de enfoque α , donde la imagen presenta bajos niveles de FPN y desenfoque. La métrica puede aplicarse a imágenes individuales de igual manera, considerando eso sí que puede confundir ruido con detalle en imagen, esto no es un problema en un arreglo de imágenes, pero puede serlo al analizar imágenes individuales.

4.2.2 Structural Similarity Index

El Structural Similarity Index (SSIM) [78] es una métrica de evaluación de calidad de imágenes basada en referencia. El SSIM se utiliza para evaluar las imágenes obtenidas mediante procesamiento de imagen en comparación con una de referencia.

El SSIM combina la estructura local de la imagen, la luminancia y el contraste en una única puntuación de calidad, y se define en la ecuación 4.2.2.

$$SSIM = \frac{(2\mu_u\mu_b + C_1) + (\sigma_{ub} + C_2)}{(\mu_u^2 + \mu_b^2 + C_1) + (\sigma_u^2 + \sigma_b^2 + C_2)} \quad (20)$$

Donde μ_u, μ_b es la luminisencia promedio de cada imagen, σ_u, σ_b la varianza de intensidad de pixel y σ_{ub} la covarianza de intensidad de pixel entre imágenes. C_1 y C_2 son constantes incluidas para evitar inestabilidad en el denominador.

4.2.3 Peak Signal to Noise Ratio

El Peak Signal to Noise Ratio (PSNR) es una métrica basada en imagen de referencia que permite determinar el nivel de ruido presente en una imagen en comparación a su imagen referencial. Entre mayor sea su valor, la

imagen es de mejor calidad y posee menos ruido. Se define con la ecuación 4.2.3.

$$PSNR = 10 \log_{10} \frac{\max(X)^2}{MSE(X, REF)} \quad (21)$$

Donde $\max(X)$ es el máximo valor de pixel alcanzado por la imagen, y $MSE(X, REF)$ es el error cuadrático medio entre la imagen revisada y su referencia.

4.3. Setup - Montaje Experimental

Con el objetivo de solucionar los problemas del IR y aumentar la resolución del SWIR, se construye un sistema completo de corrección y aumento de resolución que sigue una metodología separada en una fase de simulación y una experimental, utilizando como base la eliminación de ruido mediante reenfoque digital propuesto por Coelho. El equipo utilizado corresponde a una cámara Xeva SWIR 320x256, y un desplazador mecánico Newport con capacidad de generar matrices de imágenes mediante el software LabView. El equipo y procedimiento de captura se detalla en la figura 4.1.

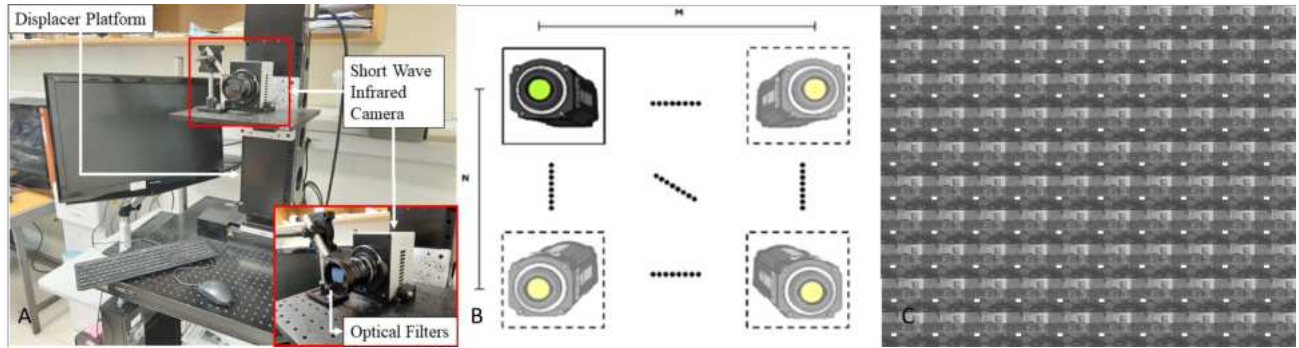


Figura 4.1: De izquierda a derecha, el equipo de captura y modo de trabajo de este. A) La cámara Xeva montada en un desplazador mecánico con su respectiva montura de filtro para generar el cubo multispectral. B) El movimiento del desplazador y la matriz $M \times N$ generada por este. C) El LF generado por el equipo, con cada imagen desplazada en el eje X e Y dependiendo de la fila y columna donde se ubique, la separación y cantidad de fotos son definidas por el usuario.

El equipo anterior permite generar matrices de imágenes donde cada imagen es de la resolución 320x256 y se captura de forma secuencial para generar un orden. Es necesario recordar que las características de utilizar un desplazador y cámara no permiten tener una escena cambiante en el tiempo, por lo que debe estar quieta durante la medición. El tamaño de la matriz utilizada y la separación entre las imágenes afecta la profundidad de campo del método de reenfoque. Por lo general es una buena idea utilizar matrices de 10x10 con una separación de 10 mm ya que así se obtiene una amplia profundidad de campo correctamente enfocada, sin embargo, estos valores pueden variar dependiendo de factores como lente utilizado, distancia de objetos y apertura de lente.

4.3.1 Simulación

La fase de simulación corresponde a la aplicación del método propuesto a una escena donde agregamos artefactos y evaluamos su desempeño. A modo de comparación con otras técnicas de SR, se compara con dos metodologías de SR en IR desarrolladas por Hardie et al[79], basada en MAP y Meza et al [80] basada en un método iterativo utilizando gradiente descendiente. Como escena de medición se utilizan dos envases de plástico colocados a 0,8 metros y 1,5 metros, a los cuales llamaremos OBJ1 para el más cercano y OBJ2 para el más lejano. Ambos están rellenos con agua y son candidatos ideales para revisión multiespectral en IR, ya que el agua puede verse a través del plástico en esta sección del espectro. La figura 4.2 muestra la escena capturada y los objetos de interés OBJ1 y OBJ2 que se analizarán. Los artefactos que se le agregan al LF son ruido, decimación y blur, donde el ruido es de forma grilla y gaussiano por separado y la decimación es de la mitad de la resolución de la imagen. La componente espectral se tradujo en utilizar 5 filtros de longitudes de onda distintas, estos son de 1160 nm, 1300 nm, 1400 nm, 1500 nm y 1600 nm, con esto, se abarca gran parte del rango de la cámara Xeva, el cual es entre 1080 nm y 1700 nm. El esquema del procedimiento se detalla en la figura 4.3.

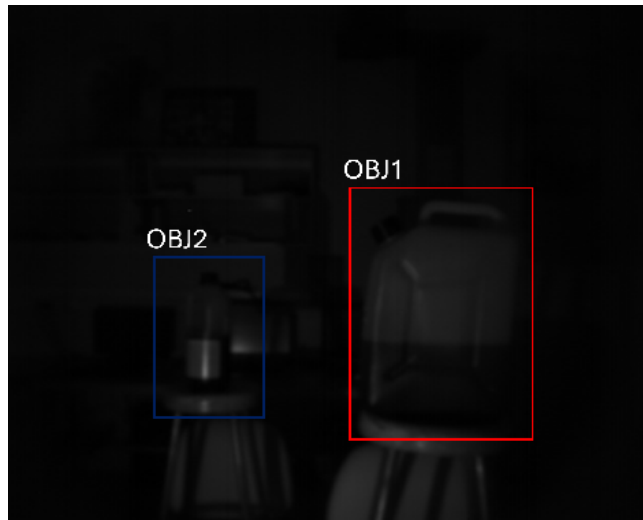


Figura 4.2: Escena IR a la cual se le aplicará la degradación de imagen y luego mejora mediante el método propuesto y los de Hardie y Meza. El OBJ1 corresponde al bidón de plástico más cercano relleno de agua, el OBJ2 corresponde al envase de plástico más lejano relleno de agua, el agua se puede apreciar como el color negro del envase.

4.3.2 Evaluación Experimental

La fase experimental consistió en aplicar directamente el método propuesto a imágenes no calibradas de la cámara Xeva y compararlo con la interpolación bicubica. Obteniendo de esta manera imágenes de resolución de 640x512 y 1280x1024 posteriormente. Para este caso, los artefactos son los naturales de la cámara y no se

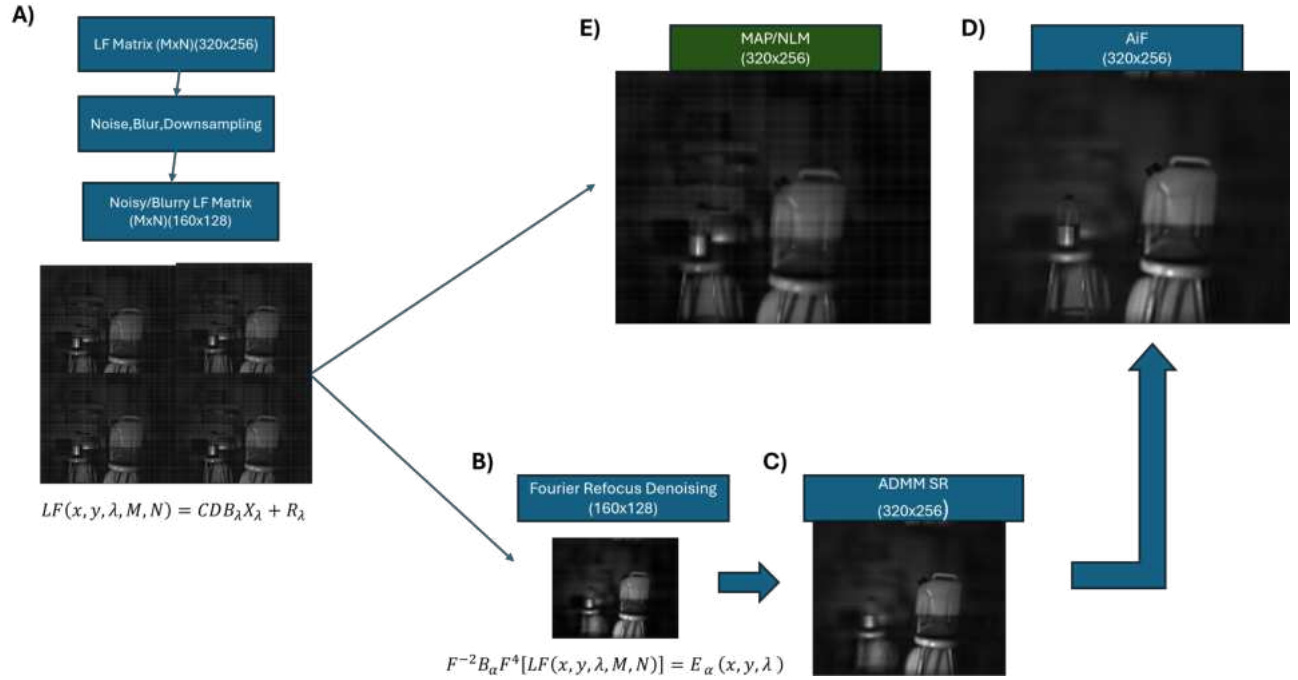


Figura 4.3: Esquema del procedimiento de captura para la fase de simulación. A) Se parte por capturar un LF de una escena SWIR calibrada sin ruido, luego se le añaden los artefactos de ruido, blur y decimación a la mitad, obteniéndose así un LF con imágenes de 160x128 y artefactos presentes. B) Se aplica el FSPT para generar Fourier Refocus Denoising (FRD) eliminando así el ruido en los objetos de interés OBJ1 y OBJ2. C) Se eligen ambos objetos enfocados y sin ruido por separado mediante el RLP para aplicarles el método de ADMM SR, con esto se obtiene cada imagen por separado con resolución aumentada, volviendo así a 320x256. D) Si bien se pueden comparar los objetos de interés por separado, se elige combinar ambos con un algoritmo de renfoque All in Focus (AiF) para tener una comparación visual más pareja con otros métodos. Esta fusión de imágenes no afecta la calidad de la imagen obtenida. E) Utilizando la misma entrada de LF con artefactos y decimación, se aplican los métodos de Hardie y Meza para comparar los resultados con el método propuesto.

posee una referencia, por lo que las métrica empleadas fueron solo libres de imagen de referencia. La escena consistió en una planta colocada a 1.1 m de la cámara y se utilizó un lente de 50 mm que entrega una escena de menor campo de visión, pero más zoom, esto fue hecho para poder apreciar de cerca el aumento de resolución y demostrar la adaptabilidad del método a distintos lentes ópticos. La figura 4.4 muestra la escena capturada con el FPN inherente del sensor y el objeto de interés (OOI). Las plantas son un buen objeto de estudio en el IR, ya que el agua presente en ellas, también llamado estrés hídrico, es visible en esta porción del espectro. De igual manera, se utilizaron los 5 filtros previamente mencionados. La figura 4.5 detalla el procedimiento seguido para esta fase.

