



Universidad
de La Laguna

Escuela de Doctorado
y Estudios de Posgrado

TÍTULO DE LA TESIS DOCTORAL

Estudio y Mejora de la Técnica en Representación de Información Tridimensional y Bidimensional sobre display

AUTOR/A

David

Carmona

Ballester

DIRECTOR/A

José Manuel

Rodríguez

Ramos

CODIRECTOR/A

Juan Manuel

Trujillo

Sevilla

DEPARTAMENTO O INSTITUTO UNIVERSITARIO

FECHA DE LECTURA

05/03/21

Estudio y Mejora de la Técnica en Representación de Información Tridimensional y Bidimensional sobre *display*



David Carmona Ballester

Director: Dr. Jose M. Rodríguez-Ramos

Codirector: Dr. Juan Manuel Trujillo Sevilla

Departamento de Ingeniería Industrial
Universidad de La Laguna

Memoria para la obtención del grado de
Doctor en Ingeniería

A día 3 de noviembre de 2020

Autorización

José Manuel Rodríguez Ramos, Doctor en Ciencias Físicas, con DNI 46441976W y como Director de la presente tesis:

AUTORIZA a D. David Carmona Ballester, Ingeniero en Electrónica, a presentar la tesis doctoral realizada bajo su supervisión.

A día 3 de noviembre de 2020 en San Cristóbal de La Laguna
Dr. Jose M. Rodríguez-Ramos

Autorización

Juan Manuel Trujillo Sevilla, Doctor en Ciencias Físicas, con DNI 78568493X y como Codirector de la presente tesis:

AUTORIZA a D. David Carmona Ballester, Ingeniero en Electrónica, a presentar la tesis doctoral realizada bajo su supervisión.

A día 3 de noviembre de 2020 en San Cristóbal de La Laguna
Dr. Juan Manuel Trujillo Sevilla

そんな毎日がキミのまわりで (*Sonna mainichi ga kimi no mawari de*)
Y rezo para que estos días
ずっとずっと続きますように... (*zutto zutto tsuzukimasu you ni ...*)
no lleguen nunca a su fin ...

Prefacio

Tras finalizar mis estudios de Grado en Ingeniería Electrónica Industrial por la Universidad de la Laguna en 2015, y ya en momento de plantear una continuación a estos, tuve la suerte de coincidir con el Dr. José Manuel Rodríguez Ramos en unas jornadas de difusión científica. A pesar de no tener un interés inicial en la investigación científica, conceptos como imagen integral y *display* 3D llamaron en aquel momento mi atención. Esto me llevó ese mismo año a cursar un Máster en dicha Universidad orientado a I+D+i.

Fue al mismo tiempo, cuando entré de forma temporal en el —actualmente desaparecido— grupo de investigación CAFADIS bajo una beca de colaboración. Los conocimientos y experiencia adquiridos durante esta época en materia de imagen plenóptica, fase de frente de onda y tratamiento computacional serían el punto de partida de una tesis doctoral que he elaborado desde el año 2016 y ha sido dirigida por el Dr. José Manuel Rodríguez Ramos y codirigida por el Dr. Juan Manuel Trujillo Sevilla; inicialmente como personal investigador de la Universidad de La Laguna, participando en múltiples proyectos relacionados con representación de imagen tridimensional, y posteriormente como Ingeniero Electrónico dentro de la empresa WOOPTIX S.L.

Esta tesis representa la culminación del esfuerzo realizado durante mi trayectoria investigadora estos últimos años, y engloba un trabajo que ha perseguido mejorar en mayor o menor medida el desempeño de sistemas de visualización —*displays*—, bien de forma computacional o modificando el *hardware* de estos. Asimismo, los resultados mostrados en este trabajo no son más que el primer paso, y servirán de apoyo para todo el trabajo posterior que como investigador continuaré realizando en los próximos años.

Agradecimientos

No puedo evitar comenzar dando las gracias a familia y amigos por el inestimable apoyo que me han brindado durante la realización de este trabajo de tesis. Sin ellos, este trabajo no tendría sentido.

Me gustaría además dar las gracias al grupo CAFADIS –actualmente Wootix– por estos años tan especiales, en los que he podido experimentar un crecimiento a nivel académico y personal que no imaginaba posible cuando comencé. En especial, a Jose Manuel Rodríguez Ramos y Juan Trujillo Sevilla, director y codirector de tesis respectivamente, que con tanta paciencia y cariño me han acompañado a lo largo de este trayecto. De ellos he adquirido un profundo interés por la investigación científica. También debo agradecer a Sergio Bonaque por las inestimables charlas y discusiones que tan positivamente han impactado en este trabajo.

Como otros tantos antes de mí, he podido experimentar el turbulento comportamiento al que uno se ve sujeto una vez dentro de la investigación científica. Tanto por los fracasos –muchas veces motivados por fallos de procedimiento o razón–, como los éxitos –gobernados por la dedicación y siempre amigos inseparables de la serendipia–, es indudable el valor y la importancia de una buena compañía. Por ello, doy gracias a todos aquellos que se han unido a mí durante esta travesía y que han hecho posible este trabajo.

Muchas gracias

Resumen

La representación de imagen en *displays* ostenta en la actualidad una posición privilegiada en la literatura científica. Esto es debido primordialmente al interés comercial que estos sistemas despiertan y su aplicación a diferentes soportes, como son la realidad virtual o los terminales móviles. La obtención de métodos que permitan superar las barreras técnicas de los sistemas de visualización potencialmente puede reducir los costes de manufactura de estos, resultando crucial su investigación.

Por tanto, en esta tesis se abordan diferentes cuestiones relacionadas con el procesamiento digital de imágenes condicionadas a la representación y visualización de las mismas. En primer lugar, se ha estudiado la representación y análisis de la función de *light-field* sobre *displays* automultiscópicos, esto es, sistemas de visualización capaces de ofrecer a un observador la sensación de profundidad sin necesidad de medios externos como son, por ejemplo, las gafas. Esto se consigue presentando a cada ojo un conjunto de imágenes similares, pero diferenciables en tanto a la disparidad espacial obtenida en relación a la distancia real de los objetos presentes en la escena.

Específicamente, esta parte del trabajo se ha elaborado sobre el concepto de *display* multicapa, siendo éste, esencialmente, un dispositivo compuesto por múltiples capas en las que es posible representar diferentes imágenes, y cuya superposición permite proyectar direccionalmente información que cumpla los requisitos anteriormente descritos. La contribución en este sentido es la mejora de los algoritmos ya existentes para la síntesis de la función de *light-field*. Los métodos numéricos comúnmente empleados son muy sensibles al mal condicionamiento del problema, y por tanto, dependen enormemente de la solución inicial del problema. Se analizan diferentes métodos, y se presentan dos alternativas que mejoran los resultados previos sobre el método *weighted non-negative matrix factorization* (WNMF). Finalmente se ha elaborado un modelo funcional del dispositivo compuesto por tres capas.

De forma paralela, se ha estudiado la viabilidad de este tipo de sistemas con el objetivo de discriminar observadores, esto es, evitar que el contenido representado pueda ser percibido en ciertas direcciones. Asimismo, se ha propuesto una metodología que permite optimizar el modelo, reduciendo enormemente los requisitos físicos del dispositivo y mejorando la calidad de las imágenes representadas, tanto para los observadores autorizados como para los que no.

La representación de imagen tridimensional y los métodos empleados en el transcurso de este trabajo pueden ser entendidos como una tomografía de ángulo limitado. Es por ello que, de forma alternativa, se ha aplicado el *weighted non-negative tensor factorization* (WNTF),

como generalización del algoritmo WNMF, en tomografía de la atmósfera en contraparte a otros métodos más tradicionales como el *randomized Kaczmarz*.

Asimismo, en esta tesis, y en la línea de *displays* tradicionales, se ha propuesto y estudiado un método que permite mejorar la resolución de las imágenes mostradas por encima del límite que el ancho de banda del dispositivo determina. Esto se ha conseguido aplicando un desenfoque artificial al sistema observado y pre-compensando la imagen representada teniendo en cuenta el diámetro del círculo de confusión. Su aplicación fue demostrada a nivel experimental tanto en un sistema de realidad virtual, como en un proyector comercial.

Finalmente, se presenta una técnica basada en la medida y caracterización de un *display* con microlentes para mejorar la percepción de un observador no emétrope. Esto consiste en evaluar la respuesta al impulso generada por cada uno de los píxeles del sistema, para posteriormente ser almacenada y empleada para pre-compensar la imagen mostrada. Se realizaron tanto pruebas en simulación, como a nivel experimental donde se demuestra la viabilidad de la técnica. Asimismo, se evalúa este método con el fin de introducir enfoque variable sobre la misma pantalla.

Índice general

Índice de figuras	XIX
Índice de cuadros	XXIII
Glosario	XXV
1. Introducción	1
1.1. Definición de sistema de visualización	1
1.2. Hipótesis de partida	2
1.3. Objetivos generales	3
1.4. Organización de la memoria	3
2. Sistema Multicapa	5
2.1. Light-field y representación	6
2.1.1. Imagen integral	7
2.1.2. Limitaciones de la imagen integral	8
2.2. Sistemas multicapa	10
2.2.1. Definición de la función de <i>light-field en sistema multicapa</i>	10
2.2.2. Sistemas multicapa y función de modulación	11
2.3. Síntesis de capas	12
2.3.1. Mínimos cuadrados	13
2.3.2. Síntesis de capas por factorización de matriz o tensor	15
2.4. Solución inicial para dos capas con WNMF	19
2.4.1. Inicialización basada en profundidad	19
2.4.2. Simulaciones	21
2.5. Optimización de reglas multiplicativas de WNMF	26
2.6. Resultados e implementación en modelo experimental	29
2.6.1. Modelo propuesto	29
2.6.2. Calibración del sistema	32
2.6.3. Capturas del sistema	34
2.7. Sumario y Conclusiones	36

3. Sistema multicapa para la prevención de <i>shoulder-surfing</i>	39
3.1. Protección de privacidad y <i>Shoulder-surfing</i>	40
3.2. Sistema multicapa para privacidad	41
3.2.1. Criterio de privacidad	42
3.2.2. Profundidad de campo del sistema de dos capas	43
3.3. Evaluación de la técnica	44
3.3.1. Recuperación del filtro impuesto como función de mérito	45
3.4. Resultados experimentales	45
3.4.1. Interdependencia entre filtro y número de <i>frames</i>	47
3.4.2. Implementación sobre modelo	51
3.5. Sumario y Conclusiones	51
4. Adaptación de WNTF a tomografía atmosférica	55
4.1. Marco teórico	56
4.2. Recuperación de fase	57
4.2.1. Kaczmarz iterativo	57
4.2.2. WNTF y Gradientes de fase	59
4.2.3. Error en la descomposición multimodal	61
4.3. Evaluación de la técnica	61
4.3.1. Modelo atmosférico	62
4.3.2. Sistema de estudio	62
4.3.3. Razón de Strehl	63
4.4. Resultados	64
4.5. Sumario y Conclusiones	67
5. Mejora de resolución por introducción de desenfoque	71
5.1. Antecedentes	72
5.2. Estimación de la función de solapamiento	73
5.2.1. Solapamiento puntual como caso de estudio	73
5.2.2. Solapamiento continuo y generalización	75
5.2.3. Puntos de operación	76
5.3. Consideraciones y limitaciones de la técnica	78
5.3.1. Relación contraste-resolución	79
5.3.2. Sistemas de visualización en color y estructuras alternativas	79
5.3.3. Círculo de confusión	80
5.4. Deconvolución como pre-compensado de la imagen	81
5.4.1. Inversión de la PSF	81
5.4.2. Inversión multimodal	82
5.5. Simulaciones	84
5.5.1. Pre-compensado simple o con desenfoque único	84
5.5.2. Pre-compensado multimodal	90

5.6.	Implementación sobre sistemas de visualización	94
5.6.1.	Caso de estudio: Proyector	94
5.6.2.	Caso de estudio: Sistema de realidad virtual	97
5.7.	Sumario y Conclusiones	99
6.	Caracterización de respuesta al impulso de sistema integral	103
6.1.	Consideraciones iniciales y trabajos previos	103
6.2.	Respuesta impulsiva (IR) del sistema de imagen integral	104
6.2.1.	Respuesta en intensidad	104
6.2.2.	Multiplicidad de la respuesta y distancia de muestreo	105
6.3.	Deconvolución basada en IR	107
6.3.1.	Adaptación de la IR como PSF del ojo humano	108
6.3.2.	Sistema multifoco	108
6.4.	Implementación y resultados experimentales	109
6.4.1.	Generación de IR	109
6.4.2.	Corrección de defectos visuales	111
6.4.3.	Mejora de la resolución	112
6.4.4.	Sistema multifocal	112
6.5.	Sumario y conclusiones	115
7.	Conclusiones y trabajo futuro	119
7.1.	Sumario	119
7.2.	Publicaciones	120
7.2.1.	Publicaciones relevantes	120
7.2.2.	Patentes	121
7.2.3.	Otras publicaciones	121
	Bibliografía	123
	Apéndice A. Pseudocódigo para las reglas multiplicativas de WNMF	135
A.1.	Generación de matriz de proyección del <i>light-field</i>	136
A.2.	Producto capa-matriz de <i>light-field</i>	138
A.3.	Producto entre capas	139
A.4.	Reglas multiplicativas para dos capas	140
A.5.	Síntesis	142
	Apéndice B. Correlación cruzada para tomografía atmosférica entre <i>Kaczmarz iterativo</i> y WNTF	143