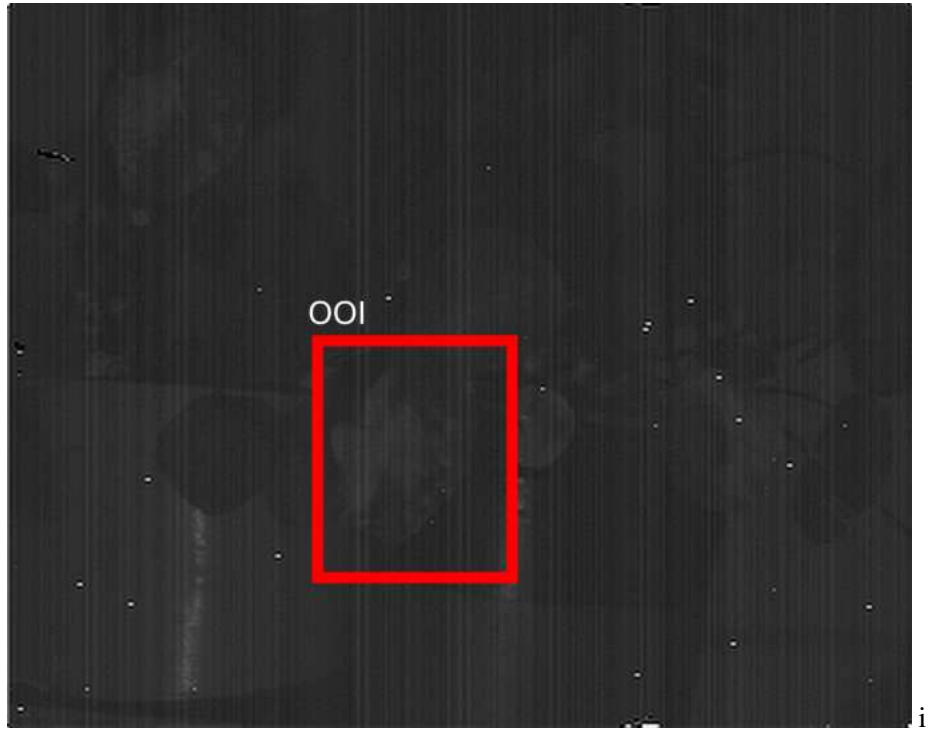


Figura 4.4: Escena IR con FPN inherente de la cámara SWIR Xeva 320x256. El OOI es una de las hojas de la planta, la cual presenta detalles elevados al utilizarse la banda IR correcta para detectar estrés hídrico.

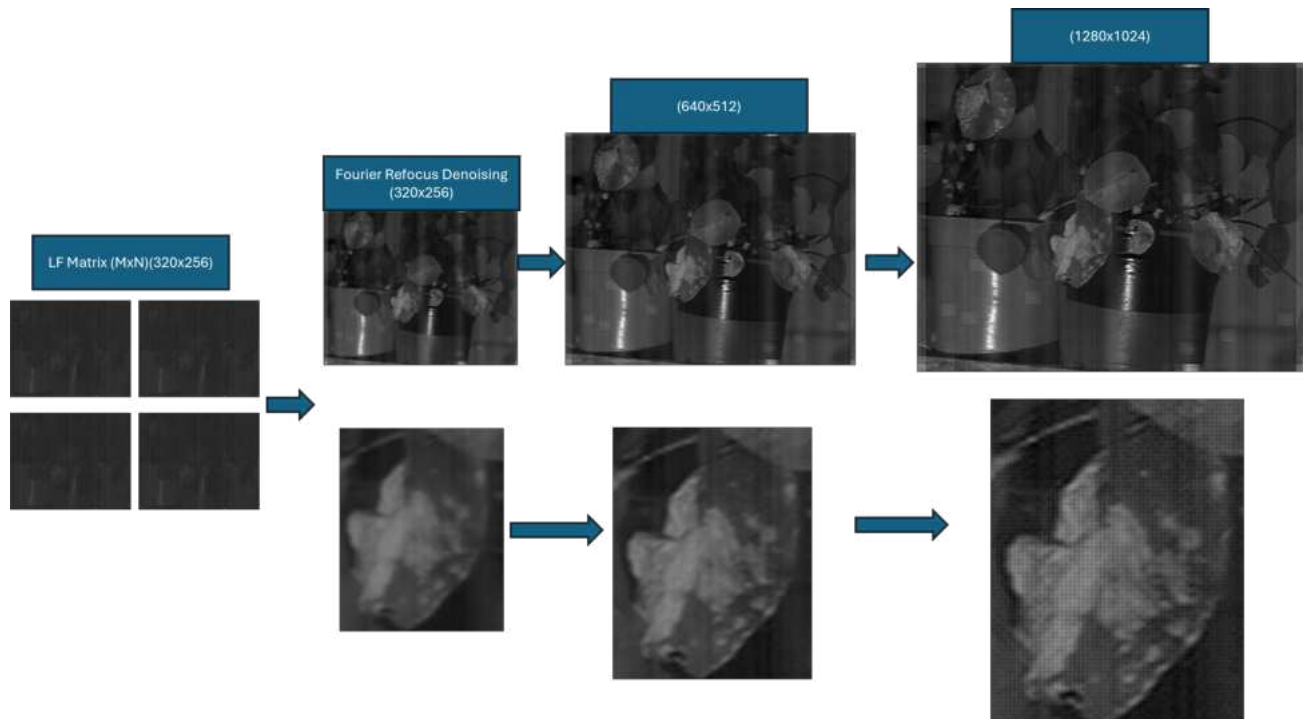


Figura 4.5: Esquema del procedimiento de captura para la fase experimental. De izquierda a derecha, se muestra como se introduce el LF con ruido de la cámara y sus sucesivos aumentos de resolución. Luego de aplicar el FRD, se aumenta la resolución a 640x512 y luego a 1280x1024, con el detalle de una de las hojas en la fila inferior.

5. Resultados

5.1. Simulación

Las pruebas de simulación consideraron una serie de factores y abarcaron cambios en una serie de variables, como lo son el tipo de ruido añadido, tamaño de la matriz utilizada, la longitud de onda y el objeto al que se le aplicaron las métricas (OBJ1, OBJ2 o la escena completa). Se divide, por ende, en 2 secciones correspondientes al tamaño de matriz empleada y al tipo de ruido. La longitud de onda común utilizada en esta sección fue de 1160 nm , ya que era de interés comparar con otros métodos en otros aspectos y un cambio de longitud de onda no aportaba en esta sección y el análisis se realizó para OBJ1, OBJ2 y la escena completa.

5.1.1 Simulación - Tamaño de Matriz

Las pruebas para el método propuesto, Hardie y Meza consideraron 4 tamaños de matrices, 2×2 , 4×4 , 8×8 y 10×10 , con el objetivo de abarcar gran parte del rango de movilidad del equipo utilizado. La figura 5.1 muestra visualmente como es la progresión de calidad obtenida, dependiendo del tamaño de matriz para cada método para un LF con ruido gaussiano añadido.

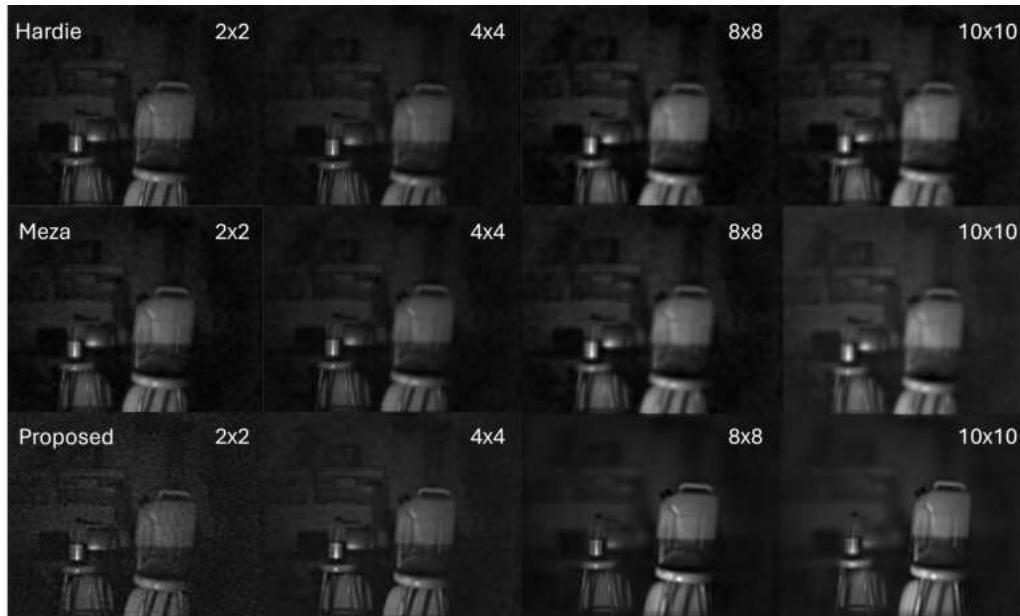


Figura 5.1: Variaciones en calidad de imagen obtenida dependiendo del tamaño de matriz empleado para cada método. Cada fila corresponde a un método distinto y las columnas separan los tamaños de matrices utilizados. Visualmente, uno puede percatarse de como los métodos varían sus resultados dependiendo del tamaño utilizado para la matriz.

La figura 5.1 da cuenta de la diferencia en calidad de imagen obtenida y muestra una clara diferencia en los resultados obtenidos por cada método al variar el tamaño de matriz. En particular, los métodos de Meza y

Hardie muestran mejores resultados en tamaños de matrices menores, mientras que el método propuesto mejora en calidad a medida que va en aumento el tamaño de matriz. Para los métodos de Meza y Hardie, entre mayor sea el tamaño de matriz, más difícil se le hace al algoritmo involucrar la gran cantidad de información y generar una imagen de buena calidad, mientras que para el método propuesto, menos imágenes no permiten al ruido eliminarse a tiempo para que los objetos estén enfocados, por ende, se enfoca con ruido. Este fenómeno se va corrigiendo a medida que disminuyen la cantidad de imágenes para los métodos Meza y Hardie, mientras que para el propuesto la mayor cantidad de imágenes eliminan el ruido a tiempo para enfocar los objetos y aplicar SR. Para entender el efecto negativo que se obtiene en tamaño de matrices pequeñas en el método propuesto, la figura 5.2 muestra las imágenes enfocadas antes de aplicar SR para los tamaños de matrices 2x2 y 10x10.

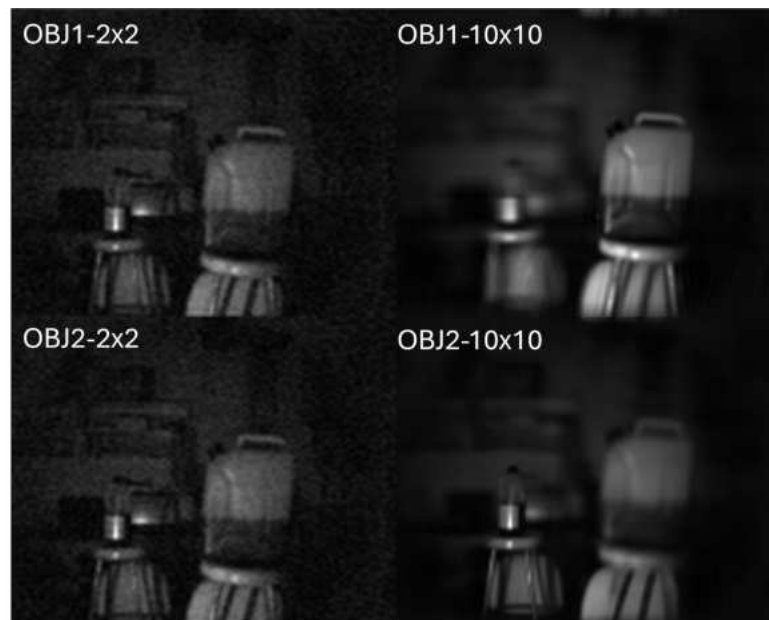


Figura 5.2: Planos focales para el OBJ1 y OBJ2 antes de aplicar SR y fusión de imagen. El tamaño de 2x2 genera planos focales con ruido para ambos objetos enfocados, mientras que el tamaño de 10x10 genera planos focales sin ruido, lo que provoca diferencias considerables en calidad al aplicar SR y fusión de imagen.