Índice de figuras

2.1. Parametrización del <i>light-field</i> como función de dos planos	7
2.2. Sistema de imagen integral	9
2.3. Esquema del sistema multicapa para n moduladores	11
2.4. Representación de la transformada de Radon directa	14
2.5. Representación esquemática de los métodos WNMF y WNTF	17
2.6. Tendencia del valor PSNR con respecto al número de iteraciones y la solución inicial	23
2.7. Tendencia del valor PSNR para las primeras 60 iteraciones de las reglas multiplicativas.	23
2.8. Diagrama de barras con el valor promedio de PSNR para cada método	25
2.9. Vista central de la escena <i>Dice</i> por varios métodos	25
2.10. Representación del <i>light-field</i> sobre una matriz dispersa para el caso de dos pantallas	27
2.11. Reacondicionamiento de la matriz dispersa en \mathbf{L}'	28
2.12. Modelo del sistema multicapa propuesto	30
2.13. Modelo experimental del sistema multicapa ajustable para tres pantallas.	31
2.14. Comparativa entre el sistema sin difusor y con difusor	31
2.15. Orientación arbitraria de planos para el sistema de tres capas propuesto	32
2.16. Patrón de calibración	33
2.17. Montaje para la calibración del dispositivo multicapa	34
2.18. Comparación antes y después de la calibración	35
2.19. Comparación antes y después de la calibración con corrección de escala	36
2.20. Vista lateral y central obtenidas sobre el sistema diseñado para <i>light-field</i> sintéticos	37
3.1. Doble Retrato del Rey Frederick IV y la Reina Louise	41
3.2. Referencia del sistema multicapa aplicado a privacidad	42
3.3. Límite superior en la profundidad de campo obtenida para un sistema de dos capas	43
3.4. <i>Light-field</i> de entrada para caso correlado y no correlado respectivamente	45
3.5. Comparativa del espectro de las vistas recuperadas para bloqueada y no bloqueada	46

3.6.	Promedio radial de la magnitud del filtro recuperado	47
3.7.	Resultado obtenido en el caso LF	48
3.8.	Tendencia del PSNR para ambas técnicas propuestas tras incrementar el rango de la descomposición.	48
3.9.	Tendencia del PSNR para el caso LF	50
3.10.	Captura del sistema implementado para privacidad	52
4.1.	Esquema simplificado de la tomografía atmosférica	58
4.2.	Descomposición de un tensor de tres dimensionaes en sus componentes para un rango M arbitrario.	60
4.3.	Distribuciones de LGS	62
4.4.	Comparativa entre las fases recuperadas	65
4.5.	Promedio radial de la razón de Strehl para los métodos propuestos	66
4.6.	Promedio radial de la razón de Strehl a lo largo del campo de visión obtenido a partir de 1000 simulaciones de atmósfera diferentes	69
4.7.	Relación de Strehl radial promedio a lo largo del campo de visión – tres capas	70
5.1.	Representación teórica de la función de intensidad I_c en SP	74
5.2.	Representación teórica de la función de intensidad I_c en SC	75
5.3.	Representación visual de la estimación de los rangos de operación del SP	77
5.4.	Propagación <i>Fresnel</i> de fuente puntual	81
5.5.	Esquema de lente líquida EL10-30d el fabricante Optotune	83
5.6.	Caso de SP ponderando la imagen resultado	86
5.7.	Tendencia de PSNR y SSIM frente a tamaño de círculo de confusión	87
5.8.	Comparativa para diferentes rangos de desenfoque	88
5.9.	Resultados obtenidos en simulación para la resolución base	89
5.10.	MTF del sistema para la técnica propuesta	91
5.11.	Recuperación para el pre-compensado multimodo en el caso del SC	92
5.12.	Descomposición en instantes de tiempo del método multimodo para el caso del SC	92
5.13.	Comparativa entre diferentes rangos de desenfoque para la implementación multimodo	93
5.14.	Montaje para la evaluación de la técnica sobre proyector	95
5.15.	Imágenes capturadas para el caso del proyector con SP	96
5.16.	Comparación en proyector para el caso SC	97
5.17.	Comparación sobre proyector del pre-compensado multimodal	98
5.18.	Montaje del sistema para realidad virtual	99
5.19.	Resultados obtenidos para el caso de SP en el montaje de realidad virtual	100
5.20.	Resultados obtenidos para el caso de SC cuadrado en el montaje de realidad virtual	101

6.1. Esquema del método <i>captura-representación</i>	105
6.2. Representación de la adquisición de múltiples respuestas en el sistema	106
6.3. Esquema del método <i>captura-representación</i>	109
6.4. Montaje experimental del sistema <i>captura-representación</i>	110
6.5. Respuesta al impulso obtenida para una rejilla de 4x4 pixeles para diferentes desenfoques	111
6.6. Imágenes obtenidas por el sistema de adquisición tras la minimización	113
6.7. Comparativa entre sistema corregido y sin corregir usando MLA	114
6.8. Carta USAF-1951 normalizada indicando el límite en resolución del sistema para cada caso	115
6.9. Resultados obtenidos en simulación para la resolución base	116
6.10. Matriz de correlación para diferentes valores de desenfoque	117
B.1. Relación estadística entre los métodos propuestos	144

Índice de cuadros

2.1. Métodos de inicialización y propuesta.	21
2.2. PSNR en dB obtenidos para dos capas y seis instantes de tiempo condicionado al método de inicialización	24
2.3. Parámetros extraídos de la matriz de homografía.	35
4.1. Parametrización del modelo atmosférico	63
5.1. Primeros cinco rangos de solapamientos en la técnica de SP obtenidos a partir de la Ecuación 5.5.	77
5.2. Valores de PSNR y SSIM del punto de operación óptimo	86
5.3. Rangos de desenfoque para el pre-compensado multimodo en el caso SC . . .	90

Glosario

Acrónimos / Abreviaturas

ALS	Mínimos cuadrados alternos – <i>Alternate least squares</i>
AO	Óptica adaptativa – <i>Adaptive optics</i>
ART	Técnica de reconstrucción algebraica – <i>Algebraic reconstruction technique</i>
AT	Tomografía atmosférica – <i>Atmospheric tomography</i>
CPL	Ciclos por letra – <i>Cycles per letter</i>
CT	Tomografía computerizada – <i>Computerized tomography</i>
DM	Mapa de profundidad – <i>Depth map</i>
ELT	Telescopios extremadamente largos – <i>Extremely large telescopes</i>
FPGA	Matriz de puertas lógicas programable en campo – <i>Field-programmable gate array</i>
GPU	Unidad de procesamiento gráfico – <i>Graphic processing unit</i>
HVS	Human visual system (Sistema visual humano)
InI	Imagen integral – <i>Integral imaging</i>
LC	Cristal líquido – <i>Lyquid crystal</i>
LCD	Pantalla de cristal líquido – <i>Lyquid crystal display</i>
LED	Diodo emisor de luz – <i>Light emissor diode</i>
LF	Letra filtrada
LGS	Estrella guía láser – <i>Laser guide star</i>
MLA	Rejilla de microlentes – <i>Micro lens array</i>
MSE	Error cuadrático medio – <i>Mean square error</i>
NC	No correlado

- NMF Factorización no-negativa de matriz – *Nonnegative matrix factorization*
- NMF Frecuencia crítica de parpadeo – *Critical flicker frequency*
- NTF Factorización no-negativa de tensor – *Nonnegative tensor factorization*
- PSNR Relación señal ruido de pico – *Peak signal to noise ratio*
- RGB Rojo, verde, azul – *Red, green, blue*
- RMS Raíz del valor medio al cuadrado – *Root mean square*
- SART Técnica de reconstrucción algebraica simultánea – *Simultaneous algebraic reconstruction technique*
- SLM Modulador de luz espacial – *Spatial light modulator*
- SSO Distribución en cuadrado con eje – *Square shaped on-axis*
- SS Distribución en cuadrado – *Square shaped*
- SVD Descomposición en valores singulares – *Singular value decomposition*
- TSO Distribución en triángulo con eje – *Triangle shaped on-axis*
- TS Distribución en triángulo – *Triangle shaped*
- WFAO Óptica adaptativa de campo extenso – *Wide-field adaptive optics*
- WFS Sensor de frente de onda – *Wavefront sensor*
- WNMF Factorización no-negativa ponderada de matriz – *Weighted nonnegative matrix factorization*
- WNTF Factorización no-negativa ponderada de tensor – *Weighted nonnegative tensor factorization*

Capítulo 1

Introducción

Al modo de muchas bellezas naturales, las obras humanas necesitan, para no perder sus encantos, ser contempladas a distancia. El análisis es el microscopio que nos aproxima al objeto y nos muestra la grosera urdidumbre del tapiz, disíbase la ilusión cuando salta a los ojos lo artificioso del bordado y los defectos del dibujo.

– Santiago Ramón y Cajal

El trabajo presentado en esta tesis engloba múltiples técnicas desarrolladas en el marco de los sistemas de visualización actuales. Con ello, se busca mejorar el desempeño de estos, bien de forma computacional y haciendo uso de técnicas basadas en el procesamiento de imagen, bien modificando *hardware* ya existente.

Las técnicas presentadas en esta memoria no se limitan a *displays* tradicionales, sino que barren un espectro más amplio, incluyendo sistemas de imagen autoestereoscópica, esto es, que permiten la percepción de profundidad sin necesidad de gafas u otro artilugio periférico.

1.1. Definición de sistema de visualización

Hace más de 60.000 años, el hombre *neandertal* realizaba las primeras representaciones pictóricas (Kim et al., 2007). Ya en estos momentos, buscaban mostrar realidades cotidianas en su mayoría. Hasta la actualidad, el ser humano ha continuado evolucionando en este sentido, buscando nuevas técnicas y medios que permitieran representar diferentes conceptos o ideas, bien desde un punto artístico o puramente ilustrativo. Esta necesidad comunicativa es la que ha propiciado dicha evolución a lo largo de la historia, materializando estos esfuerzos en todo tipo de soportes de visualización.

Debe entenderse como sistema o soporte de visualización aquel medio que permite representar una idea de manera gráfica o visual. Es posible englobar dentro de esta definición una

fotografía, un cartel, o un libro. No obstante, es un error común asumir que a esta categoría pertenecen solo los soportes electrónicos, a pesar de tener una mayor relevancia en la literatura científica.

Si bien lo anterior es cierto, la aparición de los primeros sistemas de visualización electrónicos ha dado lugar a cambios de paradigma en los métodos de representación de la información. En la actualidad, el uso de estos sistemas se ha convertido en una cuestión indispensable en el desarrollo de nuestras vidas. Desde las pantallas de nuestros ordenadores, hasta pequeños accesorios como *smartwatches* o incluso paneles publicitarios, vemos que su presencia aumenta con cada nuevo avance tecnológico. Es por ello que la investigación en sistemas de visualización (comúnmente denominados *displays* en la literatura) ha aumentado de manera significativa en las últimas décadas. Asimismo, las mejoras en canales de comunicación, microelectrónica y semiconductores continúan abriendo nuevas vías de estudio anteriormente inabarcables.

En el marco de los sistemas de representación electrónicos –primero analógicos, y posteriormente digitales– ha existido siempre un vínculo muy estrecho con la adquisición de imagen. No en vano, la generación del contenido a representar en estos soportes provenía en su mayoría de grabaciones realizadas en diferentes ámbitos, bien televisivos o cinematográficos. En las últimas décadas, hemos sido testigos de múltiples hitos tecnológicos que han propiciado nuevas alternativas a como crear este contenido. Entre ellos, con la aparición de sistemas de cómputo de mayores prestaciones, hemos podido asistir al auge de la generación de gráficos por computadora.

A pesar de que este trabajo se centra principalmente en la representación de imagen, resulta interesante tomar como referencia los sistemas de adquisición actuales, pues son realidades mutuamente entrelazadas, donde las limitaciones, necesidades y avances técnicos vienen condicionados en su mayoría por dicha mutualidad.

1.2. Hipótesis de partida

Como en otros campos de investigación, la representación de imagen y su análisis toma como punto de partida diferentes términos. Específicamente, la digitalización de estos medios da lugar a un conjunto de limitaciones técnicas, surgidas entre otros, de la propia manufactura de estos dispositivos. Expandiendo sobre lo anterior, podemos concretamente hablar del término calidad como una métrica generalista en torno a estos aspectos. No obstante, este nuevo término de calidad se apoya en el sistema visual humano –o *Human Visual System* en inglés (HVS)–, buscando una representación que tome como referencia las cualidades de éste, trasladando de manera fiel la realidad observable a estos dispositivos. Por ello, se plantea un análisis de conjunto y se proponen mejoras sobre la técnica actual tanto en sistemas de visualización tridimensionales como en bidimensionales.

1.3. Objetivos generales

Los objetivos de este trabajo de tesis se centran en la mejora de la calidad de los sistemas de visualización por medio de la implementación de nuevas técnicas que contribuyan a la resolución de los problemas clásicos presentes en los sistemas actuales. Con esto, se pretende incrementar la calidad en términos del observador por diferentes medios computacionales basados en las implementaciones existentes en la bibliografía. No obstante, y de manera paralela, también se proponen nuevas aplicaciones para técnicas que, si bien ya están consolidadas en la literatura, pueden ser transferidas a otros campos donde se apliquen unos principios similares.

1.4. Organización de la memoria

La presente memoria está organizada en dos bloques de contenido esencialmente diferenciables. El primer Bloque está formado por los Capítulos 2, 3 y 4, y se centra en la mejora de la técnica en sistemas de visualización tridimensional. En el Capítulo 2 se realiza una descripción detallada del trabajo realizado sobre el sistema multicapa para la generación de imagen tridimensional. Se abordará inicialmente desde un marco esencialmente teórico, indicando las técnicas más comúnmente empleadas a la hora de representar imagen. De igual manera, se tratará de forma extensa cómo es posible la mejora de las técnicas actuales atendiendo al punto de partida computacional que se plantea a dichos algoritmos. Finalmente se mostrará un modelo diseñado y ensamblado en laboratorio, además de una breve disertación sobre los métodos de calibración de éste. En el Capítulo 3 se estudiará de forma alternativa la aplicación del sistema multicapa como elemento protector de privacidad, esto es, discriminar de entre un conjunto de observadores quién tiene acceso a la información representada. En el Capítulo 4 se adaptará el algoritmo *weighted non-negative tensor factorization* (WNTF) –empleado en los Capítulos previos–, al campo de la óptica adaptativa para la síntesis de gradientes de fase en un perfil atmosférico.

El segundo Bloque de esta memoria abarca los Capítulos 5 y 6. El contenido de los mismos está enfocado principalmente a algoritmos de aplicación directa a sistemas de visualización bidimensionales. En el Capítulo 5 se presenta una técnica que permite mejorar la resolución de este tipo de sistemas induciendo un desenfoque artificial sobre el contenido representado, y pre-compensando dicha imagen de acuerdo a éste para superar el límite de resolución lateral o espacial. Finalmente, en el Capítulo 6 se introduce la caracterización de la respuesta impulsiva de un sistema de manera más general, y se hace uso de ella para representar una versión corregida frente a un observador, derivándose múltiples aplicaciones de ello.

Capítulo 2

Sistema Multicapa

Para la obra científica, los medios son casi nada y el hombre casi todo.

– Santiago Ramón y Cajal

Los contenidos presentados en este Capítulo sirven como antesala del primer bloque de esta memoria. Se introduce el uso del sistema multicapa como alternativa a sistemas de representación tridimensional tradicionales enmarcados en imagen integral. Asimismo, se exploran las técnicas actuales en relación a la síntesis de información tridimensional y su representación en los mismos, con especial interés en la propuesta de métodos alternativos y mejoras de estos. El trabajo realizado incluye el diseño de un prototipo funcional, desde la generación de imagen hasta su representación, la mejora de uno de estos algoritmos con base en su punto de partida, y una reordenación de este para su implementación optimizada sistemas de cómputo paralelo.

Publicaciones relacionadas

En la Sección 2.4 de este Capítulo se recogen los resultados correspondientes a la publicación:

- D. Carmona-Ballester, J. M. Trujillo-Sevilla, Sergio Bonaque-González, A. Hernández-Delgado, y J. M. Rodríguez-Ramos. Analyzing the impact of the initialization for the nonnegative tensor factorization in tensor displays (in press). *Optical Engineering*, 57 (6), 2017. doi: 10.1117/1.OE.57.6.061603

Asimismo, los resultados presentados en la Sección 2.6 se corresponden con la comunicación realizadas en el congreso *SPIE Defence+Commercial Sensing 2019*.

- David Carmona-Ballester, Viana L Guadalupe-Suárez, Juan M Trujillo-Sevilla, Sergio Bonaque-González, Ricardo Oliva-García, Jose G Marichal-Hernández, Miguel J Sicilia-

Cabrera, Lara Díaz-García, y Jose M Rodríguez-Ramos. Calibration method applied to a tunable tensor display system. In *Proc.SPIE*, volume 10997, may 2019a. URL <https://doi.org/10.1117/12.2518865>

2.1. Light-field y representación

Para definir el concepto de *light-field* –o campo de luz en español–, debemos partir antes de la idea de función plenóptica. Esta función hace referencia, siguiendo su etimología, a la representación de la luz en un sentido pleno, esto es, su procedencia con respecto a los objetos dentro de un espacio de estudio o escena (Adelson y Wang, 1992). Alternativamente, y por simplicidad, es posible entender el enunciado anterior como una función que contiene todos los rayos de luz presentes en una escena.

De forma general, y en un contexto de óptica geométrica, la unidad fundamental *portadora* de luz es el rayo¹. La proyección de uno de estos rayos sobre un plano puede definirse como $I(x, y)$, representando la intensidad total de dicha proyección en las coordenadas del plano (x, y) . Alternativamente, si obviamos la proyección sobre un plano, y consideramos un punto en el espacio como la fuente de éste, es posible definir cada rayo a partir de los ángulos acimutal y cenital de la esfera que rodea al punto respectivamente, esto es $I(\theta, \phi)$. Tiene cabida dentro de la definición condicionar esta función a otros tantos parámetros. Por ejemplo, tomando en cuenta el origen de los rayos de luz, esto es, cuya fuente se encuentra situada en el espacio en torno a un vector de coordenadas \vec{r} . Es posible también considerar la longitud de onda de la luz emitida, o en otras palabras, el color si estamos en el espectro visible. Finalmente, la inclusión de la dimensión temporal cabe en dicha definición.

Tras el enunciado anterior surge una pregunta esencial: "*¿Cómo podemos trasladar esta función a un conjunto de datos que exprese de manera simplificada la información tridimensional de una escena?*". Aún ignorando la dimensión temporal y la cromaticidad –o color–, para un volumen tendríamos una función 5-dimensional; $I(r_x, r_y, r_z, \theta, \phi)$. Siguiendo la propuesta de Levoy y Hanrahan (1996), y de manera similar a Gortler et al. (1996), una forma de reparametrizar la función plenóptica pasa por considerar que los puntos de vista 2-dimensionales observables a partir de una escena forman parte de una función 4-dimensional. Más concretamente, proponen definir cada uno de los rayos de luz procedentes de esta a partir de la intersección de dos planos. En la Figura 2.1 se muestra un esquema de la parametrización del *light-field* a partir de dos planos².

¹En óptica encontramos diferentes aproximaciones al análisis comportamental de la luz. Por un lado, la aproximación paraxial considera la propagación de la luz en forma de diferentes rayos desde una fuente, siendo estos los encargados de formar imagen en un sistema óptico. No obstante, otras aproximaciones tienen base en la naturaleza ondulatoria de la luz, y toman en cuenta un conjunto de efectos que desde una aproximación geométrica no pueden ser estudiadas, como son por ejemplo los fenómenos de difracción (Goodman, 2005).

²A diferencia de lo presentado en Gortler et al. (1996) y Levoy y Hanrahan (1996), se ha decidido sustituir la nomenclatura empleada en los ejes de coordenadas para mantener la coherencia con el resto de la memoria. Por ello, los ejes de coordenadas (u, v) y (s, t) en el trabajo de Levoy y Hanrahan (1996) se han sustituido por (x, y) y (u, v) respectivamente.

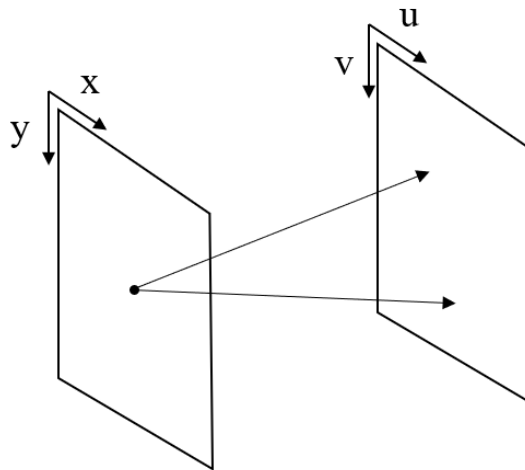


Figura 2.1 Parametrización del *light-field* como función de dos planos

2.1.1. Imagen integral

Hasta ahora nos hemos centrado en definir el concepto de *light-field*. En su mayoría, la formulación anteriormente introducida tiene históricamente origen en la fotografía, especialmente en el campo del procesamiento de imagen, donde la obtención de la función plenóptica resulta de gran valor. Es el físico Gabriel Lippman el primero en plantear un método que permite obtener dicha información (Lippmann, 1908). Esto lo consigue con la introducción de un sistema de fotografía sobre el cual se añade una rejilla de microlentes – o *microlens array* en inglés (MLA)– lentes de tamaño muy pequeño– delante de la película fotosensible. De esta forma, al tomar la imagen, detrás de cada microlente se formaría una pequeña imagen compuesta por fragmentos de los diferentes puntos de vista de la escena. La capacidad de discernir el origen e inclinación de los rayos de luz que forman cada parte de esta imagen permitió extraer la información tridimensional de la escena. Posteriormente, Herbert E. Ives (1930) continuaría evolucionando el concepto, añadiendo una lente sobre el sistema, y sustituyendo las microlentes por una rejilla sobre la que se proyectaría la información de manera similar. Este dispositivo se asemejaría bastante al actual sistema plenóptico.

A partir de esta información comienza a desarrollarse toda una teoría basada en el tratamiento computacional del *light-field*, dando lugar a nuevos operadores de reconstrucción que permiten trasladar la carga de información contenida en la función plenóptica a diferentes espacios. Ren Ng. (2005) presentaba de esta manera el operador fotográfico, demostrando como a partir de la información extraída del *light-field* es posible generar diferentes desenfoques de una escena. Este efecto se denomina apertura sintética. De manera alternativa, ponía de manifiesto que este mismo operador existe en el sentido inverso, dando lugar a un conjunto de vistas pertenecientes al *light-field*. No en vano, esta definición toma una relevancia elevada en el campo de los sistemas de visualización. El replanteamiento de la información tridimensional de una escena en forma de *light-field* presenta un enfoque alternativo a su análisis.

El objetivo de un sistema de visualización tridimensional radica en la correcta representación de la función del *light-field*. Uno de los primeros trabajos en la materia es realizado por Frederic Ives (1903). Él presentaba un sistema sobre el que era posible percibir una imagen tridimensional estática. Dicho efecto se conseguía por medio de unas barras paralelas situadas sobre una imagen, bloqueando de forma selectiva la información representada para cada ojo de un observador. Adecuando el contenido posterior a esta rejilla, el par de imágenes observables sugería un cierto efecto de profundidad. El trabajo de Ives constituye el primer sistema de imagen tridimensional sin necesidad de gafas.

No obstante, este trabajo da lugar a lo que él mismo denomina como un *estereograma*, es decir, la representación de dos imágenes diferentes. En su trabajo posterior, Lippman demuestra que, a partir de la información obtenida por medio de la cámara plenóptica, y realizando la operación inversa, esto es, superponiendo una rejilla de microlentes sobre la imagen resultante, un observador sería capaz de percibir todos los puntos de vista de la escena (Lippmann, 1908). Este sistema, que se ha continuado aplicando hasta la actualidad, fue denominado imagen integral (InI), debido a que representaba la totalidad de la escena. Un requisito indispensable de este sistema, que posteriormente sería planteado por Ives (1930), consistiría en que la relación de focales entre la lente principal y la rejilla debía ser constante con el fin de garantizar la no superposición de las imágenes elementales formadas tras esta última.

2.1.2. Limitaciones de la imagen integral

Un sistema de imagen integral está sujeto a una limitación severa. Volvamos por un momento al sistema planteado por Lippmann (1908). Esta metodología plantea mostrar selectivamente la información representada sobre una superficie, de tal forma que es posible observar por cada ojo un objeto desde dos perspectivas diferentes. La disparidad espacial percibida –o el desplazamiento relativo entre los objetos de ambas imágenes– es posteriormente analizada por diferentes mecanismos del cerebro en lo que conocemos por *estereopsis* determinando la profundidad de la escena (Deangelis, 2000)³. Debido a ello, el espacio efectivo de representación en el soporte de la imagen queda reducido de manera efectiva a la mitad. De la misma manera, en un teórico sistema en el que tuvieramos un conjunto de tres perspectivas para un soporte de las mismas dimensiones, estaríamos dividiendo el espacio efectivo en ese mismo número de secciones. Este efecto es conocido como la relación espacio-angular del sistema, y de manera simplificada, determina un intercambio entre el detalle de menor tamaño representable sobre el dispositivo y la profundidad aparente (Chai et al., 2000; Zwicker et al., 2006).