Se aplicaron las métricas descritas anteriormente (PSNR, SSIM y RLP) tanto para el OBJ1, OBJ2 y la escena completa. Cabe destacar que la diferencia entre los resultados del OBJ1 y OBJ2 son únicamente una diferencia en la magnitud de los valores, con el comportamiento manteniéndose igual en ambas (es decir, el OBJ2 es proporcional a los resultados del OBJ1). Esto es producto de que son el mismo objeto, solo que uno tiene menos detalle al estar más alejado, por lo que no aportan información nueva más allá de otro punto de referencia. Es por esto que se muestran en los resultados las métricas aplicadas al OBJ1 y a la escena completa, si se desea ver aplicadas las métricas al OBJ2 en el anexo estarán disponibles los resultados. La figura 5.3 muestra como se comporta el RLP para los stack de 2x2, 4x4, 8x8, 10x10 para el OBJ1.

Ignorando los frames iniciales donde el ruido aún está fuertemente presente, el plano focal se identifica

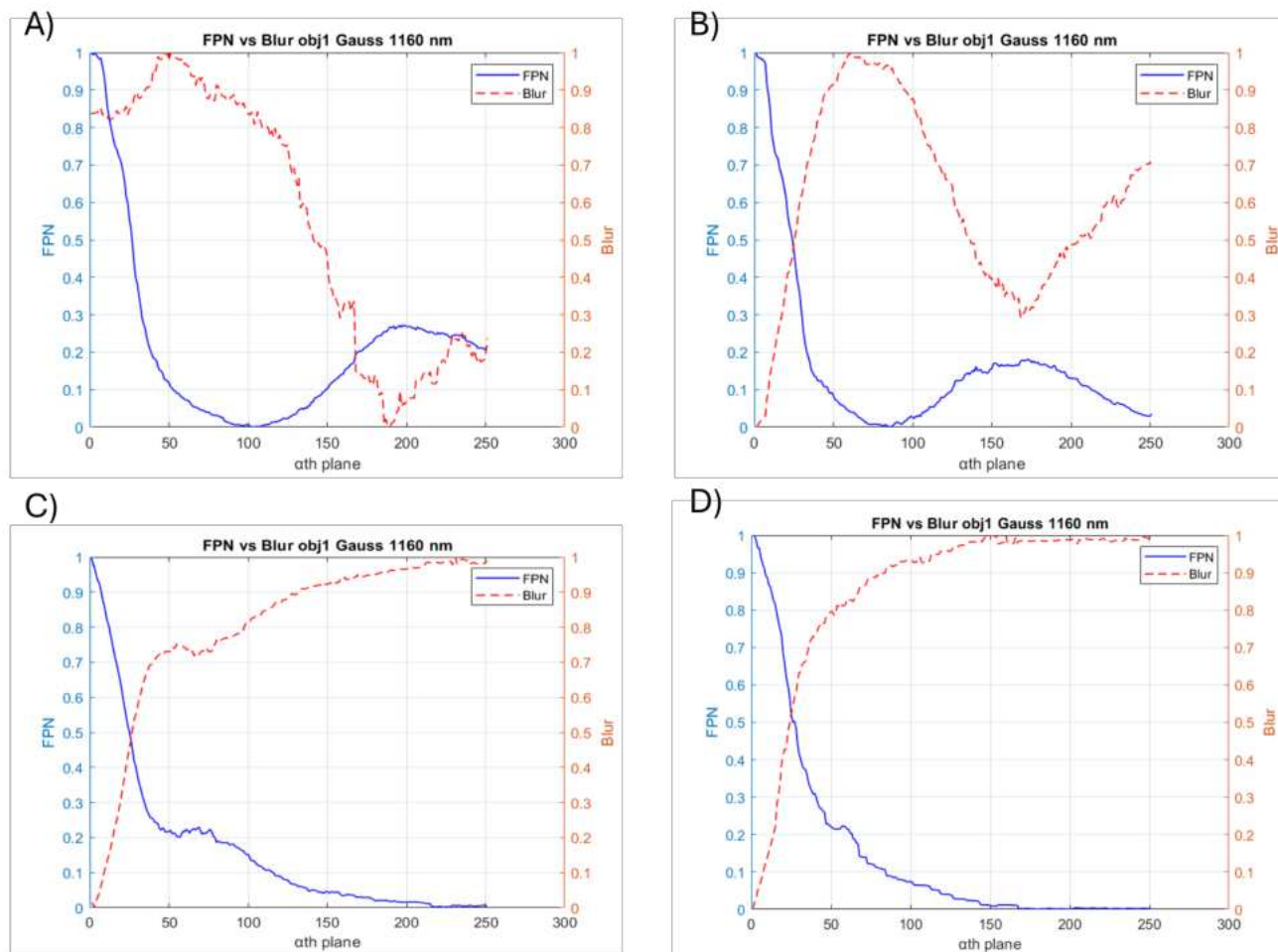


Figura 5.3: Comportamiento del RLP a lo largo de distintos tamaños de matriz tanto para blur como para el detalle (NU). Es necesario recordar que se está buscando el menor blur y un peak de ruido que lo acompañe, el motivo de esto es que esto será el plano enfocado, ya que interpretará el ruido como detalle la métrica. A) Para la matriz de 10x10 el plano focal está marcado por el peak de FPN y declive en blur en el frame 198. B) Para la matriz de 8x8 el plano es el 170. C) Para la matriz de 4x4 el plano focal es el 80. D) Para la matriz de 2x2 el plano focal es el 50

claramente en los gráficos, sin embargo, a medida que disminuye el tamaño de matriz, el plano focal se va desplazando hacia el fondo. Lo anterior debido a que el tamaño de matriz afecta directamente la profundidad de campo obtenida y por ende va desplazándolo hacia el fondo, acercándolo al plano del FPN. Esto es un factor importante a considerar, ya que a medida que se acerca al plano del FPN, la imagen al estar enfocada tendrá más o menos ruido presente, haciendo ineficiente enfocar al objeto. Es por esto que parámetros como apertura de lente, distancia de los objetos, entre otros, pueden alterarse para mejorar estos resultados, pero afectan los otros tamaños de matriz empleados igualmente. Independiente de esto, se realiza el análisis por métricas al OBJ1 y a la escena completa, los que se pueden ver desde la figura 5.4 hasta la 5.6.

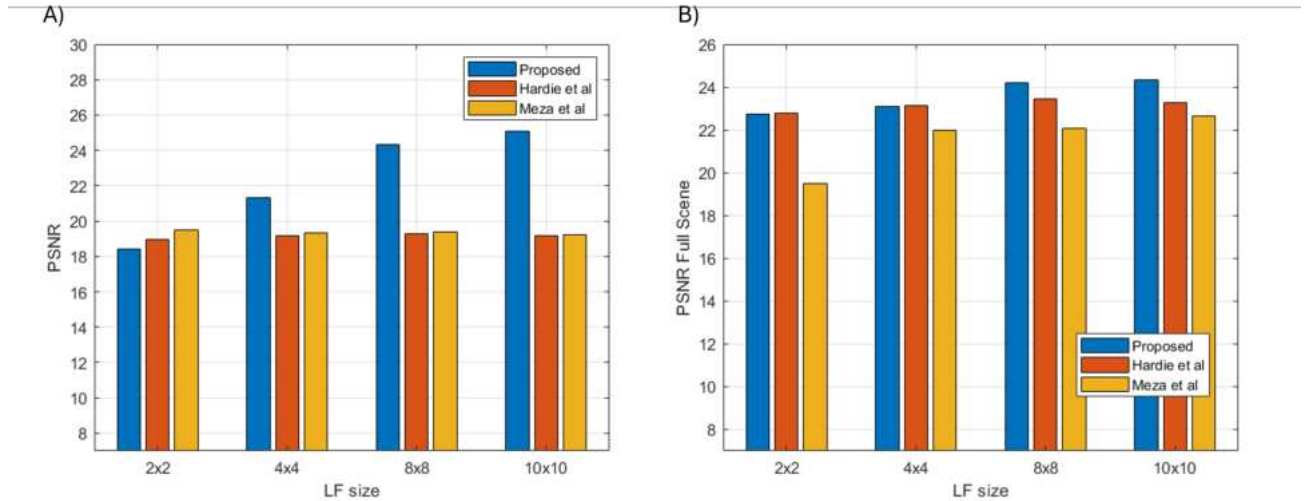


Figura 5.4: PSNR de OBJ1 y OBJ2 luego de aplicar el procedimiento descrito en la figura 4.3. A) el OBJ1 debido a su cercanía parte en valores más elevados de PSNR para los 3 métodos. Se observa que a medida que el tamaño de matriz aumenta, el método propuesto mejora en calidad de imagen, apreciándose la mejora en comparación a otros métodos desde 4x4 en adelante. B) La escena completa si bien sigue el mismo comportamiento, tiene valores de PSNR más cercanos entre los métodos, esto es producto de que los métodos de Hardie y Meza no lidian con el desenfoco y tienen la escena completa enfocada, mientras que el método propuesto realizó fusión de imagen para solo los dos objetos de interés, sin considerar el resto de la escena desenfocada. Pese a esto, el PSNR del método propuesto es superior debido a la eliminación de ruido.

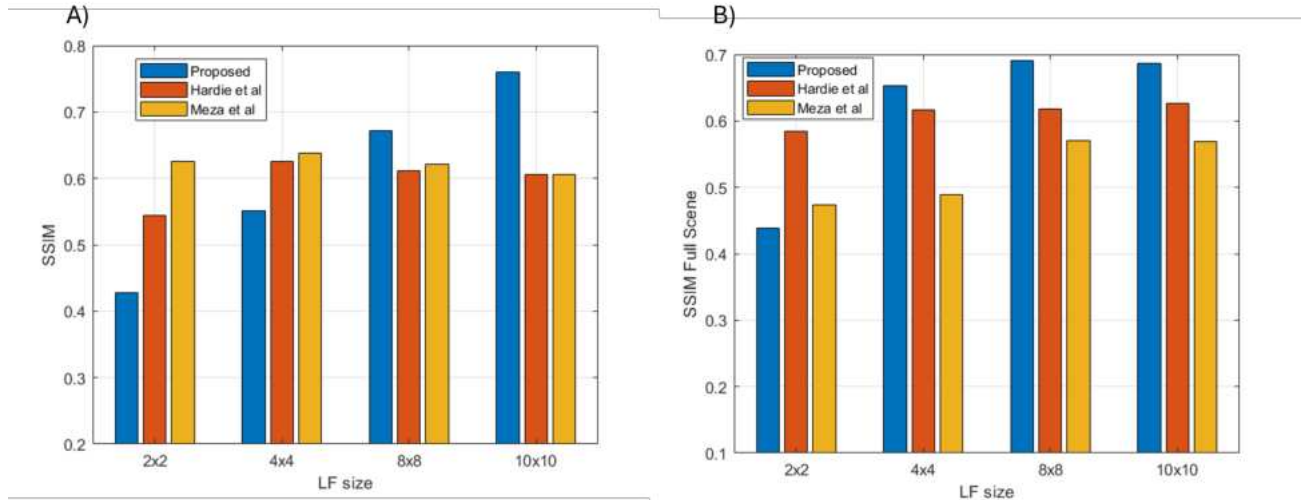


Figura 5.5: SSIM de OBJ1 y escena completa luego de aplicar el procedimiento descrito en la figura 4.3. A) el OBJ1 muestra un comportamiento similar al PSNR, donde los valores van en aumento a medida que el tamaño de matriz aumenta para el método propuesto y disminuyen para los otros métodos. En esta ocasión, la mejora del método propuesto se nota a partir de la matriz de 8x8 en adelante B) La escena completa sigue el mismo comportamiento de ir aumentando conforme el tamaño de matriz aumenta, sin embargo, los valores son mayores en general, esto es producto de que al considerar más de la escena, los pequeños detalles como bordes tienen menos incidencia en el valor final del SSIM.