En los soportes digitales actuales esto se convierte en un gran inconveniente. De manera ilustrativa, se muestra en la Figura 2.2 un sistema de imagen integral clásico reducido a dos dimensiones. Como se puede observar, cada par de puntos (x, u) da lugar a un rayo, y estos a su vez, coinciden en el espacio formando cada uno de los puntos que componen la imagen

³La percepción de profundidad no está limitada a análisis de la disparidad presente en las imágenes formadas en la retina del observador. Existen diferentes pistas visuales que ayudan a estimar esta profundidad. Entre ellas, cabe destacar la perspectiva, las pistas focales –esto es el desenfoque de los objetos en relación a su distancia con el observador–, el desplazamiento relativo en relación al movimiento y oclusiones ocasionadas por los objetos de esta.

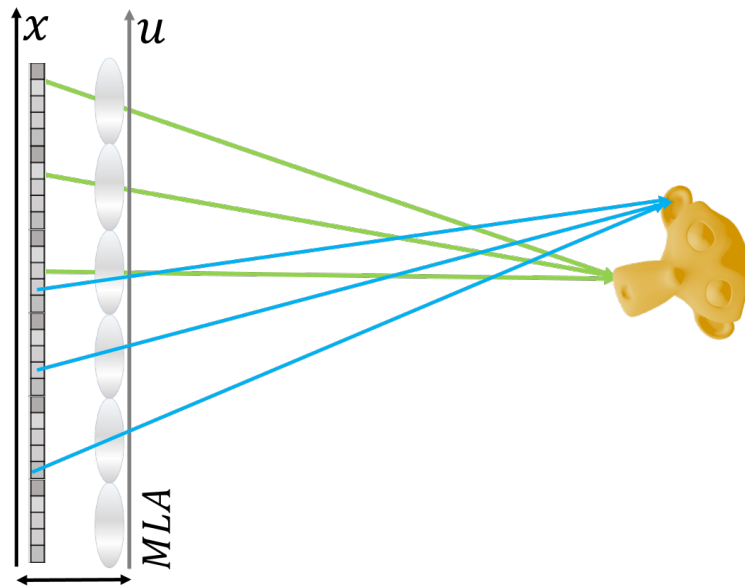


Figura 2.2 Sistema de imagen integral compuesto por una pantalla situada en el eje x y a una distancia de esta superpuesto un MLA en el eje u . La proyección de rayos desde la pantalla a través del MLA da lugar a un volumen situado a una cierta distancia del dispositivo.

tridimensional. Todos estos puntos de vista observables parten de la proyección de los píxeles que subtiende cada lente, de tal forma que, siguiendo el razonamiento anterior, la limitación en resolución aumenta a razón del número de píxeles por lente. Una de las líneas de investigación más prominentes en el campo de la imagen integral se centra en la mejora de esta relación por diferentes técnicas, abarcando metodologías basadas en modificaciones del sistema o técnicas esencialmente computacionales (Xiao et al., 2013).

Por ejemplo, algunos trabajos como los realizados por Kim et al. (2007) y Jung et al. (2012) sustituyen el MLA por una pantalla LCD, donde los propios píxeles actúan como pequeños orificios, y por tanto se comportan de manera similar a una *cámara oscura* o *pinhole*. Esto permite controlar electrónicamente la posición del orificio temporalmente, y por tanto, mejorar la resolución alcanzable por el sistema. Otros trabajos plantean técnicas basadas en la localización del observador, dando especial prioridad o peso a aquellas vistas que son potencialmente percibidas por este (Dorado et al., 2018; Shen et al., 2016, 2017). Si bien este planteamiento no busca *a priori* resolver la limitación en resolución espacial, permite incrementar el ángulo de visión de estos dispositivos, y por tanto, mejora la relación espacio-angular del dispositivo.

2.2. Sistemas multicapa

Uno de los intereses principales de este trabajo de tesis se encuentra en el uso de los sistemas multicapa como alternativa en crecimiento frente a los tradicionales de InI. Esto viene principalmente dado por las limitaciones intrínsecas de la técnica.

Un sistema multicapa no se aleja del concepto de InI en tanto a que su objetivo es el de la representación de un *light-field* conocido. Sin embargo, la diferencia esencial es la forma en que se obtiene la información para representarlo. En el caso de la InI es necesario obtener un conjunto de imágenes elementales, de tal forma que, al ser proyectadas a través de una rejilla de microlentes o *pinhole*, se obtenga una reconstrucción de la información tridimensional. Como ya se indicó anteriormente, a esta técnica subyacen una serie de desventajas fundamentales, entre ellas, la relación espacio-angular del sistema.

La alternativa multicapa fue presentada por primera vez por Gotoda (2010), si bien el concepto de superponer láminas para conseguir un efecto tridimensional es muy anterior. Más específicamente, lo que se presentó en dicho trabajo es la formulación que permite, en menor o mayor medida, sintetizar un *light-field* en un conjunto de pantallas, tales que, al ser superpuestas, se pueda observar una escena tridimensional previamente conocida. Como en el caso de imagen integral, el sistema multicapa parte de la definición del *light-field* en dos planos (Gortler et al., 1996; Levoy y Hanrahan, 1996).

2.2.1. Definición de la función de *light-field* en sistema multicapa

El sistema multicapa presenta una estructura compuesta por un número arbitrario de capas sobre las cuales es posible controlar la cantidad de luz que pasa por unidad de píxel. Si se retoma la definición de *light-field* propuesta anteriormente, se puede observar que la morfología del sistema multicapa es directamente una traslación de esta idea sobre un dispositivo físico (Gotoda, 2010). Por ello, la representación del *light-field* con base en dicho sistema resulta inmediata. Definamos por tanto el *light-field* como $l(x, u)$. De la misma forma, sea $f_n(x)$ la función que modula la intensidad de un rayo de luz al atravesar la superficie de un plano n en una posición arbitraria x a una distancia d_n del origen de un sistema multicapa. Siguiendo la proyección de los rayos a través del sistema, se puede definir la relación entre el *light-field* y la contribución de cada capa como sigue (Wetzstein et al., 2011):

$$l(x, u) = \prod_{n=1}^N f^{(n)}(x + (d_n/d_r)u), \quad (2.1)$$

con N siendo el número total de capas que compone el sistema, y d_r la distancia entre el eje x y u . De forma esquemática, en la Figura 2.3 se muestra una representación del sistema para un número arbitrario de capas. Obsérvese que se ha definido la relación entre *light-field* y capas como un producto. No obstante, esto no es necesariamente cierto como se verá más adelante.

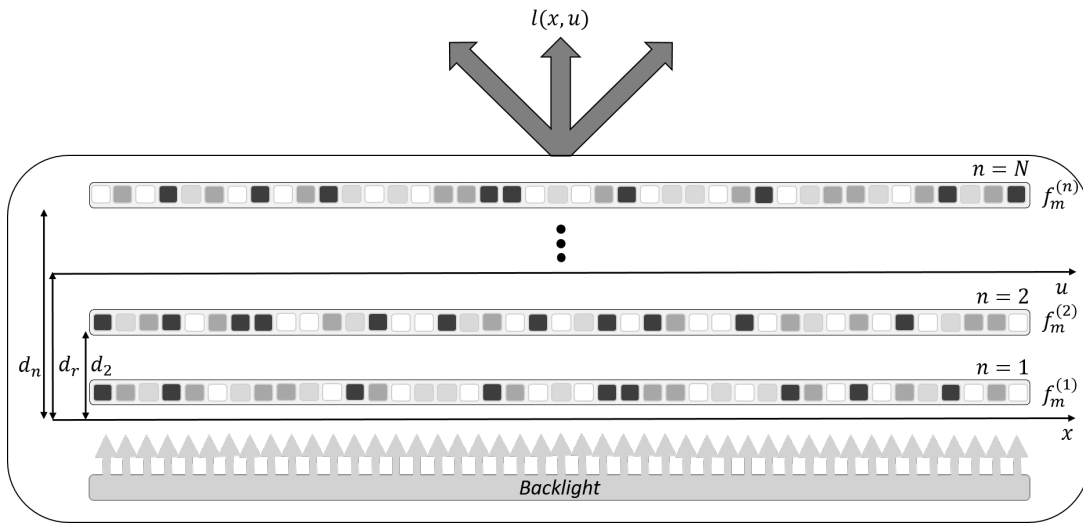


Figura 2.3 Esquema del sistema multicapa para n moduladores $f_m^{(n)}$. Una fuente de iluminación trasera o *backlight* se sitúa en uno de los extremos, produciendo de forma global un *light-field* $l(x, u)$.

2.2.2. Sistemas multicapa y función de modulación

Si bien hemos partido de una relación entre *light-field* y capas en forma de producto, esto tiene una razón puramente práctica, y se basa en lo presentado por Wetzstein et al. (2012c). De manera general, el problema de la representación tridimensional es muy similar a la reconstrucción de imagen tomográfica. La tomografía computerizada –del inglés Computerized Tomography (CT) es una técnica que busca reconstruir por medios computacionales un objeto tridimensional a partir de información bidimensional, esto es, proyecciones desde diferentes ángulos (Kak et al., 2002). Se obtiene la información bidimensional en forma de la atenuación recibida por un campo electromagnético al atravesar un cierto medio, en este caso el sujeto de estudio, para diferentes ángulos. Con esta información, es posible obtener una aproximación del objeto en tres dimensiones Herman y Meyer (1993).

Como punto de partida, y extendiendo la definición previa, Wetzstein et al. (2011) analiza el problema considerando que cada una de las capas atenúa el valor de intensidad de los rayos de luz de acuerdo a la ley de *Beer-Lambert* (Goodman, 2005):

$$I = I_0 \exp\left(-\int_c \mu(x, z) dr\right), \quad (2.2)$$

con I siendo la intensidad del haz al atravesar el medio, I_0 como la intensidad previa, $\mu(x, z)$ el volumen de atenuación. Por simplicidad, definimos un eje z como profundidad, siendo éste perpendicular al sistema. En definitiva, se entiende que la intensidad final no es más que la integral de línea en la dirección del haz de luz que atraviesa el medio de atenuación.

Todo ello permite reacondicionar la Ecuación 2.1 como la suma de las contribuciones de cada una de las atenuaciones. Nótese que en este caso la función de atenuación se descompone en el número de capas del sistema.

$$I(x, u) = \sum_{n=1}^N \ln \left[f^{(n)}(x + (d_n/d_r)u) \right] \quad (2.3)$$

Con el objetivo de simplificar, se asume que la intensidad producida por la iluminación trasera es unitaria en todas las direcciones.

Este planteamiento parte de suponer que las capas que componen el sistema tienen un comportamiento atenuante sobre la intensidad de la luz. Wetzstein et al. (2011) hace uso de transparencias o pantallas de cristal líquido (LCD), siendo el último caso especialmente interesante debido a la posibilidad de representar contenido de manera dinámica por un medio electrónico.

Concretamente, una pantalla LCD está compuesta por una lámina de cristal líquido (LC) encerrado normalmente en un soporte de vidrio y un polarizador en cada cara de la lámina desfasados 90 grados. El LC permite alterar la polarización de una onda electromagnética que lo atraviese, y es dependiente de una tensión aplicada. De esta manera, y junto a los polarizadores es posible por medio de una fuente de iluminación trasera y de forma electrónica, atenuar la luz procedente de la fuente trasera, y por tanto, representar una imagen.

No obstante, una limitación a la superposición de múltiples láminas de LC es la atenuación pasiva que introduce cada uno de los polarizadores al sistema, resultando en el oscurecimiento de la imagen, y en definitiva, siendo ópticamente ineficiente. Por ello, Lanman et al. (2011) analiza el mismo problema pero retirando todos los polarizadores del sistema excepto dos; el frontal y el trasero del dispositivo. De esta manera, es posible estudiar la polarización introducida por cada una de las capas del sistema de forma individual, en lugar de considerar atenuaciones. En este caso, y siguiendo la ley de *Mallus*, la intensidad al atravesar el sistema se define como:

$$I = I_0 \sin^2 \theta, \quad (2.4)$$

siendo θ el ángulo de polarización. De una forma similar al caso de atenuación, es posible representar el *light-field* reformulando la Ecuación 2.1:

$$I(x, u) = \sin^2 \left[\sum_{n=1}^N \gamma^{(n)}(x + (d_n/d_r)u) \right], \quad (2.5)$$

con $\gamma^{(n)}$ siendo la rotación al campo aportada de manera global por cada una de las capas.

2.3. Síntesis de capas

Como se ha visto anteriormente, es posible reformular el problema de la generación de *light-field* en un sistema con base en las aportaciones de cada una de las capas de este, bien desde el punto de vista de la atenuación introducida o como rotaciones del campo electromagnético

para ser posteriormente atenuadas con base en dicho valor. No obstante, resta comprender como es posible a partir de estas definiciones estimar el valor de $f^{(n)}$ para cada capa.

2.3.1. Mínimos cuadrados

Una limitación del sistema multicapa, y de forma más general de la imagen integral, radica en que el *light-field* a representar no es completo, o dicho de otra manera, es discreto y está limitado en ángulo. Esto significa que el conjunto de vistas a representar se encuentra enmarcado en un cierto ángulo y comprende una discretización, normalmente equidistante, de estos. La reconstrucción de objetos tridimensionales a partir de este conjunto de datos suele denominarse en la bibliografía como *tomografía de ángulo limitado* y es un problema mal condicionado (Bakushinsky y Goncharsky, 1994).

De forma general, y siguiendo las Ecuaciones 2.3 y 2.5, se observa que el problema puede descomponerse en un número finito de rayos atravesando un conjunto de píxeles. Por simplicidad, consideraremos que la discretización del *light-field* da lugar a un número de elementos $X \times Y \times U \times V$, coincidiendo esta nomenclatura con sus respectivas dimensiones en el caso continuo. Wetzstein et al. (2011) y Lanman et al. (2011) definen de forma matricial el siguiente sistema de ecuaciones lineales:

$$\mathbf{l} = \mathbf{A}\mathbf{f} - \mathbf{e}, \quad (2.6)$$

donde \mathbf{l} es un vector de tamaño $X \times Y \times U \times V$ que contiene todos los valores del *light-field* que se quiere sintetizar, \mathbf{f} es un vector solución $X \times Y \times N$ con los valores a representar en cada pantalla tal que $\mathbf{f} = [\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n]$ y \mathbf{A} la matriz de transferencia que relaciona los rayos proyectados a través de cada pantalla con las coordenadas del *light-field*. Asimismo, se introduce un término \mathbf{e} de error, ya que, como se indica anteriormente, se trata de un sistema mal condicionado, y por tanto no tiene necesariamente una solución exacta. Nótese que en el caso de atenuación se debe introducir un signo negativo sobre los valores de las capas obtenidas, de tal forma que se cumpla la Ecuación 2.2. Asimismo, en el caso rotacional, el *light-field* se define en función de rotaciones y no como valores de atenuación.

Este tipo de sistemas de ecuaciones suelen ser motivo de estudio en la literatura en cuanto a las aproximaciones elegidas para su resolución. En un plano teórico, y partiendo del ejemplo de la CT, es necesario contar con un conjunto denso de proyecciones para poder reconstruir el objeto tridimensional de estudio. Recordemos que estas proyecciones se obtienen al atravesar con un haz paralelo de algún tipo de radiación electromagnética el objeto en cuestión, manifestando una modificación en intensidad debido al medio. Este proceso es equivalente a la transformada de *Radón* del cuerpo y puede ser entendida como una reducción en dimensionalidad del problema. En la Figura 2.4, extraída del Capítulo 3 de Kak et al. (2002), se muestra de forma esquematizada dicha transformada.

La transformada inversa de Radon puede obtenerse a partir del Teorema de Corte de *Fourier* (o *Fourier Slice Theorem*), que postula que la transformada de Fourier de cada una de

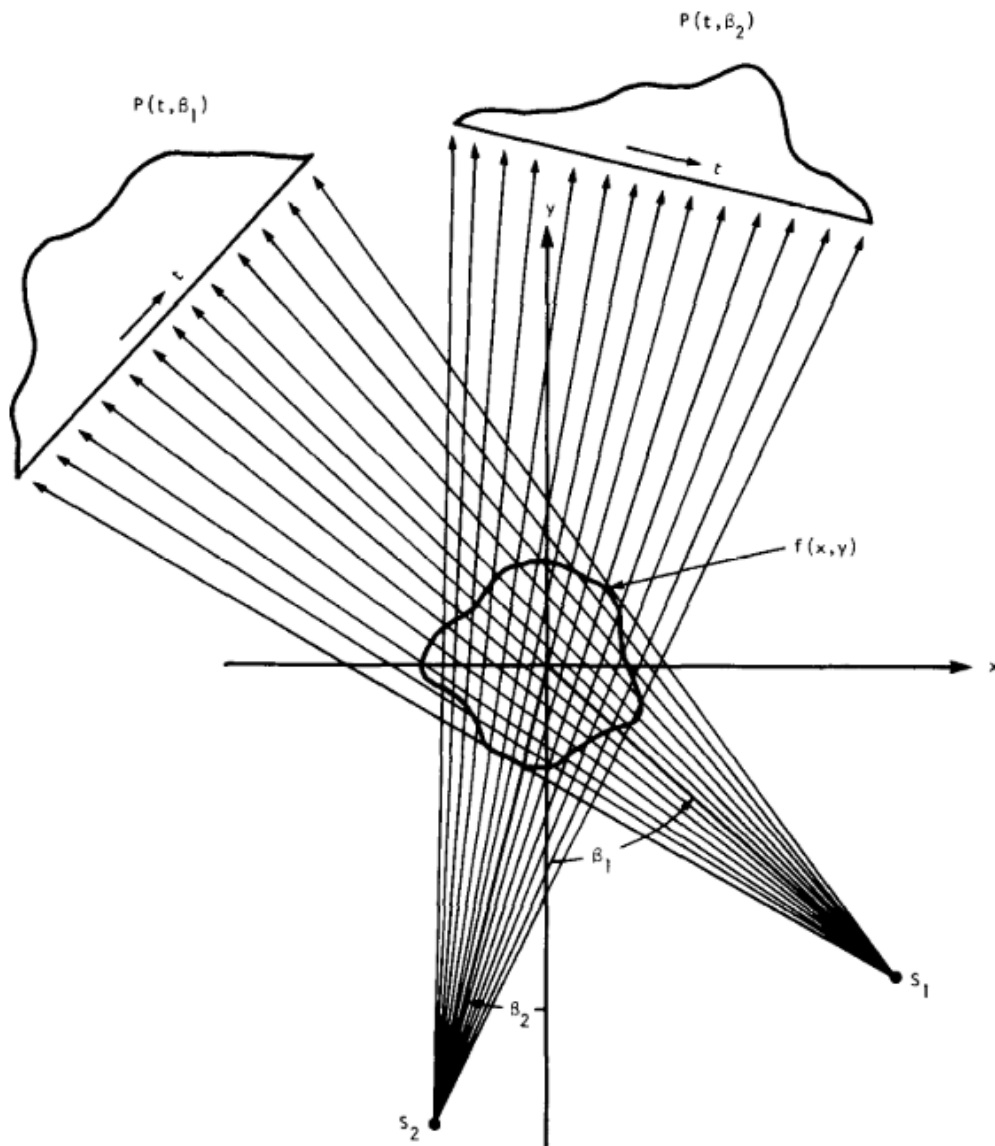


Figura 2.4 Representación de la transformada directa de Radón de una función $f(x, y)$. Imagen extraída del Capítulo 3 de Kak et al. (2002)

las proyecciones es un corte en la transformada de Fourier del objeto en estudio. Por ello, el resultado de dicha inversa es dependiente de la densidad de ángulos muestreados. Volviendo al caso del sistema multicapa, contamos con un número discreto de vistas muestradas en un cierto ángulo de visión. Por ello, es necesario la aplicación de métodos algebraicos (Gordon et al., 1970; Herman y Meyer, 1993).

Wetzstein et al. (2011) propone la minimización de la norma euclídea ($l_2 - norm$) del error entre el *light-field* original y el resultado de la proyección para las capas sintetizadas. Esto se puede expresar de la siguiente manera:

$$\arg \min_{\mathbf{f}} \|\mathbf{l} - \mathbf{A}\mathbf{f}\|^2, \quad \text{para } \mathbf{f} \geq 0 \quad (2.7)$$

Esta metodología, comúnmente denominada Técnica de Reconstrucción Algebraica (ART) puede aplicarse tanto para el caso de rotaciones como atenuaciones. No obstante, Lanman et al. (2011) propone como alternativa la Técnica de Reconstrucción Algebraica Simultánea (SART).

2.3.2. Síntesis de capas por factorización de matriz o tensor

De forma paralela, Lanman et al. (2010) proponían una metodología alternativa para la representación de imagen tridimensional por medio de un sistema multicapa. Este planteamiento nace de considerar que, comparativamente al uso de MLA, una segunda lámina LCD situada encima de una pantalla actúa de forma similar, subtendiendo un conjunto de píxeles por cada píxel en la lámina superpuesta. No obstante, esto plantea una ventaja interesante, pues teóricamente se tienen tantas lenticulas como píxeles, y por tanto se evita la limitación en resolución impuesta por la imagen integral.

Sin detrimento de lo anterior, el planteamiento técnico es el mismo, esto es la superposición de un número finito de capas sobre el cual es posible la representación de un *light-field*. En el trabajo de Lanman et al. (2010) se propone concretamente el uso de dos capas. Comparativamente, se introduce además la dimensión temporal como elemento de la síntesis de capas. Más concretamente, se añade la subdivisión de cada *light-field* en un subconjunto cuya suma devuelve este. Esto se formula de la siguiente manera:

$$\tilde{l}(x, u; N, M) = \frac{1}{M} \sum_{m=1}^M \prod_{n=1}^N f_m^{(n)}(x + (d_n/d_r)u), \quad (2.8)$$

siendo M el número de elementos que forman el subconjunto generado a partir del *light-field*. En el caso particular de Lanman et al. (2010), el número de capas se limita a $N = 2$.

En un sentido puramente físico, esta descomposición se asocia directamente a la dimensión temporal. La aproximación presentada tiene como objetivo tomar ventaja sobre la integración temporal de la retina del observador. Para un observador sin defectos visuales, el HVS tiene un tiempo de integración o umbral de detección de parpadeo –del inglés *critical flicker frequency*

(CFF)– que define el tiempo mínimo en el que es capaz de diferenciar una señal de otra (Krauzlis y Lisberger, 1994)⁴.

No obstante, comparativamente a la resolución por ART, Lanman et al. (2010) propone como alternativa la factorización de matriz no-negativa ponderada –del inglés *weighted non-negative matrix factorization*, WNMF–. Esta idea parte de la descomposición de una matriz en sus componentes principales, en este caso, dos vectores correspondientes a las capas del sistema. Esta metodología, también fuertemente ligada al ámbito de la CT, es presentada por Paatero y Tapper (1994) y posteriormente adaptada al caso del *light-field* de la siguiente manera:

$$\mathbf{L} = \mathbf{F}\mathbf{G} \quad (2.9)$$

siendo \mathbf{L} matriz a descomponer, y \mathbf{F} y \mathbf{G} los vectores o matrices resultantes de la descomposición. En la Figura 2.5 se muestra de forma esquematizada la operación de factorización. Como se observa, es posible ajustar de forma arbitraria el rango⁵ de \mathbf{F} y \mathbf{G} . Lanman et al. (2010) toma ventaja de esta representación matemática para obtener los valores en cada capa. Asimismo, el número de instantes de tiempo coincide con el rango de la factorización, asignando cada columna o *modo* M de las matrices a los instantes de tiempo del *light-field*.

Las dimensiones de la matriz \mathbf{L} concuerdan con el tamaño (XY) para un *light-field* 4-dimensional $l(x, y, u, v)$. De esta forma, \mathbf{F} y \mathbf{G} tienen ($XY \times M$) y ($M \times XY$) dimensiones respectivamente. Sobre la matriz \mathbf{L} se disponen los valores de dicho *light-field* de tal forma que el resultado de la Ecuación 2.9 haga corresponder los elementos de \mathbf{F} y \mathbf{G} con los píxeles atravesados por cada rayo de luz considerado en la representación.

Una generalización de esta metodología es planteada por Wetzstein et al. (2012c), donde es posible considerar más de dos capas. En este caso, la Ecuación 2.9 se redefine como:

$$\mathbf{L} = \sum_{m=1}^M \bigotimes_{n=1}^N \mathbf{f}_{n,m}, \quad (2.10)$$

siendo \mathbf{L} la extensión n -dimensional o tensor de la definición propuesta anteriormente para el *light-field*. El operador \bigotimes se corresponde con el producto vectorial recursivo y se aplica sobre las columnas correspondientes de cada una de las capas $\mathbf{F}^{(n)}$. Esta generalización se denomina factorización no-negativa ponderada de tensores –del inglés *weighted nonnegative tensor factorization* (WNTF)– (Blondel et al., 2007). En este caso, el tensor \mathbf{L} pasa a tener

⁴El CFF depende de múltiples factores. Entre ellos, se puede destacar el valor de contraste e intensidad de la señal percibida, si esta es proyectada en la fovea o su periferia, cambios relativos o sensibilidad al contraste. No obstante, por motivos prácticos, se parte de un valor de 24Hz como frecuencia límite para el diseño de sistemas de visualización debido a la predominancia de conos en la fovea, siendo éstos menos sensibles al cambio en condiciones determinadas.