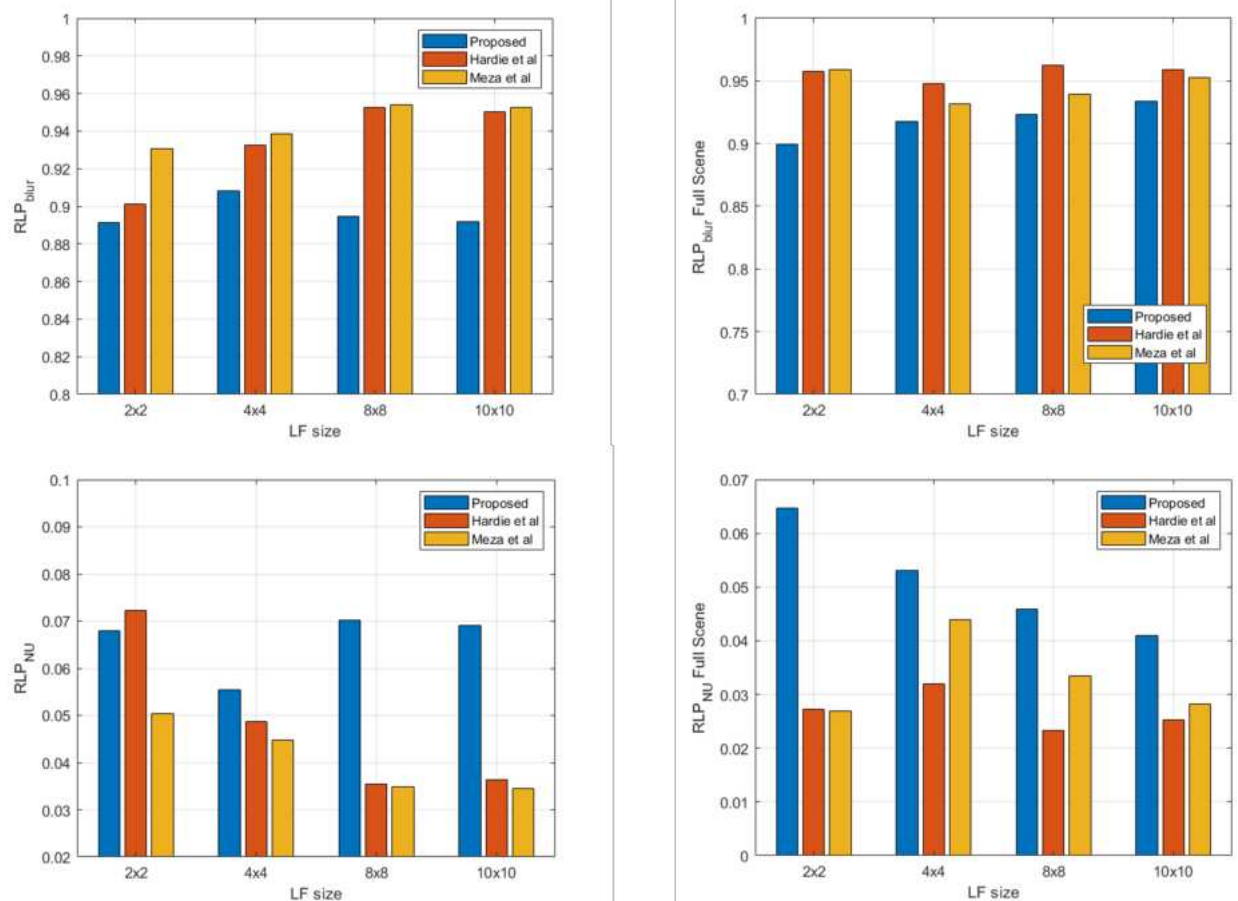


Figura 5.6: RLP de OBJ1 y escena completa luego de aplicar el procedimiento descrito en la figura 4.3, con la columna de la izquierda representando el OBJ1 y la de la derecha la escena completa. En el recuadro superior izquierdo, el OBJ1 exhibe un comportamiento de blur bajo para el método propuesto en comparación a los otros métodos, debido principalmente a la eliminación de ruido y mayor detalle. En el recuadro inferior izquierdo, el RLP_{NU} aumenta a medida que el tamaño de matriz aumenta para el método propuesto, mientras que disminuye para los otros métodos, señalando así un mayor detalle obtenido en la imagen para el propuesto. Resultados similares se pueden observar para la escena completa en la columna derecha.

Una revisión más exhaustiva de los resultados permite darse cuenta de que el fenómeno de mejor calidad de imagen para matrices de mayor tamaño es notorio para el método propuesto, en cambio, ocurre lo inverso para los métodos de comparación. El PSNR y SSIM son métricas claras en sus resultados, sin embargo, no ocurre lo mismo para las métricas relacionadas con el RLP, las que deben interpretarse basados en el contexto de la imagen analizada. Esto es debido a 2 motivos, el primero, es que la métrica del RLP fue diseñada para analizar arreglos de imágenes con distinto punto focal, por lo que su análisis en imágenes individuales no es su objetivo inicial. El segundo es que es fácil para el RLP confundir ruido con detalle, lo que no es problema cuando se utiliza en el método de eliminación de ruido, ya que se pueden separar los peaks de ruido y detalle. Sin embargo, para imágenes individuales, el valor del RLP_{NU} no tiene referencia del ruido y no se logran separar contextualmente ruido y detalle en una imagen, provocando así por ejemplo que para el RLP_{NU} del OBJ1 en la figura 5.6 el valor

para matrices de 2x2 sea mayor que el de 4x4 para el método propuesto. Lo anterior ocurre, ya que el fuerte valor de ruido presente en reenfoque digital en matrices de 2x2 del método propuesto es confundido con detalle en la imagen.

Resumiendo los resultados, el mejor resultado posible para el método propuesto involucra una matriz de 10x10, mientras que los otros dos métodos utilizan matrices de 2x2. Al compararlos, el método propuesto entrega mejores resultados en comparación, además, supera a los métodos propuestos en tamaños de matrices de 4x4 en adelante. Lo anterior solidifica los buenos resultados obtenidos por el método propuesto y su capacidad de trabajar con varios tamaños de matriz.

5.1.2 Simulación - Tipo de Ruido

El ruido de un sensor IR puede presentarse de distintas formas, por lo que esta sección consideró al menos 2 tipos de ruido, de grilla y de tipo gaussiano (es decir, en forma de valores aleatorios de pixel grilla y como puntos específicos). Manteniendo el tamaño de matriz constante, se revisaron los resultados para ambos tipos de ruido, los que se pueden ver en la figura 5.7.

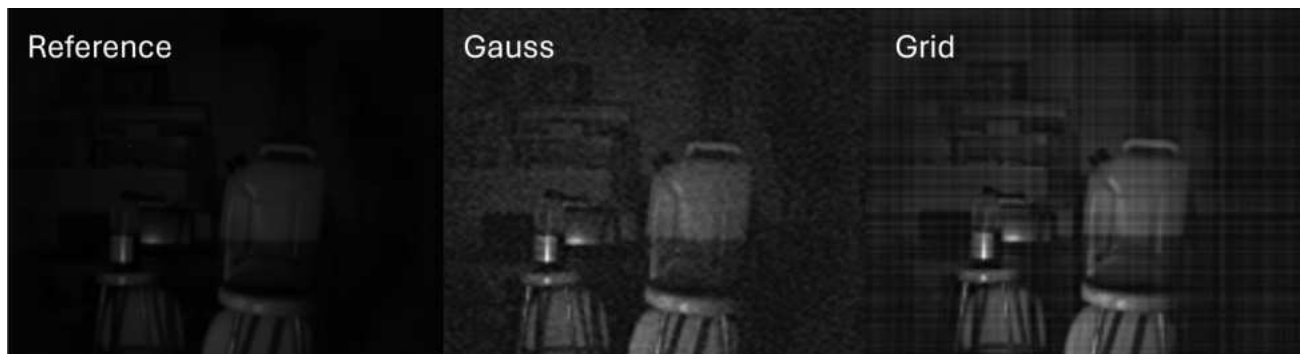


Figura 5.7: De izquierda a derecha se tiene la imagen de referencia, la imagen con ruido gaussiano y la imagen con ruido de grilla. El ruido para ambos casos considero una desviación estándar de 0.01.

Se le aplicaron las mismas métricas anteriormente descritas a ambos tipos de ruido, obteniéndose los gráficos en la figura 5.8 y 5.9.

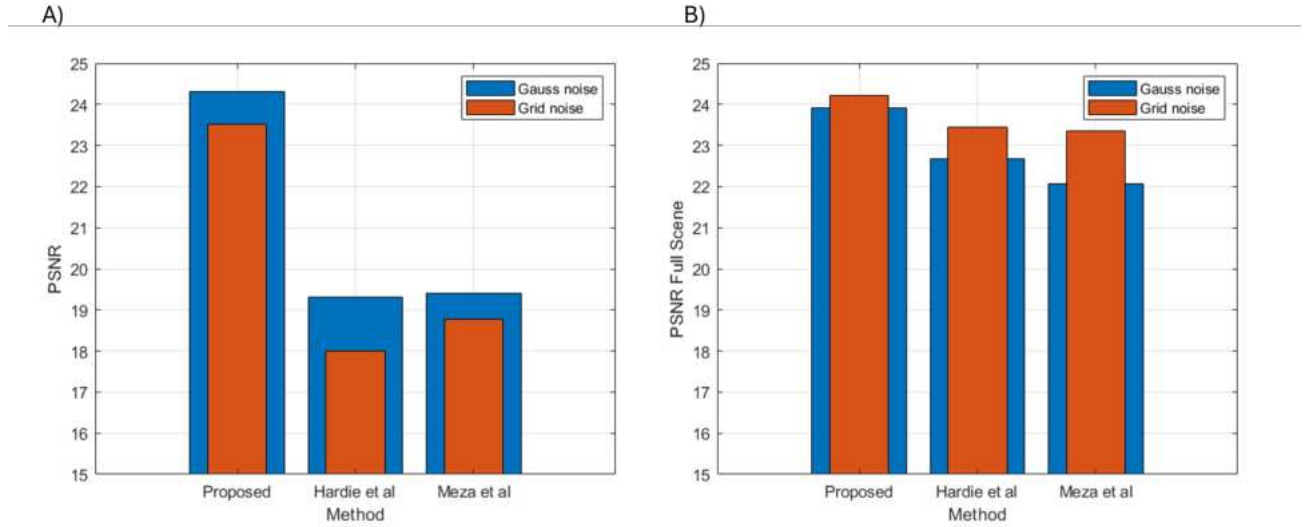


Figura 5.8: PSNR de OBJ1 y escena completa luego de aplicar el procedimiento descrito en la figura 4.3 para una matriz de 8x8. A) Para el OBJ1, si bien el valor de PSNR de gauss es superior al de grilla, ambos mantienen superioridad en el método propuesto sobre los métodos de Meza y Hardie. B) Para la escena completa, los métodos obtienen valores más cercanos entre ellos, manteniendo eso si la superioridad del método propuesto.