⁵Se entiende por rango de una matriz o determinante el número de columnas linealmente independientes. En el marco de la factorización de matrices y tensores resulta de especial interés, pues condiciona la exactitud de la separación de los factores. Esta propiedad se ha empleado extensivamente en el análisis de componentes principales (PCA) y *compressed sensing* con el fin de obtener información inherente a los datos contenidos en la matriz de estudio (Cichocki et al., 2009; Langville y Meyer, 2005; Paatero y Tapper, 1994).

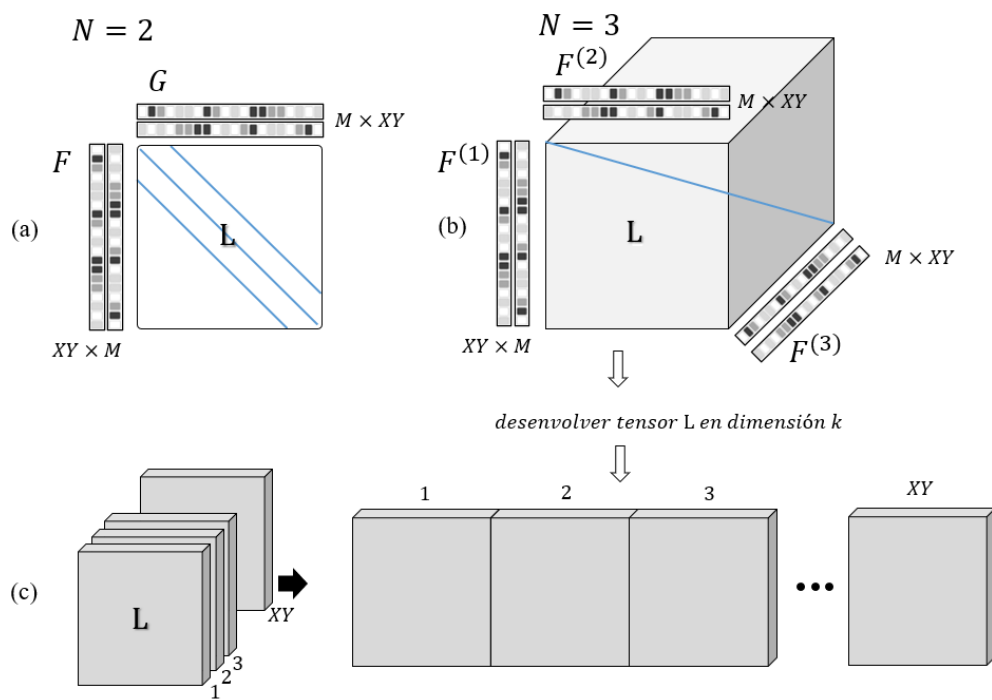


Figura 2.5 Representación esquemática de los métodos (a) WNMF y (b) WNTF, ambos para un sistema de dos y tres capas respectivamente y un número finito de modos. A partir del WNTF se muestra (c) la operación por el tensor es desvenuelto durante los productos intermedios de la técnica para una dimensión.

en sus N dimensiones un tamaño de XY . Asimismo, las matrices $\mathbf{F}^{(n)}$ tienen un número de elementos $XY \times M$.

De la misma manera que se plantea para el WNMF, la distribución de los valores de intensidad del *light-field* se realiza en L a razón de lo planteado en la Ecuación 2.8. En otras palabras, por ejemplo, para un conjunto de tres capas definidas como \mathbf{f} , \mathbf{g} y \mathbf{h} , los valores (f_i, g_j, h_k) se corresponderían con la posición $l_{i,j,k}$ del tensor L , siendo (i,j,k) un conjunto arbitrario de índices de la matriz.

Existen múltiples maneras de obtener los factores que componen L . La propuesta de Lanman et al. (2010) para dos capas y posteriormente de Wetzstein et al. (2012c) para N capas hace uso de las reglas multiplicativas (Blondel et al., 2007; Cichocki et al., 2009). Estas consisten en la minimización alterna del error ajustando capa a capa y se expresa como sigue:

$$\mathbf{F}^{(n)} \leftarrow \mathbf{F}^{(n)} \otimes \left(\frac{(\mathbf{W}_{(n)} \otimes \mathbf{L}_{(n)}) \mathbf{F}_{\odot}^n}{(\mathbf{W}_{(n)} \otimes (\mathbf{F}^{(n)} (\mathbf{F}_{\odot}^n)^T)) \mathbf{F}_{\odot}^n} \right), \quad (2.11)$$

siendo $L_{(n)}$ el resultado de *desenvolver* el tensor L en torno a la dimensión n . Para entender esto de forma sencilla, se muestra en la Figura 2.5 de forma esquemática. \mathbf{F}_{\odot}^n se corresponde con la siguiente definición:

$$\mathbf{F}_{\odot}^n \equiv \mathbf{F}^{(N)} \odot \dots \odot \mathbf{F}^{(n+1)} \odot \mathbf{F}^{(n-1)} \odot \dots \odot \mathbf{F}^{(1)}, \quad (2.12)$$

entendiéndose \odot como el operador de producto de *Khatri-Rao*. Este se define para dos matrices \mathbf{A} y \mathbf{B} como:

$$\mathbf{A} \odot \mathbf{B} = \begin{bmatrix} a_1 \otimes b_1 & a_2 \otimes b_2 & \dots & a_M \otimes b_M \end{bmatrix}. \quad (2.13)$$

donde \otimes es el producto de *Kronecker* de dos matrices⁶. Esta operación se define para dos matrices \mathbf{C} y \mathbf{D} como:

$$\mathbf{C} \otimes \mathbf{D} = \begin{bmatrix} c_{11} \mathbf{D} & c_{12} \mathbf{D} & \dots & c_{1J} \mathbf{D} \\ c_{21} \mathbf{D} & c_{22} \mathbf{D} & \dots & c_{2J} \mathbf{D} \\ \vdots & \vdots & \ddots & \vdots \\ c_{I1} \mathbf{D} & c_{I2} \mathbf{D} & \dots & c_{IJ} \mathbf{D} \end{bmatrix} \quad (2.14)$$

La aplicación de las reglas multiplicativas de forma alterna se demuestra en la bibliografía como un método que converge, esto es, reduce sistemáticamente el valor del error con cada iteración (Blondel et al., 2007). No obstante, esta convergencia se demuestra local, y por tanto, es posible obtener una solución fuera de los criterios deseados. Esto viene dado por la dependencia con el valor de inicio o semilla del método, y debe por tanto ser analizado con el fin de obtener un mejor resultado final.

⁶El producto de *Kronecker* es una generalización sobre las matrices del producto vectorial

2.4. Solución inicial para dos capas con WNMF

En la Sección 2.2 se introduce a los métodos de WNMF y WNTF como operación para la síntesis del *light-field* en un número finito de imágenes cuya superposición y posterior proyección en torno a un eje permite la reconstrucción visual de dicho *light-field*. Sin embargo, debido a que dicha síntesis se trata de un problema mal condicionado (Wang y Zhang, 2013), es imperativo dotar a estos algoritmos de un punto de partida que permita obtener un resultado más acorde con lo esperado. De no ser así, existe la probabilidad de que la solución estimada sea subóptima.

En esta Sección se plantea un método de inicialización para el algoritmo WNMF basado en ruido dependiente de la profundidad de la escena. Con esto, se busca mitigar los artefactos producidos al hacer únicamente uso de una señal de ruido. Asimismo, se evalúan diferentes métodos y se comparan con la propuesta presentada.

2.4.1. Inicialización basada en profundidad

El problema de la inicialización radica en la obtención de un punto de partida inicial adecuado al caso de estudio. Con esto se busca limitar la posibilidad de que la solución resultante sea sub-óptima o no se adecúe a los objetivos que subyacen al problema en cuestión. Se pueden encontrar en la bibliografía múltiples propuestas de inicialización para las reglas multiplicativas del WNMF. Cichocki et al. (2009) propone una heurística basada en la generación de múltiples matrices con ruido, o como resultado de un algoritmo NMF por mínimos cuadrados alternos –del inglés *alternate least squares* (ALS)–. Se obtuvo como solución inicial aquella que tuviera la función de menor coste tras 15-20 iteraciones. Por otro lado, Langville et al. (2014) estudia diferentes métodos de inicialización, encontrándose entre ellos el de *random-centroid*, *SVD-centroid*, y *random Acol*, resultando generalmente efectivos y de coste computacional reducido. Sin embargo, en la aplicación de las reglas multiplicativas sobre el sistema multicapa, existen pocos trabajos que planteen un método más allá de la inicialización por ruido (Lanman et al., 2010).

Como alternativa a métodos previos, se propone condicionar la solución inicial a la profundidad de la escena. Cuando se representa un *light-field* en un sistema multicapa, se está sintetizando la información en un conjunto de planos de tal manera que su solapamiento en diferentes direcciones dé lugar a una muestra válida de éste. En cierta manera, se puede entender esta operación como una tomografía de ángulo limitado, esto es, que las proyecciones se realizan dentro de un intervalo angular. Debido a este comportamiento, la resolución del problema equivale a la reconstrucción tridimensional del objeto en el espacio del sistema de visualización designado. Se puede demostrar que, dicha reconstrucción, es equivalente al operador de desenfoque presentado por Ng (2005), existiendo trabajos de referencia donde se parte de un conjunto de desenfoques para una escena a representar (Takahashi et al., 2018). Esto se puede observar en los resultados obtenidos en estudios previos (Gotoda, 2010; Lanman et al., 2010; Wetzstein et al., 2011, 2012b), de tal forma que, el conjunto de todas las capas obtenidas

después de la síntesis, es visualmente muy próximo a un desenfoque sintético de la vista central del *light-field* en la profundidad correspondiente. Más concretamente, la descripción del problema en la Sección 2.2 no es más que la aplicación de los conceptos básicos de la CT, es decir, la retroproyección filtrada, sobre un volumen discreto.

Este planteamiento ha sido empleado en sistemas multicapa previamente por Wang et al. (2015), pero no se ha evaluado sobre las reglas multiplicativas del algoritmo WNMF. El uso de ruido para inicializar las capas del sistema permite dar un punto de inicio diferente para cada uno de los instantes, por tanto permitiendo que el resultado sea multiplexado en el tiempo. Sin embargo, esto ocasiona en muchos casos, la presencia de ruido residual en ciertas partes de la imagen, especialmente en esas con frecuencias angulares bajas. Para mitigar este efecto, se propone limitar el ruido introducido en relación a la disparidad de la escena. Además, y siguiendo el razonamiento anterior, asumiendo el comportamiento de cada capa como un operador de desenfoque a diferentes alturas. Si tenemos en cuenta todo lo anterior, podemos definir el vector de soluciones iniciales \mathbf{F}_0^n como:

$$\mathbf{F}_0^n(x, y) = \text{abs}(\text{DM}(x, y)) \times \eta(x, y) + \mathbf{I}_n, \quad (2.15)$$

donde $\text{DM}(x, y)$ es el mapa de profundidades de la escena partiendo de la vista central del *light-field*, $\eta(x, y)$ es una función de ruido, e \mathbf{I}_n es la imagen central enfocada a la altura teórica asignada a la capa n .

Con el fin de evaluar posibles candidatos a solución inicial para el problema particular del sistema multicapa, se han elegido varias metodologías basadas en la bibliografía previa (Blondel et al., 2007; Cichocki et al., 2009; Lanman et al., 2010). Éstas quedan recogidas en la Tabla 2.1. Entre ellas se puede distinguir entre dos tipos: de valor fijo, y con ruido. Esencialmente, las funciones que son fijas son aquellas que consisten en inicializar con un valor único cada una de las capas. En la síntesis para el caso de multiplexado temporal, esto supone que cada uno de los instantes de tiempo partiría del mismo valor, por lo que las reglas multiplicativas no permitirían una síntesis multimodal, esto es, cuando $M > 1$. Por ello, predomina en la literatura el uso de ruido como inicialización, ya que permite en este caso diferenciar sutilmente cada uno de los modos de cada capa.

Cuadro 2.1 Métodos de inicialización y propuesta.

Métodos	Parámetros
Unidad $F_0^n(x, y) = ones(x, y)$	-
Gaussiano $F_0^n(x, y) = \eta(x, y)$	$\mu = 0$ y $\sigma = 1$
Sal&Pimienta $F_0^n(x, y) = S P_{noise}(x, y)$	0.1 ratio
Reenfoque $F_0^n(x, y) = I_n$	-
Gaussiano+Reenfoque $F_0^n(x, y) = \eta(x, y) + I_n$	$\mu = 0$ y $\sigma = 1$
Sal&Pimienta $F_0^n(x, y) = S P_{noise}(x, y) + I_n$	0.1 ratio
Acol	-
Métodos propuestos	
DM+Desenfoque+Gaussiano $F_0^n(x, y) = abs(DM(x, y)) \times \eta(x, y) + I_n$	$\mu = 0$ and $\sigma = 1$
DM+Desenfoque+Sal&Pimienta $F_0^n(x, y) = abs(DM(x, y)) \times S P_{noise}(x, y) + I_n$	0.1 ratio

2.4.2. Simulaciones

Se ha realizado la comparación de los métodos propuestos en la Tabla 2.1 sobre un conjunto de cuatro *light-fields*: *Dice*, *Green Dragon* y *Messerschmit* y *Dice2*. Los tres primeros fueron obtenidos del *MIT Synthetic Light-Field Archive* ⁷, mientras que el último fue generado en *Blender*, un programa de renderizado 3D (Blender Foundation, 2018).