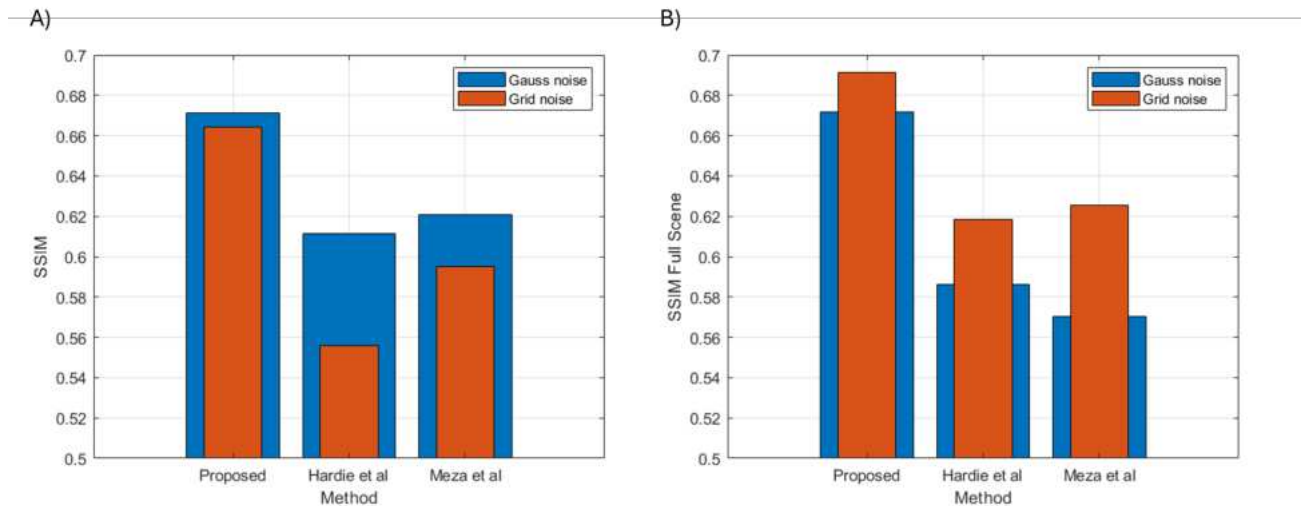


Figura 5.9: SSIM de OBJ1 y escena completa luego de aplicar el procedimiento descrito en la figura 4.3 para una matriz de 8x8. A) Para el OBJ1, los valores de SSIM siguen una tendencia similar a los valores de PSNR, con una clara superioridad del método propuesto sobre los otros dos B) Para la escena completa, los métodos obtienen valores más cercanos entre ellos, manteniendo eso si la superioridad del método propuesto.

Donde se mantiene el comportamiento tanto para el ruido gaussiano como el ruido de grilla, resaltando así la adaptabilidad del método propuesto ante los distintos tipos de ruido.

5.2. Experimental

Las pruebas experimentales utilizaron imágenes obtenidas sin calibración de la cámara Xeva 320 en distintas longitudes de onda. Como objeto de interés se utilizó una planta con hojas bajo evidente estrés hídrico, que es visible en el espectro SWIR. En este escenario se probó aumentar la resolución en x2 y x4, obteniéndose así imágenes de 640x512 y 1280x1024 píxeles. El aumento de resolución se realizó de forma progresiva, es decir, se aumentó primero a 640x512 y luego a 1280x1024 con distintas longitudes de onda, mostrando así la adaptabilidad del método propuesto ante diferentes escenarios. Como comparación al método propuesto, se incluyó la interpolación bicubica para comprobar con un método ya existente. El análisis se realizó para una hoja en específico denominada OOI y se aplicó solamente la métrica RLP, ya que no se tiene una imagen de referencia bajo la cual comparar para utilizar otras métricas.

5.2.1 Experimental - Resoluciones

Se utilizó una configuración estándar consistente en un tamaño de matriz de 10x10, con una longitud de onda de 1400 nm. Esta longitud de onda se eligió debido a que la presencia de estrés hídrico es prevalente en bandas cercanas a este valor. La figura 5.10 muestra los resultados en la banda de 1400 nm para el OOI y como es en comparación a la interpolación bicúbica aplicada al mismo escenario.

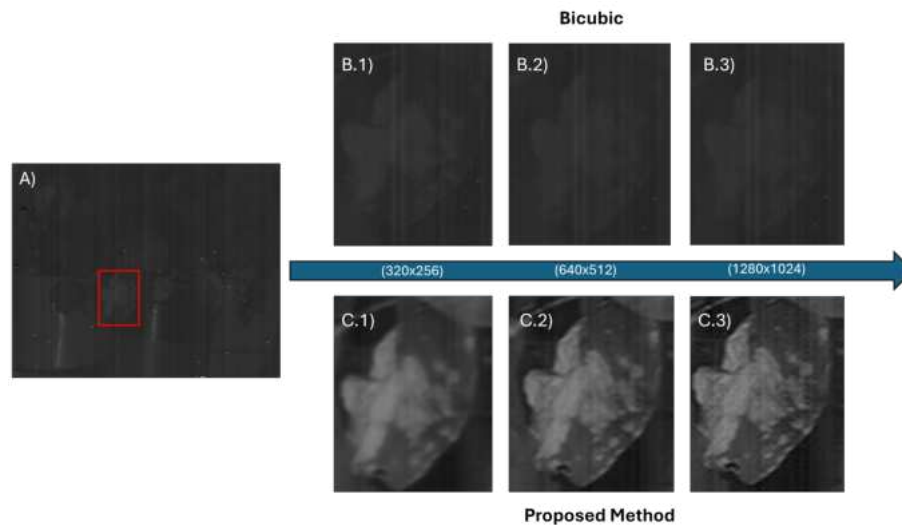


Figura 5.10: Aumento en resolución de imagen aplicando la interpolación bicúbica y el método propuesto a una imagen A, con aumentos de resolución de x2 y x4. Para las imágenes B 1-3, la interpolación bicúbica no logra superar el FPN presente en la imagen y los resultados son pobres para el aumento de resolución. Para las imágenes C 1-3 Se logra eliminar el ruido y lograr un aumento de resolución exitoso, aumentando el detalle de la imagen original y por ende su calidad.

Se aprecia claramente como la figura 5.10 el método propuesto entrega imágenes de calidad mejorada a medida que se aumenta la resolución, apreciándose detalles mejorados en la hoja. Para comprobar los efectos

visuales, se aplicaron las métricas de RLP, las cuales se pueden ver en la figura 5.11.

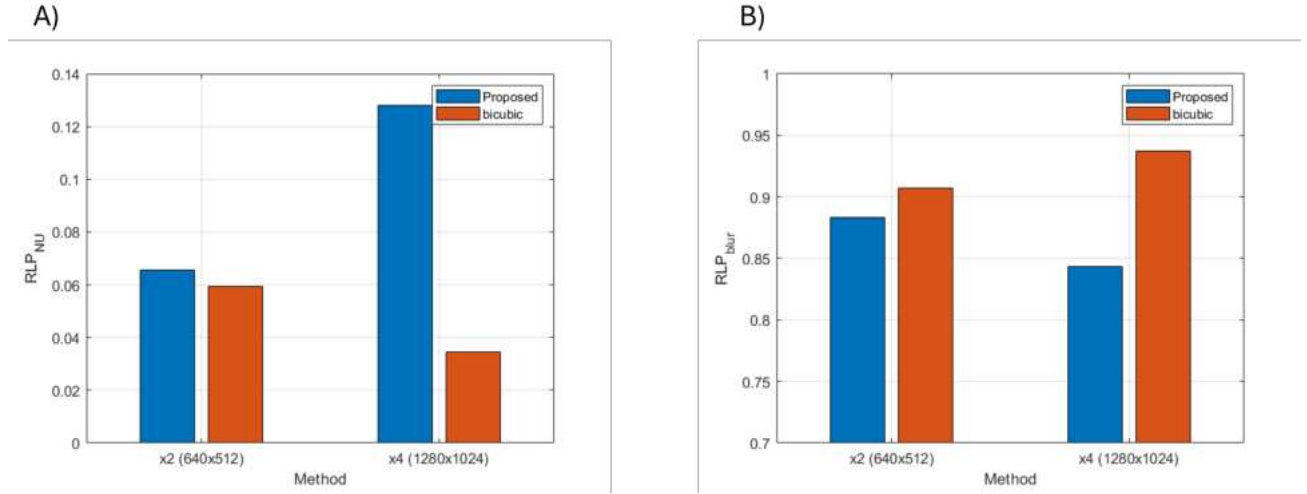


Figura 5.11: RLP de OOI después de aplicar el método bicúbico y el método propuesto. A) El RLP_{NU} de la imagen con 640x512 muestra la superioridad del método propuesto, que se ve aumentada al pasar a la resolución de 1280x1024, cabe destacar que el contexto de porque se asemejan los valores en la resolución x2 es debido a la fuerte presencia de ruido en el método bicúbico que es interpretado como detalle por la métrica. B) El blur es notoriamente prevalente para el método bicúbico, con el método propuesto disminuyendo el valor a medida que aumenta la resolución.

5.2.2 Experimental - Longitudes de onda

Utilizando la misma configuración anterior, ahora se utilizaron distintos filtros ópticos para comparar el comportamiento a lo largo del espectro SWIR, en particular, las longitudes de onda de 1160 nm, 1300 nm, 1400 nm, 1500 nm y 1600 nm. La figura 5.12 muestra como se comporta el OOI al aplicar los métodos bicúbico y propuesto.

En la figura 5.12 se muestra como la resolución aumenta para cada una de las escenas para distintas longitudes de onda. Independiente de la longitud de onda utilizada, el aumento de resolución logra aumentar la calidad de la imagen mediante una mejora en los detalles de esta. Lo anterior se puede apreciar mejor en longitudes de onda donde el estrés hídrico está presente, como lo es desde los 1400 nm en adelante.

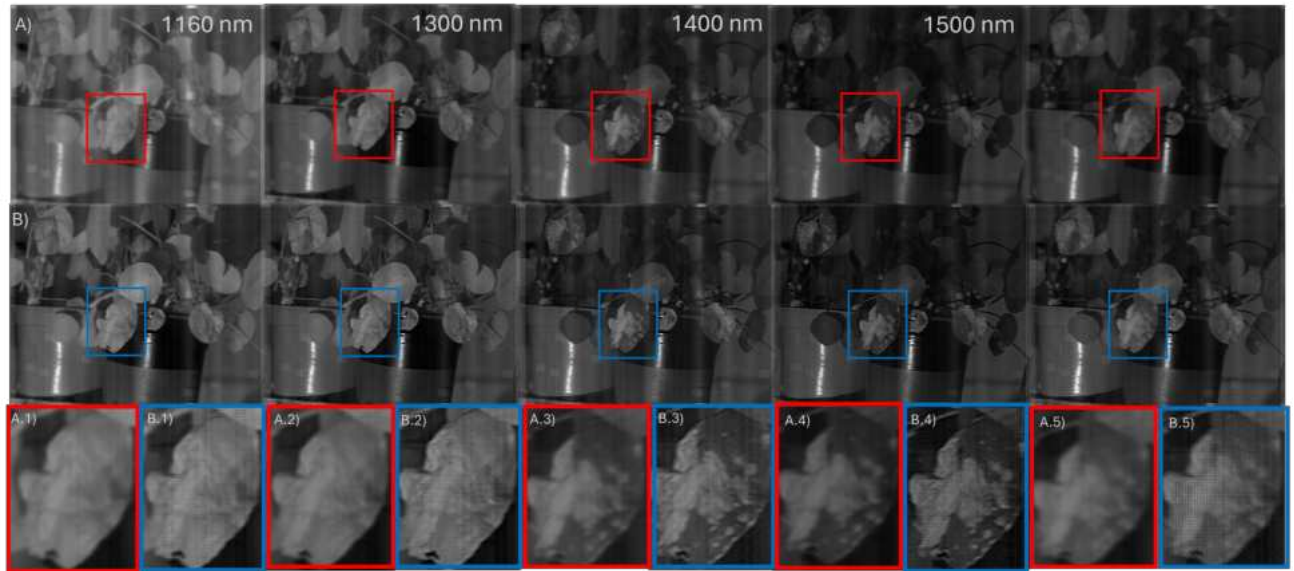


Figura 5.12: Diferentes longitudes de onda para la misma escena, mostrando la imagen inicial sin ruido, la imagen con el método propuesto y un aumento en el OOI. A) La imagen inicial con el ruido ya eliminado mediante el método propuesto con una resolución de 320x256 para 5 distintas longitudes de onda. B) Imágenes resultantes de aplicar la sección de aumento de resolución del método propuesto. Para este caso se llega a la resolución de 1280x1024. A 1-5) Acercamiento del objeto de interés OOI para las imágenes de la fila A, es decir, sin ruido tras la primera parte del método propuesto. B 1-5) Acercamiento del objeto de interés OOI para las imágenes de la fila B, es decir, con la parte de aumento de resolución incluida.

6. Discusión

Para el caso de simulación, los resultados destacan la adaptabilidad del método propuesto a distintos tipos de ruido y su comportamiento en distintos tamaños de matriz. Con eficacia consistente en las variaciones de artefactos y un comportamiento creciente en calidad respecto al tamaño de matriz, que comparativamente es superior a otros métodos, al menos en matrices de 4x4 en adelante. Las métricas empleadas favorecen al método propuesto, sin embargo, para el caso del RLP se debe tener en consideración el efecto que el ruido puede tener. Es por esto que en conjunto al SSIM y PSNR se puede llegar a resultados coherentes con el RLP cumpliendo principalmente su función de determinar el plano focal, su función de determinar la calidad de imagen posterior a esto es un apoyo a las otras métricas.

En el caso de los resultados experimentales, se prueba con aumentos de resolución de x2 y x4, con los resultados de x4 visiblemente superiores en comparación al método bicúbico. Nuevamente, cabe destacar que el RLP como métrica de calidad puede confundir el ruido con detalle en imagen, por lo que su uso debe ser consciente de que aplicar un método de SR que no elimine el ruido puede obtener valores mayores. La adaptabilidad nuevamente es revisada en esta sección al analizar diferentes longitudes de onda mediante el uso de filtros ópticos y los coeficientes de Zernike. El método de SR original empleado no consideraba PSF variable ni ruido fuerte presente, por lo que era relativamente rígido. Con la inclusión de Zernike se pueden usar diferentes PSF para

cada filtro óptico y así obtener un modelo de imagen fiel a lo capturado y por ende mejores resultados.

Los resultados obtenidos son demostraciones de como el método propuesto se desarrolla en diferentes condiciones, ya sea con distintos ruidos, longitudes de onda variables, tamaños de matrices, entre otros. Desde un punto de vista de desempeño en comparación a otros métodos de SR, el propuesto permite obtener resultados de mejor calidad respecto a otros métodos ya existentes que utilizan LF como su entrada. El método propuesto logra combinar la información entregada por el LF para generar no solo una imagen sin ruido, sino que aprovecha además de realizar SR y fusión de imagen con la información resultante, aprovechando así la totalidad del resultado obtenido.

Lo anterior permite desarrollar líneas de investigación avanzadas con equipo sin la resolución adecuada, ejemplo de esto es la combinación de dos espectros IR con resoluciones distintas, como lo son el SWIR y el LWIR. A modo de ejemplo de su uso, se realizó una captura de la misma escena con 2 cámaras IR distintas, una Xeva SWIR (320x256) y una Gobi LWIR (640x480), ambas de distinta resolución y distinto espectro. Normalmente, combinar ambos espectros en una sola imagen es dificultoso debido a la enorme diferencia en resolución espacial, sin embargo, al utilizar el método propuesto en la SWIR, se puede llegar a la resolución superior del LWIR y combinar ambos espectros. A modo de ejemplo, se utilizó como escena una botella rellena de agua caliente, que si bien su nivel en el envase es visible en el SWIR, su temperatura no lo es. La figura 6.1 muestra la configuración utilizada.

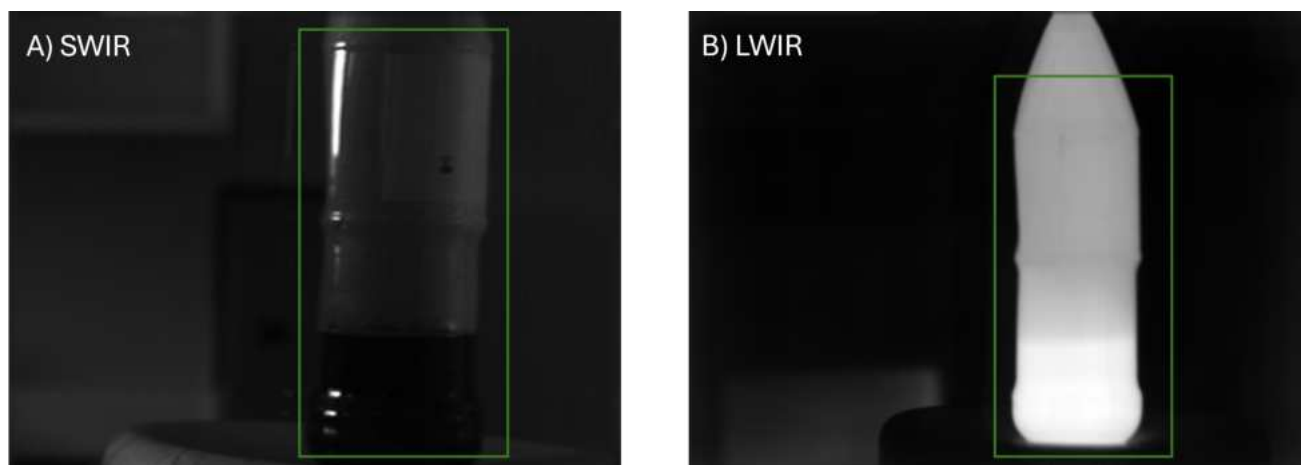


Figura 6.1: Escena de agua caliente capturada con dos cámaras distintas A) La cámara Xeva SWIR captura el nivel de agua en la botella pero con una resolución de 320x256. B) La cámara Gobi LWIR captura la temperatura del agua con una resolución mayor de 640x480 pero sin determinar el nivel del agua.

Ambas imágenes fueron obtenidas de un LF con ruido inherente de la cámara, es decir, con FPN que hacía difícil de identificar los objetos y sus propiedades. Ahora, al aplicar la sección de SR del método a la cámara SWIR, se obtiene una imagen de una resolución cercana al LWIR (640x512) con la cual se puede generar fusión de imagen y obtener una escena con la información de ambos espectros. La figura 6.2 detalla como se ven ambas

imágenes por separado y la combinación de ambas, ahora coherente debido a la similitud en resolución.

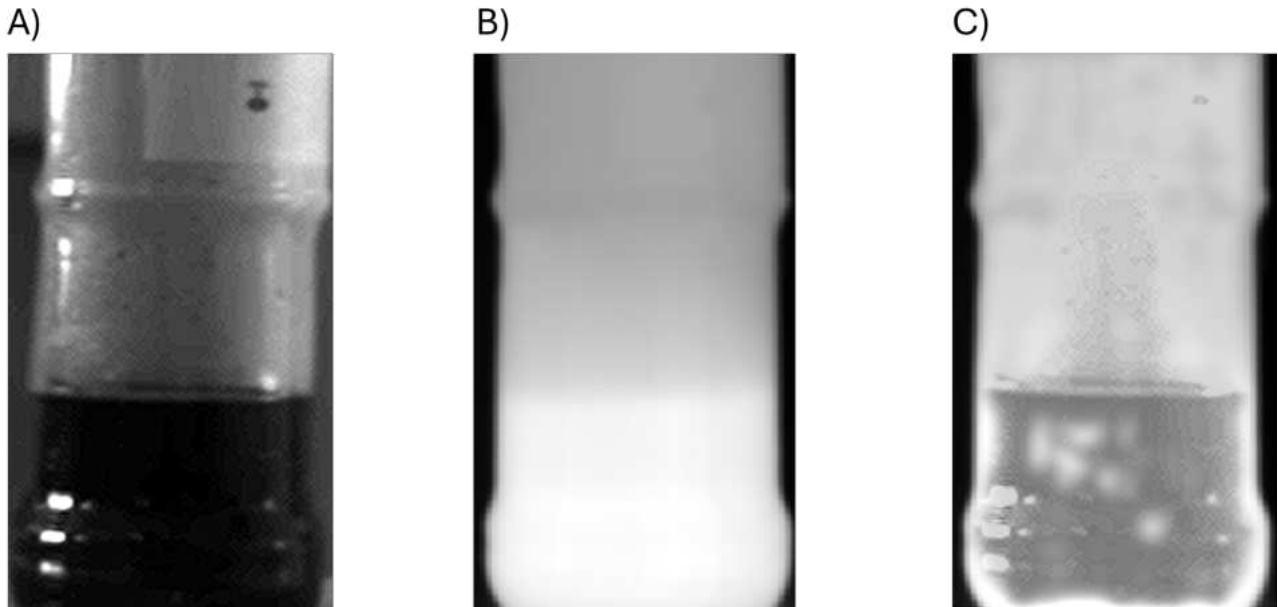


Figura 6.2: Escena capturada por distintas cámaras y su combinación en fusión de imagen. A) La botella en SWIR con el nivel de agua identificado por el color negro. B) La botella en LWIR con la temperatura irradiada en blanco. C) la fusión de ambas imágenes con el detalle del nivel y la emisión.

Como se puede ver, en la imagen resultante se pueden apreciar tanto el nivel del líquido en la botella como su forma original a través de las emisiones de calor, que muchas veces hacen difícil determinar la forma original del envase. Si bien visualmente los resultados son útiles, es necesario destacar que las propiedades físicas de los objetos se pierden al fusionarlas, es decir, la intensidad de pixel ahora no es indicador de temperatura u otro fenómeno, sino que meramente de forma u alguna característica física atenuada producto de la fusión de imagen.

Resultados como el anterior respaldan la utilidad de tener un sistema y método de SR con eliminación de ruido y fusión de imagen, que permita utilizar la mayor cantidad de información disponible en el LF para generar resultados de calidad.

7. Conclusiones y trabajo a futuro

La demostración de la solidez de la metodología propuesta, en conjunto con ejemplo de aplicaciones, muestran como los objetivos propuestos al principio fueron logrados. Revisando la hipótesis entonces y basándonos en lo expuesto, es posible aumentar la resolución del SWIR a resoluciones mayores como lo puede ser el NIR el 1024x1080, el LWIR en 640x480 u otras cámaras del mismo espectro con distintas resoluciones espaciales. No solo eso, sino que además el método propuesto es adaptable a ruido en las escenas y consistente a lo largo

del espectro, permitiendo así utilizar un sistema de captura completo que lidie con el FPN, la baja resolución y el enfoque digital del IR.

Como trabajo a futuro quedan unas cuantas interrogantes que revisar, principalmente una inclusión más directa de la información del LF al proceso de SR, si bien se utilizaron las imágenes entregadas por una captura plenóptica, aún se puede acortar más el procesamiento de imagen e incluir el FSPT a la SR directamente. Otro punto a destacar es el hecho de que el tamaño de 2×2 puede ser mejorado para superar a los otros métodos, pero requeriría un análisis más exhaustivo de la física involucrada y parámetros como apertura de lente, distancia de objetos, enfoque del lente, entre otros, lo que se desvía del enfoque del trabajo entregado. Aparte de lo anterior mencionado, sería interesante utilizar el método propuesto en aplicaciones donde la fusión de imágenes de baja calidad con imágenes de alta calidad sea de evidente ayuda, para así poder refinar un proceso que para este trabajo fue revisada de manera superficial.

8. Anexo

8.1. Simulación - Tamaño Matriz - OBJ2

A continuación de la figura 8.1 hasta la 8.5, se muestran los resultados obtenidos para el OBJ2 en la sección de simulación.