Las simulaciones realizadas se corresponden con los siguientes parámetros:

- Resolución de *light-field*: 512×384 píxeles; 7×7 vistas
- Campo de visión del *light-field*: 10grados
- Número de capas: 2
- Modos o instantes de tiempo: 6
- Número de iteraciones: 1000

⁷<http://web.media.mit.edu/~gordonw/SyntheticLightFields/>

En el caso de las técnicas que requieren mapa de profundidad, es posible obtener una aproximación a partir del *light-field*. De las implementaciones actuales, se hace uso concretamente de la propuesta presentada por Lüke et al. (2015), basada en la obtención de profundidad a partir de los gradientes de los planos que conforman el *light-field* en un espacio 4D. Cada uno de los mapas se ha normalizado a un rango de [0,1], asignando dichos valores respectivamente a las partes más cercanas y lejanas al centro.

Se ha establecido como métrica para la evaluación en la mejora de la calidad la proporción máxima de señal a ruido, del inglés *peak signal to noise ratio* (PSNR) y que se define de la siguiente manera:

$$\text{PSNR} = 20 \times \log \left(\frac{\max(I)}{\sqrt{\text{MSE}}} \right) \quad (2.16)$$

donde $\max(I)$ se corresponde con el máximo valor que un píxel puede tomar dentro de la imagen, y MSE es el error cuadrático medio.

El análisis se ha llevado a cabo partiendo de un sistema de dos capas y asumiendo una retro-iluminación estática como se presenta en Wetzstein et al. (2011). Se ha procedido de esta forma a favor de la simplicidad, ya que el incremento del número de capas afecta sustancialmente a la complejidad computacional del algoritmo. Sin embargo, debido a la generalización de las reglas multiplicativas, es posible teorizar que los resultados que se presentan a continuación son extensibles a un número superior de capas. Se ha implementado el algoritmo de tomografía presentado en dicho estudio sobre MATLAB R2016a (Mathworks, Natick, MA) en un procesador Intel(R) Core(TM)2 Quad2.50GHz y 8.00GB de RAM. Siguiendo lo propuesto, se han generado un conjunto de soluciones iniciales correspondientes a los métodos indicados en la Tabla 2.1.

En la Figura 2.6 se muestra la relación del PSNR con el número de iteraciones para cada uno de los métodos estudiados. La métrica se ha aplicado sobre la vista central por simplicidad. Cada una de las curvas representa la media de los resultados obtenidos para el total de las cuatro escenas propuestas. En la gráfica se observa que, con la salvedad de los métodos estáticos, se percibe una mejora continua en PSNR pasadas las 10 iteraciones. Se puede observar que, las curvas, tienden a converger a un valor para un número elevado. En relación a los métodos estáticos, se alcanza un máximo local a las pocas iteraciones, quedando siempre por debajo de las demás curvas. Este comportamiento queda mejor ilustrado en la Figura 2.7. Cabe destacar que, los métodos propuestos alcanzan un valor de PSNR superior durante las primeras iteraciones, seguidos por el método de ruido *Gaussiano*+Desenfoque. Estos continuaron mejorando el valor señal ruido sensiblemente hasta alcanzar las 1000 iteraciones.

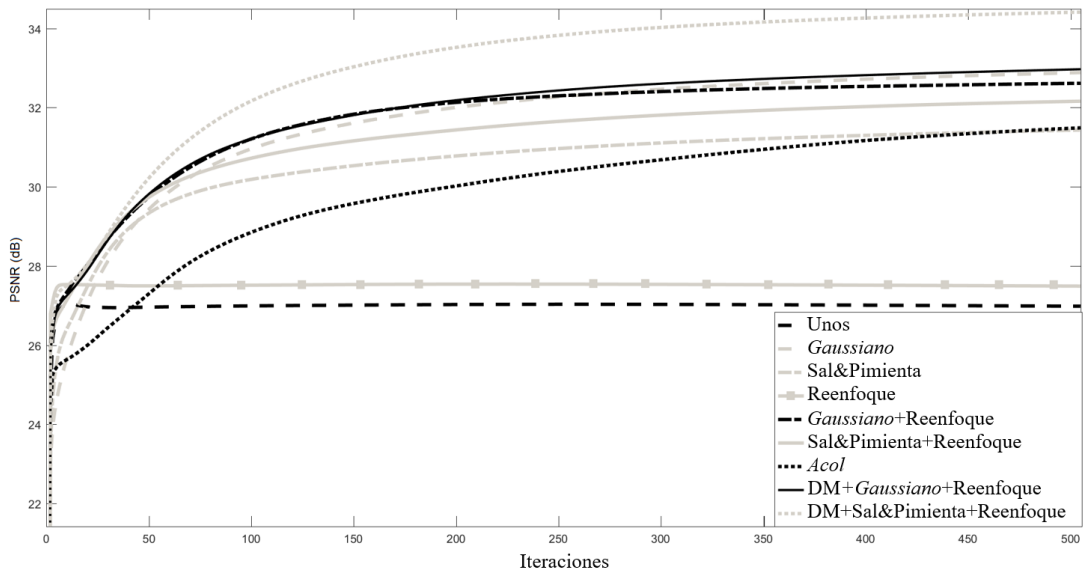


Figura 2.6 Tendencia del valor PSNR con respecto al número de iteraciones para cada uno de los métodos propuestos como solución inicial en el caso de WNMF

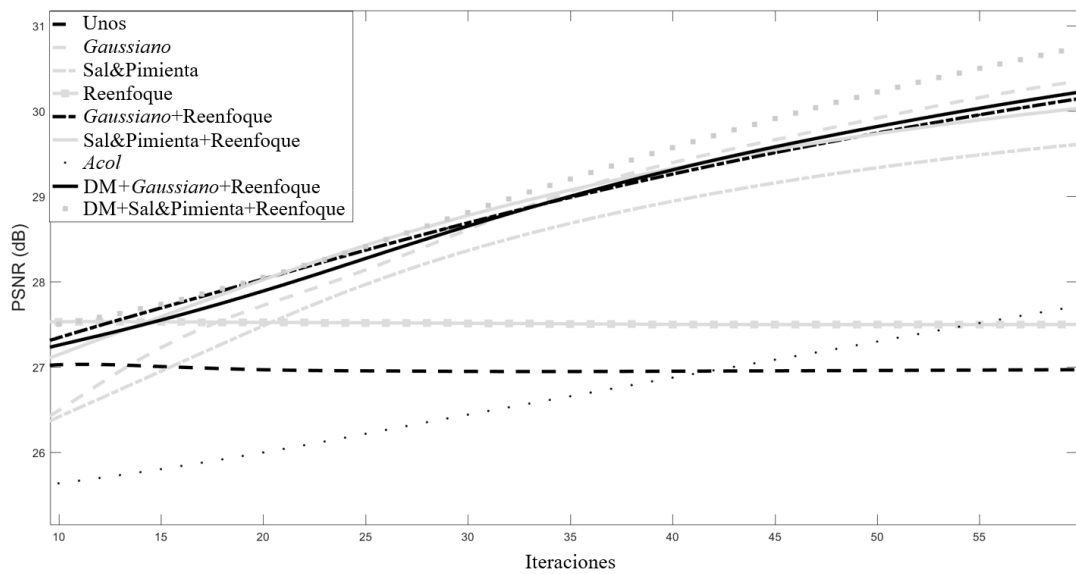


Figura 2.7 Tendencia del valor PSNR para las primeras 60 iteraciones de las reglas multiplicativas.

Los valores finales para cada método quedan recogidos en la Tabla 2.2. Junto a cada técnica se indica el valor de la desviación típica, siendo esta definida como:

$$\sigma_{method} = \sqrt{\frac{\sum(\text{PSNR}_{LF} - \overline{\text{PSNR}})}{N_c}}, \quad (2.17)$$

con σ_{method} como la desviación típica, \overline{PSNR} siendo el valor medio del PSNR para el total de casos N_c . Comparativamente, los métodos basados en profundidad refieren una mejora con respecto a los demás. Específicamente, los valores inferiores de PSNR se corresponden con las técnicas de inicialización estáticas.

Cuadro 2.2 Valores de PSNR en dB obtenidos para un sistema de dos capas y descomposición en seis instantes de tiempo tras 1000 iteraciones. En la última columna se indica la desviación estándar (STD) para cada uno de los métodos.

Métodos de inicialización	Dice (dB)	Dice2 (dB)	Dragon (dB)	Messer. (dB)	STD(dB)
Unos	23.35	29.43	27.94	26.51	±2.59
<i>Gaussiano</i>	26.61	34.43	37.65	34.22	±4.68
Sal&Pimienta	25.29	33.81	35.10	32.46	±4.38
Reenfoque	25.29	29.45	27.96	26.57	±1.79
<i>Gaussiano</i> +Reenfoque	27.84	32.92	36.66	33.64	±3.66
Sal&Pimienta+Reenfoque	28.29	33.84	35.08	32.48	±2.95
<i>Acol</i>	25.60	34.64	37.39	34.27	±5.10
Métodos propuestos					
DM+Reenfoque+ <i>Gaussiano</i>	26.29	34.56	37.95	34.39	±4.95
DM+Reenfoque+Sal&Pimienta	31.28	34.68	38.25	34.41	±2.85

De manera ilustrativa, en la Figura 2.8 se muestra diagrama de barras donde cada elemento se corresponde con el promedio del PSNR acompañado por su desviación típica. Se puede observar como de los dos métodos propuestos, el que emplea ruido tipo *Sal&Pimienta* mejora considerablemente comparado a todos los demás, mientras que el que hace uso de ruido *Gaussiano* muestra un valor similar al de otros métodos que no hacen uso de mapa de profundidad. Esta discrepancia surge de la escena *Dice*, donde el ruido tiende a generar una serie de halos en torno a los objetos desenfocados. Este tipo de efecto se muestra mitigado en los métodos que hacen uso de profundidad. Esto puede verse en la Figura 2.9. Sin embargo, como se ha indicado, el caso de *Sal&Pimienta* con DM resulta en una atenuación más significativa que el *Gaussiano*. Se puede teorizar que este efecto tiene origen en el carácter binario de este método de *Sal&Pimienta*. Cabe destacar que esta discrepancia tiene lugar solo en este caso, comportándose el resto de forma favorable a los métodos propuestos.

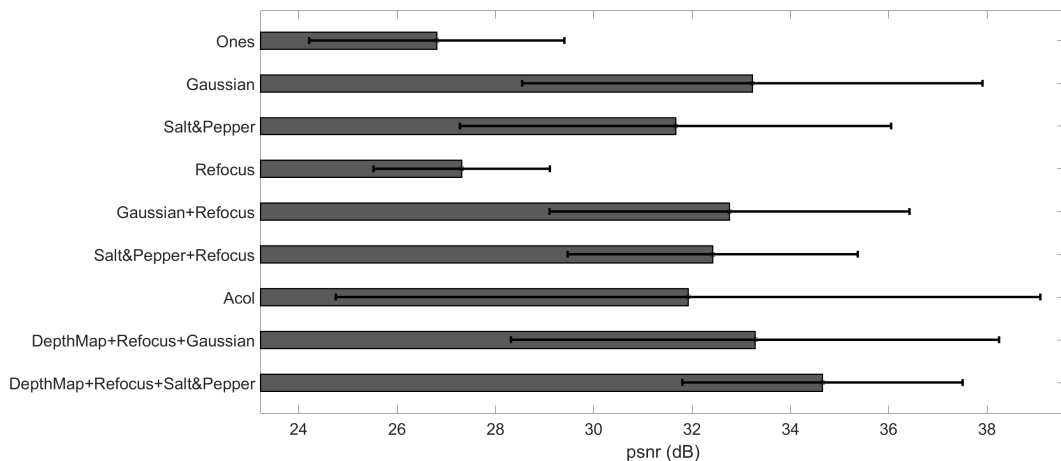


Figura 2.8 Diagrama de barras con el valor promedio de PSNR para cada método una vez alcanzadas las 1000 iteraciones. La desviación estándar se representa por medio de una línea recta sobre cada barra.