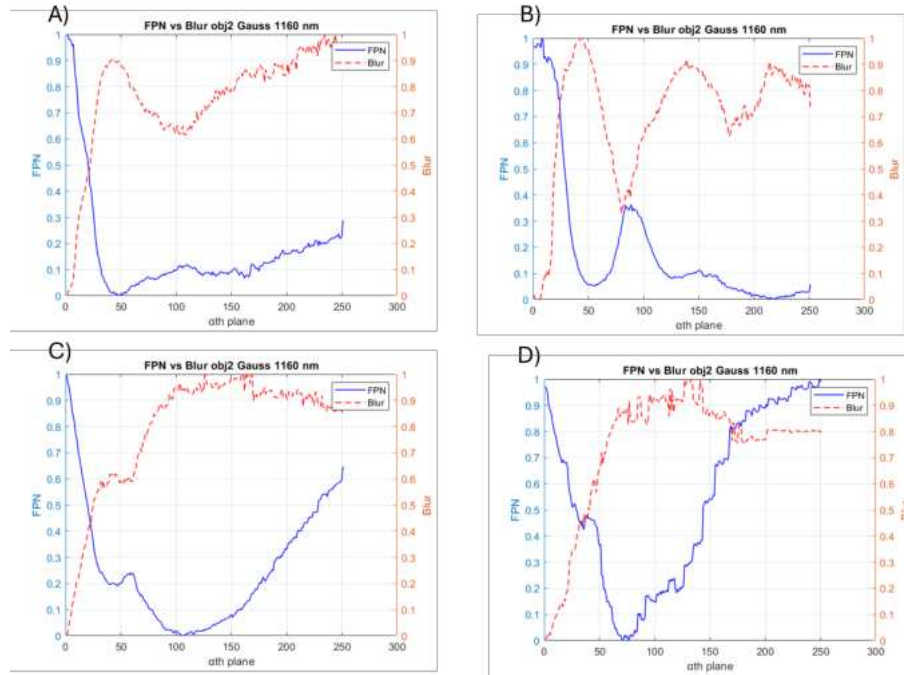


Figura 8.1: Comportamiento del RLP a lo largo de distintos tamaños de matriz tanto para blur como para el detalle (NU). Es necesario recordar que se está buscando el menor blur y un peak de ruido que lo acompañe, el motivo de esto es que esto será el plano enfocado, ya que interpretará el ruido como detalle la métrica. A) Para la matriz de 10x10 el plano focal está marcado por el peak de FPN y declive en blur en el frame 110. B) Para la matriz de 8x8 el plano es el 80. C) Para la matriz de 4x4 el plano focal es el 60. D) Para la matriz de 2x2 el plano focal es el 23.

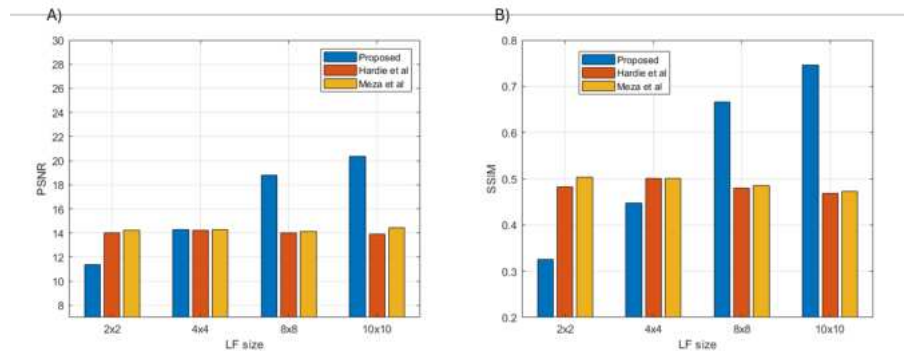


Figura 8.2: SSIM y PSNR del OBJ2. A) El PSNR sigue el comportamiento de ir en aumento a medida que se aumenta el tamaño de matriz para el método propuesto, llegando a valores de 22 dB. B) El SSIM aumenta considerablemente para los tamaños de 8x8 y 10x10 para el método propuesto, alcanzando valores de 0.67 y 0.75 respectivamente.

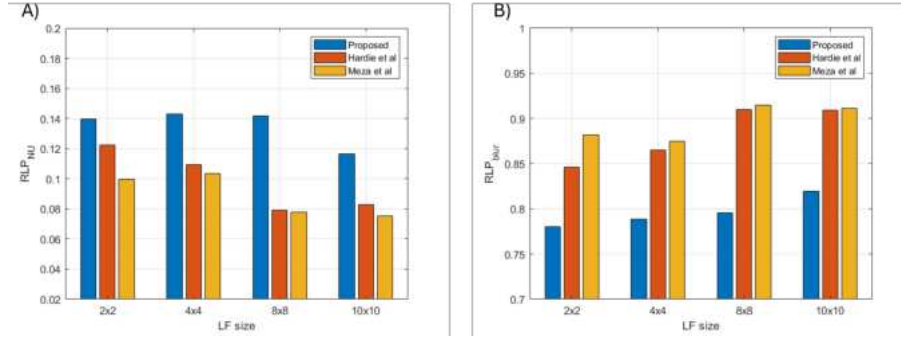


Figura 8.3: Comportamiento del RLP para el OBJ2. A) El detalle de la imagen es consistente en ser mayor para el método propuesto para todos los tamaños de matriz, sin embargo, cabe recordar que el detalle con ruido pueden confundirse en estos tamaños pequeños. B) El blur se mantiene en niveles bajos para el método propuesto a lo largo de los distintos tamaños de matrices, demostrando la versatilidad en enfocar objetos a distancia.

8.2. Simulación - Tipos de Ruido - OBJ2

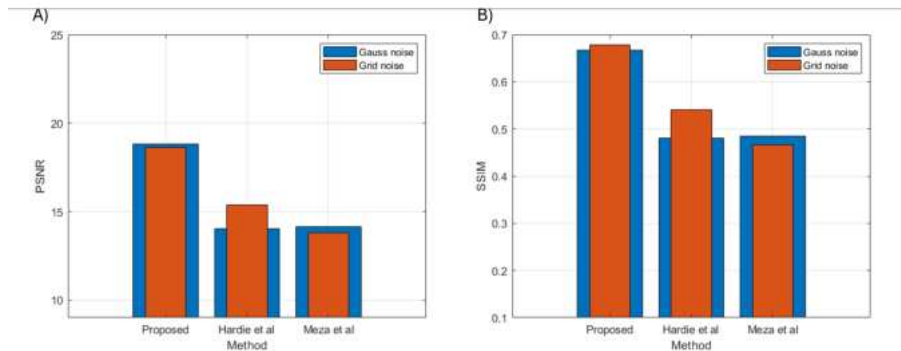


Figura 8.4: SSIM y PSNR del OBJ2 para ruido gaussiano y de grilla. A) Los valores de PSNR son superiores para el método propuesto tanto para ruido gaussiano como de grilla. B) El SSIM mantiene valores mejores para el método propuesto en ambos tipos de ruido.

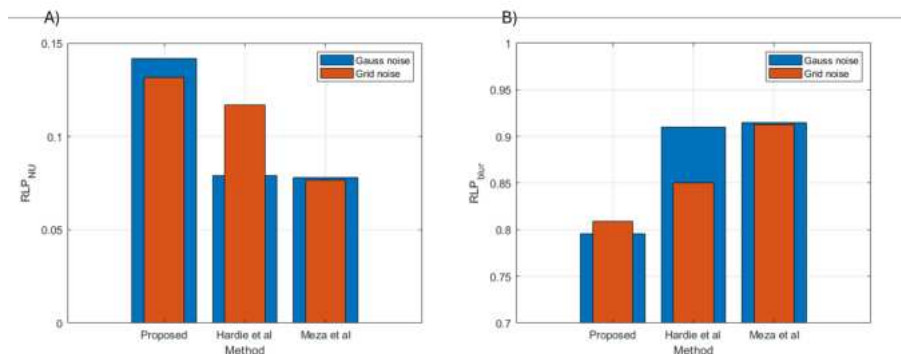


Figura 8.5: Comportamiento de RLP para ambos tipos de ruido. A) el detalle de la imagen NU es mayor para ambos tipos de ruido en el caso del método propuesto. B) El blur mantiene valores menores para el método propuesto para ambos tipos de ruido.

Referencias

- [1] Minchul Kim and Taeweon Suh. Eavesdropping vulnerability and countermeasure in infrared communication for iot devices. *Sensors*, 21(24):8207, 2021.
- [2] Pablo A. Coelho and Jorge E. Tapia et al. Infrared light field imaging system free of fixed-pattern-noise. *Scientific Report*, 2017.
- [3] Laury Chaerle and Dominique Van Der Straeten. Imaging techniques and the early detection of plant stress. *Trends in plant science*, 5:495–501, 2000.
- [4] M. Rodrigues, R. B. de Oliveira, G. L. A. A. Dos Santos, K. M. de Oliveira, A. S. Reis, R. H. Furlanetto, and M. R. Nanni. Rapid quantification of alkaloids, sugar and yield of tobacco (*nicotiana tabacum* l.) varieties by using vis–nir–swir spectroradiometry. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 274:121082, 2022.
- [5] F. Hu, M. Zhou, P. Yan, K. Bian, and R. Dai. Multispectral imaging: A new solution for identification of coal and gangue. *IEEE Access*, 7:169697–169704, 2019.
- [6] R. A. Oliveira, R. Näsi, P. Korhonen, et al. High-precision estimation of grass quality and quantity using uas-based vnir and swir hyperspectral cameras and machine learning. *Precision Agriculture*, 2023.
- [7] G. Bonifazi, L. Fiore, R. Gasbarrone, R. Palmieri, and S. Serranti. Hyperspectral imaging applied to weee plastic recycling: A methodological approach. *Sustainability*, 15(14):11345, 2023.
- [8] L. Wiley, J. Follansbee, P. Leslie, O. Furxhi, R. Pimpinella, D. Brady, and R. Driggers. Target discrimination in the extended swir (eswir) band (2–2.5 μm) compared to vis, nir, and swir in degraded visual environments. In *Infrared Imaging Systems: Design, Analysis, Modeling, and Testing XXXIII*, volume 12106, pages 41–51. SPIE, May 2022.
- [9] M. Levoy and P. Hanrahan. Light field rendering. In *Proceedings of SIGGRAPH*, 1996.
- [10] E. H. Adelson and J. R. Bergen. The plenoptic function and the elements of early vision. *Computational Models of Visual Processing*, 1(2):3–20, 1991.
- [11] A. Daveri, S. Paziani, M. Marmion, H. Harju, A. Vidman, M. Azzarelli, and M. Vagnini. New perspectives in the non-invasive, in situ identification of painting materials: The advanced mwir hyperspectral imaging. *TrAC Trends in Analytical Chemistry*, 98:143–148, 2018.

- [12] Y. Li, M. Tofighi, J. Geng, V. Monga, and Y. C. Eldar. Efficient and interpretable deep blind image deblurring via algorithm unrolling. *IEEE Transactions on Computational Imaging*, 6:666–681, 2020.
- [13] J. Rim, H. Lee, J. Won, and S. Cho. Real-world blur dataset for learning and benchmarking deblurring algorithms. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXV 16*, pages 184–201. Springer International Publishing, 2020.
- [14] M. Labussière, C. Teulière, and O. Ait-Aider. Blur aware metric depth estimation with multi-focus plenoptic cameras. *Computer Vision and Image Understanding*, 235:103802, 2023.
- [15] Z. Cai, X. Liu, G. Pedrini, W. Osten, and X. Peng. Light-field depth estimation considering plenoptic imaging distortion. *Optics Express*, 28(3):4156–4168, 2020.
- [16] J. Zhang, Y. Zhou, and F. Fang. Depth error correction for plenoptic cameras based on an innovative ar system combining geometrical waveguide and micro-lens array. *Optics and Lasers in Engineering*, 167:107629, 2023.
- [17] Z. Wang, J. Chen, and S. C. H. Hoi. Deep learning for image super-resolution: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(10):3365–3387, Oct. 2021.
- [18] Y. Huang, Z. Jiang, R. Lan, S. Zhang, and K. Pi. Infrared image super-resolution via transfer learning and psrgan. *IEEE Signal Processing Letters*, 28:982–986, 2021.
- [19] N. Ren. Fourier slice photography. *ACM Transactions on Graphics*, 24(3):735–744, Jul 2005.
- [20] Jürgen Hartmann et al. Blackbody and other calibration sources. In *Experimental Methods in the Physical Sciences*, volume 42, pages 241–295. Elsevier, 2009.
- [21] M. Kłosowski and Y. Sun. Fixed pattern noise reduction and linearity improvement in time-mode cmos image sensors. *Sensors*, 20(20):5921, 2020.
- [22] N. Chen, J. Zhang, S. Zhong, W. Mao, D. Hu, and L. Yao. Two-dimensional calibration for fixed-pattern noise reduction of thermal images. In *2019 IEEE International Symposium on Circuits and Systems (IS-CAS)*, pages 1–5, Sapporo, Japan, 2019.
- [23] Shahla Keyvan, Rodney Rossow, and Carlos Romero. Blackbody-based calibration for temperature calculations in the visible and near-ir spectral ranges using a spectrometer. *Fuel*, 85(5-6):796–802, 2006.
- [24] J. Guan, R. Lai, A. Xiong, Z. Liu, and L. Gu. Fixed pattern noise reduction for infrared images based on cascade residual attention cnn. *Neurocomputing*, 377:301–313, 2020.

- [25] S. H. Kim, J. Lee, H. Kwen, J. H. Park, K. I. Lee, and J. K. Shin. Pixel dark current calibration method to reduce dark fixed pattern noise in amorphous silicon bolometer-type uncooled infrared image sensor. *Sensors and Materials*, 32, 2020.
- [26] H. I. Ashiba, N. Sadic, E. S. Hassan, S. El-Dolil, and F. E. Abd El-Samie. New proposed algorithms for infrared video sequences non-uniformity correction. *Wireless Personal Communications*, 126(2):1051–1073, 2022.
- [27] V. F. Utkin. Scene-based non-uniformity fixed pattern noise correction algorithm for infrared video sequences. (*unpublished*), 2020.
- [28] T. Zhang, X. Li, J. Li, and Z. Xu. Cmos fixed pattern noise removal based on low rank sparse variational method. *Applied Sciences*, 10(11):3694, 2020.
- [29] A. K. Gurralla, G. Rahul, S. Amalan, S. Preejith, and M. Sivaprakasam. Eliminating vertical fixed pattern noise in cmos-based endoscopic images using modified dark frame subtraction. In *2023 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, pages 1–6, Jeju, Korea, Republic of, 2023.
- [30] J. G. Harris and Y. M. Chiang. Nonuniformity correction using constant average statistics constraint: Analog and digital implementations. *Proceedings of SPIE*, 3061:10–17, 1996.
- [31] John G. Harris and Yu-Ming Chiang. An analog implementation of the constant average statistics constraint for sensor calibration. 9:699, 1997.
- [32] D. A. Scribner, K. A. Sarkady, M. R. Kruer, J. T. Caulfield, J. D. Hunt, , and C. Herman. Adaptive nonuniformity correction for ir focal-plane arrays using neural networks. *Proc. SPIE*, 1541:100–109, 1991.
- [33] B. M. Ratliff, M. M. Hayat, , and R. C. Hardie. An algebraic algorithm for nonuniformity correction in focal-plane arrays. *J. Opt. Soc. Am. A*, 19:1737–1747, 2002.
- [34] B. M. Ratliff, M. M. Hayat, , and J. S. Tyo. Generalized algebraic scene-based nonuniformity correction algorithm. *Journal of the Optical Society of America A*, 22, 2005.
- [35] Ratliff BM; Hayat MM; Tyo JS. 'radiometrically accurate scene-based nonuniformity correction for array sensors. *Journal of the Optical Society of America A*, 20:1890–1899, 2003.
- [36] S. N. Torres and M. M. Hayat. Kalman filtering for adaptive nonuniformity correction in infrared focal-plane arrays. *J. Opt. Soc. Am. A*, 20:470–480, 2003.

- [37] Carlos Saavedra and Sergio Torres et al. Process allowing the removal through digital refocusing of fixed-pattern-noise in effective images formed by electromagnetic sensor arrays in a light field, 2021.
- [38] J. Xie, G. Hou, G. Wang, and Z. Pan. A variational framework for underwater image dehazing and deblurring. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(6):3514–3526, June 2022.
- [39] V. Rengarajan, S. Zhao, R. Zhen, J. Glotzbach, H. Sheikh, and A. C. Sankaranarayanan. Photosequencing of motion blur using short and long exposures. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pages 510–511, 2020.
- [40] G. Jiang, X. Jin, R. Deng, K. Sun, J. Yang, and W. Lv. Underwater plenoptic cameras optimized for water refraction. *Optics Express*, 31(13):21464–21481, 2023.
- [41] R. D. Fischer, M. Moaven, D. Kelly, S. Morris, B. Thurow, and B. C. Prorok. 3d tracking velocimetry of l-pbf spatter particles using a single high-speed plenoptic camera. *Additive Manufacturing Letters*, 3:100083, 2022.
- [42] A. C. Maia and J. A. Gaspar. Plenoptic camera viewpoints rotation to improve 3d face reconstruction. In *2023 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)*, pages 163–168, Tomar, Portugal, 2023.
- [43] Sihao Gao, Jun Li, Ping Wang, and Xiaojun Xu. Learning to detect 3d planes from a single plenoptic image. *Applied Optics*, 62(14):5050–5056, 2023.
- [44] M. F. Eberhart, S. Loehle, F. Grigat, J. Connolly, M. McGilvray, and D. Gillespie. 3d measurement of ice crystal accretion using a plenoptic camera. In *AIAA SCITECH 2023 Forum*, page 0806, 2023.
- [45] Takahiro Sasaki and James R. Leger. Non-line-of-sight object location estimation from scattered light using plenoptic data. *Journal of the Optical Society of America A*, 38(2):211–228, 2021.
- [46] Mingce Chen, Mao Ye, Zhe Wang, Chai Hu, Taige Liu, Kewei Liu, Jiashuo Shi, and Xinyu Zhang. Electrically addressed focal stack plenoptic camera based on a liquid-crystal microlens array for all-in-focus imaging. *Optics Express*, 30(25):34938–34955, 2022.
- [47] Q. Qi, M. M. Hossain, B. Zhang, T. Ling, and C. Xu. Flame temperature reconstruction through a multi-plenoptic camera technique. *Measurement Science and Technology*, 30(12):124002, 2019.
- [48] M. Zhang, M. Vogelbacher, K. Aleksandrov, H. J. Gehrmann, D. Stapf, R. Streier, and J. Matthes. A novel plenoptic camera-based measurement system for the investigation into flight and combustion behavior of refuse-derived fuel particles. *ACS omega*, 8(19):16700–16712, 2023.

- [49] C. H. Lin, P. M. Lin, C. W. Chen, and M. F. Chen. Design and validation of aberration-free plenoptic imaging system for industrial inspection. In *2022 IEEE International Conference on Imaging Systems and Techniques (IST)*, pages 1–6, Kaohsiung, Taiwan, 2022.
- [50] A. Scagliola, F. Di Lena, A. Garuccio, M. D’Angelo, and F. V. Pepe. Correlation plenoptic imaging for microscopy applications. *Physics Letters A*, 384(19):126472, 2020.
- [51] E. Longo, D. Alj, J. Batenburg, O. de La Rochefoucauld, C. Herzog, I. Greving, and P. Zeitoun. Flexible plenoptic x-ray microscopy. *Photonics*, 9(2):98, February 2022.
- [52] P. Matysiak, M. Grogan, M. Le Pendu, M. Alain, E. Zerman, and A. Smolic. High quality light field extraction and post-processing for raw plenoptic data. *IEEE Transactions on Image Processing*, 29:4188–4203, 2020.
- [53] T. Zhong, X. Jin, and K. Tong. 3d-cnn autoencoder for plenoptic image compression. In *2020 IEEE International Conference on Visual Communications and Image Processing (VCIP)*, pages 209–212, Macau, China, 2020.
- [54] F. Bhatti and T. Greiner. Design of an fpga hardware optimizing the performance and power consumption of a plenoptic camera depth estimation algorithm. *Algorithms*, 14(7):215, 2021.
- [55] Kamal Nasrollahi and Thomas B Moeslund. Super-resolution: a comprehensive survey. *Machine Vision and Applications*, 25:1423–1468, 2014.
- [56] R. Morin, A. Basarab, and D. Kouame. Alternating direction method of multipliers framework for super-resolution in ultrasound imaging. In *Proceedings of IEEE International Symposium on Biomedical Imaging*, Spain, 2012.
- [57] T. Bengtsson, I. Y-H. Gu, M. Viberg, and K. Lindstrom. Regularized optimization for joint super-resolution and high dynamic range image reconstruction in a perceptually uniform domain. In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 1097–1100, Japan, 2012.
- [58] S. Zhang. Application of super-resolution image reconstruction to digital holography. *EURASIP Journal on Advances in Signal Processing*, 2006(7):90358, 2006.
- [59] Syed Muhammad Arsalan Bashir et al. A comprehensive review of deep learning-based single image super-resolution. *PeerJ Computer Science*, 7:e621, 2021.
- [60] K. Kim and Y. Kwon. Example-based learning for single-image sr and jpeg artifact removal. Technical Report 173, MPI-TR, 08 2008.

- [61] Shengyang Dai, Mei Han, Wei Xu, Ying Wu, and Yihong Gong. Soft edge smoothness prior for alpha channel super resolution. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–8, 2007.
- [62] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Image super-resolution using deep convolutional networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(2):295–307, 2015.
- [63] Y. Liang, R. Timofte, J. Wang, Y. Gong, and N. Zheng. Single image super resolution-when model adaptation matters. *arXiv preprint arXiv:1703.10889*, 2017.
- [64] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, et al. Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4681–4690, 2017.
- [65] C. Latry and B. Rouge. Super resolution: quincunx sampling and fusion processing. In *Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 315–317, Toulouse, France, 2003.
- [66] H. Wang and D. Wen. The progress of sub-pixel imaging methods. In *Proceedings of the SPIE Conference Series*, pages 91420K–1, 2014.
- [67] P. Netinant, P. Vasprasert, and M. Rukhiran. Evaluations of effective on lwir micro thermal camera iot and digital thermometer for human body temperatures. In *Proceedings of the 2021 The 5th International Conference on E-Commerce, E-Business and E-Government*, pages 20–24, April 2021.
- [68] Diego A. Socolinsky, Andrea Selinger, and Joshua D. Neuheisel. Face recognition with visible and thermal infrared imagery. *Computer Vision and Image Understanding*, 91(1–2):72–114, 2003.
- [69] R. Vadivambal and D. S. Jayas. Applications of thermal imaging in agriculture and food industry—a review. *Food and Bioprocess Technology*, 4:186–199, 2011.
- [70] M. T. Sánchez, J. A. Entrenas, I. Torres, M. Vega, and D. Pérez-Marín. Monitoring texture and other quality parameters in spinach plants using nir spectroscopy. *Computers and Electronics in Agriculture*, 155:446–452, 2018.
- [71] P. Hauff. An overview of vis-nir-swir field spectroscopy as applied to precious metals exploration. *Spectral*, 2008.

- [72] D. W. Messinger, J. Van Aardt, D. McKeown, M. Casterline, J. Faulring, N. Raqueño, and M. Velez-Reyes. High-resolution and lidar imaging support to the haiti earthquake relief effort. In *Imaging Spectrometry XV*, volume 7812, pages 177–188. SPIE, August 2010.
- [73] A. A. Gowen and et al. Hyperspectral imaging—an emerging process analytical tool for food quality and safety control. *Trends in Food Science and Technology*, 18(12):590–598, 2007.
- [74] J. Kuula and et al. Using vis/nir and ir spectral cameras for detecting and separating crime scene details. In *Proceedings of SPIE*, volume 8359, page 83590P, 2012.
- [75] Stanley H. Chan, Xiran Wang, and Omar A. Elgendy. Plug-and-play admm for image restoration: Fixed point convergence and applications, 2016.
- [76] Guillermo Machuca, Pablo Meza, and Esteban Vera. Single-shot super-resolution and non-uniformity correction through wavefront modulation in infrared imaging systems. *Journal of Applied Remote Sensing*, 18(2):022205, March 2024.
- [77] A. Jara, G. Machuca, S. N. Torres, P. A. Coelho, and F. Perez. A reference-free image index to simultaneously quantify infrared-imaging fixed-pattern-noise and blur artifacts. *IEEE Access*, 9:121593–121607, 2021.
- [78] Zhou Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, April 2004.
- [79] R. C. Hardie and D. R. Droege. A map estimator for simultaneous superresolution and detector nonuniformity correction. *EURASIP Journal on Advances in Signal Processing*, 2007(1):089354, 2007.
- [80] P. Meza, G. Machuca, S. Torres, C. S. Martin, and E. Vera. Simultaneous digital super-resolution and nonuniformity correction for infrared imaging systems. *Applied Optics*, 54(21):6508–6515, 2015.