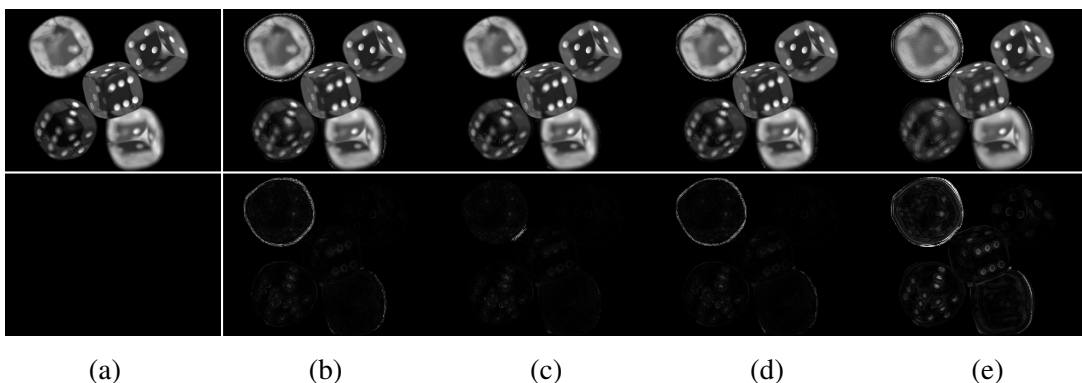


Figura 2.9 Vista central de la escena *Dice*. De izquierda a derecha: referencia (a) ruido *Sal&Pimienta* (b) *DM+Sal&Pimienta+Reenfocue* (c) *DM+Gaussiano+Reenfocue* (d) y *Acol* (e). La primera fila contiene el resultado de la proyección, mientras que la segunda muestra el error entre la referencia y el método.

Este análisis permite concluir que la inicialización en el sistema multicapa juega un papel, que si bien no es definitivo, puede suponer una mejoría considerable sobre la representación del *light-field* en contraposición a trabajos previos (Lanman et al., 2010). Asimismo, cabe destacar que los métodos propuestos no ofrecen una mejoría transferible a cualquier representación, sino que pueden depender del contenido de la escena, como se ha visto en el caso de los desenfocues. Este primer análisis abre las puertas al estudio de métodos de mayor complejidad, y su extensión a un conjunto superior de casos: tomando diferentes consideraciones en relación a ponderaciones por posicionamiento espacial de los observadores (Chen et al., 2014; Ji y Park, 2016).

2.5. Optimización de reglas multiplicativas de WNMF

El método de factorización empleado en la síntesis del *light-field* conlleva un coste computacional elevado, principalmente asociado a los productos intermedios de las reglas multiplicativas (Lanman et al., 2010; Paatero y Tapper, 1994). En esta Sección se introduce a una implementación alternativa de los algoritmos descritos en la Sección 2.2 para el caso de WNMF.

Si se toma como base la descomposición de la matriz dispersa del *light-field* siguiendo las reglas multiplicativas (véase Expresiones 2.11 y 2.12 para el sistema de dos y tres capas respectivamente) se observa que se realiza de forma alterna una corrección sobre los valores de los vectores en los que ésta se descompone, compensando el resultado intermedio de su producto vectorial. No en vano, dichas reglas se derivan de la minimización de la norma euclídea de forma alterna y se demuestra su convergencia a un mínimo local (Cichocki et al., 2009; Grippo y Sciandrone, 2000; Paatero y Tapper, 1994). Un inconveniente de esta metodología es la aparición de productos intermedios. A pesar de que la matriz a descomponer es dispersa, los productos generados no lo son: por tanto, dependiendo de las dimensiones del problema, esto puede repercutir de forma negativa en el coste computacional y recursos necesarios para llevarla a cabo.

Partiendo de la Ecuación 2.11, se puede extraer del denominador lo siguiente:

$$\mathbf{W}_{(n)} \otimes (\mathbf{F}^{(n)}(\mathbf{F}_{\odot}^n)^T) \quad (2.18)$$

El resultado de $(\mathbf{F}^{(n)}(\mathbf{F}_{\odot}^n)^T)$ es una matriz temporal que sirve para dar cuenta del error con respecto a la matriz objetivo, y es de las mismas dimensiones que la matriz dispersa del *light-field*. Sin embargo, en este caso la matriz se encuentra completamente poblada, y es de hecho el producto de *Hadamard* con la matriz de pesos, el que da lugar a la matriz dispersa nuevamente.

De manera ilustrativa, se representa en la Figura 2.10 la matriz dispersa para un *light-field* arbitrario de 7x7 vistas (U,V) y una resolución por vista de 384x512 píxeles (X, Y). El tamaño total de la matriz dispersa se puede computar como $K \times K = (X \times Y)^2$, siendo el número de elementos no nulos el total de valores comprendidos por el *light-field* o expresado de manera algebraica $X \times Y \times U \times V$. Para el supuesto presente esto implica los siguientes valores:

- $Elementos_{Total} = (384 \times 512)^2 = 38,654,705,664$
- $Elementos_{NoNulos} \leq 384 \times 512 \times 7 \times 7 = 9,633,792$

Como se puede observar, el número de elementos no nulos es cinco órdenes de magnitud inferior al total de elementos de la matriz, resultando en una razón de dispersión muy elevada.

Existe una amplia bibliografía destinada a la optimización de los algoritmos WNMF y WNTF, y su implementación sobre *hardware* como GPU o FPGA. Muchos de estos trabajos abordan el problema desde una perspectiva aséptica, no tomando en consideración la aplicación particular del método (Antikainen et al., 2011; Cerina et al., 2017; Zhang et al., 2009). Sin embargo, es posible tomar ventaja sobre la representación del *light-field* en este método.

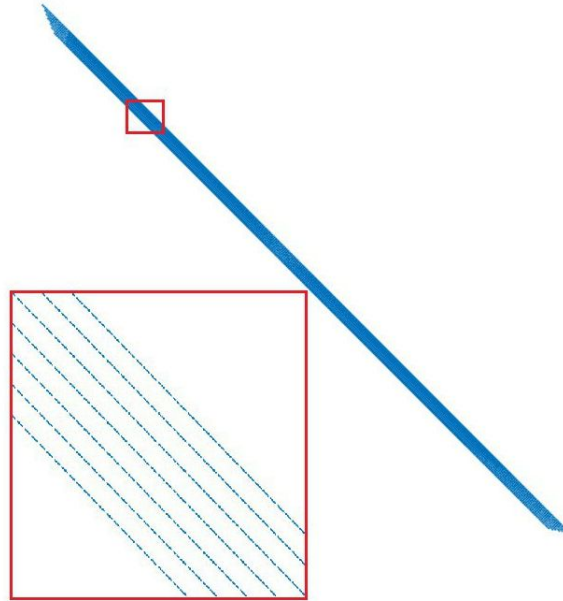


Figura 2.10 Representación del *light-field* sobre una matriz dispersa para el caso de dos pantallas

Siguiendo la formulación del *light-field* en basada en la topología del sistema multicapa (véase Ecuación 2.8), es posible redimensionar la matriz dispersa tal que su representación se limite al tamaño en memoria de éste. Para entender esta operación, conviene observar la distribución de las vistas sobre la matriz dispersa representada en la Fig. 2.10 correspondiéndose cada una de éstas con las diagonales. De manera intuitiva, y siguiendo la Ecuación 2.8, se puede observar que para la vista central corresponde la diagonal principal de la matriz, mientras que las vistas laterales cuentan con un desplazamiento lateral con respecto a la misma. Esta distribución contiene de forma implícita la trayectoria de los rayos de luz y como interactúan estos con los píxeles de cada pantalla. Por tanto, para poder reducir la dimensionalidad de este sistema, es necesario tomar en cuenta dicha relación.

Sea \mathbf{L}' una nueva matriz de dimensiones $(U \times V, X \times Y)$ sobre la que se representa en cada fila una vista del *light-field*. Ahora, y de manera similar al método previo, se introduce un desplazamiento lateral correspondiente a cada una de estas filas a razón de sus coordenadas (u, v) . Para mayor claridad, en la Figura 2.11 se muestra de manera esquemática este proceso. Partiendo de un sistema compuesto por dos capas denotadas como \mathbf{F} y \mathbf{G} , y manteniendo la representación de las capas en formato vectorial, es posible reacondicionar el numerador de la Ecuación 2.8 de la siguiente manera:

$$num[\mathbf{F}]_j = \begin{cases} \sum_{i=1}^{X \times Y} \mathbf{L}'_{i,j} \times \mathbf{G}_{j+\text{offset}(i)} & 1 \geq j + \text{offset}(i) \leq X \times Y \\ 0 & \text{cualquier otro caso} \end{cases}, \quad (2.19)$$

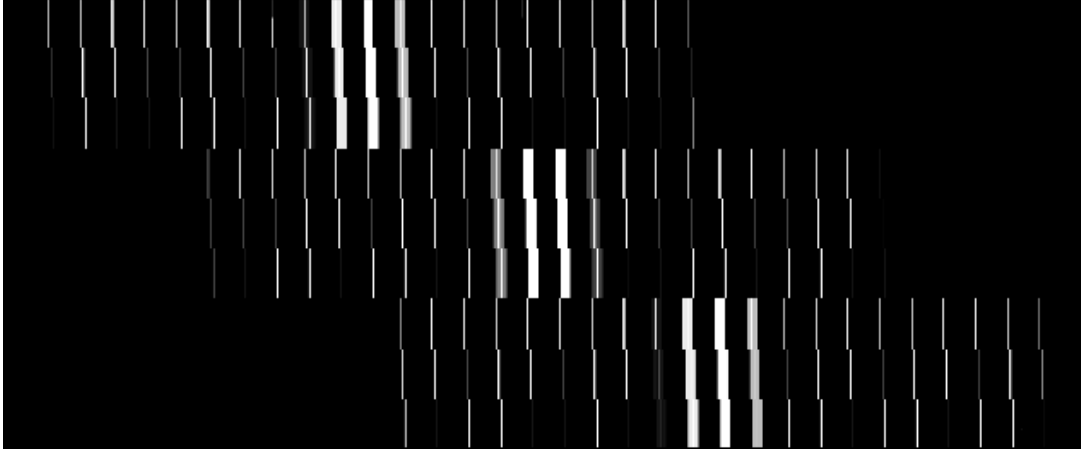


Figura 2.11 Reacondicionamiento de la matriz dispersa de un *light-field* de 3×3 vistas en L' . Obsérvense las franjas horizontales correspondientes a cada una de las vistas.

siendo (i,j) los índices de la nueva matriz. El vector **offset** se define de manera auxiliar para alojar el desplazamiento lateral correspondiente a cada una de las vistas (u, v) . Nótese que, al igual que en la matriz original, se pierde información de las vistas laterales en tanto a las áreas en las que la propagación de la luz no interactúa con todas las pantallas. Por ello, se asume que $num[\mathbf{F}]_j = 0$ si no se cumple $1 \geq j + \mathbf{offset}(i) \leq X \times Y$. La operación anteriormente indicada es igualmente válida para la capa \mathbf{G} , pero hay que tener en cuenta que el vector de desplazamientos, así como la distribución de la matriz es opuesta, esto es, que se obtiene a raíz de la matriz dispersa traspuesta.

De manera análoga, los productos intermedios surgidos a partir del producto vectorial entre las capas \mathbf{F} y \mathbf{G} en cada iteración pueden redefinirse de la siguiente manera:

$$L'[\mathbf{F}]_{i,j} = \begin{cases} \mathbf{F}_j \mathbf{G}_{j+\mathbf{offset}(i)} & 1 \geq j + \mathbf{offset}(i) \leq X \times Y \\ 0 & \text{cualquier otro caso} \end{cases}, \quad (2.20)$$

siendo $L'[F]_{(i,j)}$ el resultado temporal del producto vectorial entre capas. El mismo razonamiento se sigue en el caso de la actualización de la capa \mathbf{G} .

Finalmente, siguiendo el razonamiento previo, es posible adaptar las reglas multiplicativas para el sistema de dos capas:

$$\mathbf{F}_j \leftarrow \mathbf{F}_j \times \left(\frac{\sum_{i=1}^{X \times Y} L'_{i,j} \mathbf{G}_{j+\mathbf{offset}(i)}}{\sum_{i=1}^{X \times Y} L'[\mathbf{F}]_{i,j} \mathbf{G}_{j+\mathbf{offset}(i)}} \right) \quad (2.21)$$

$$\mathbf{G}_j \leftarrow \mathbf{G}_j \times \left(\frac{\sum_{i=1}^{X \times Y} L'_{i,j} \mathbf{F}_{j+\mathbf{offset}(i)}}{\sum_{i=1}^{X \times Y} L'[\mathbf{G}]_{i,j} \mathbf{F}_{j+\mathbf{offset}(i)}} \right) \quad (2.22)$$

Un concepto importante de esta conversión radica en la desaparición de la matriz de pesos. Esto se debe a que, por lo general, esta matriz actúa como una máscara binaria (Lanman et al., 2010; Wetzstein et al., 2012c), a pesar de que en algunos trabajos se ha empleado para ponderar

ciertas vistas en función a diferentes criterios (Chen et al., 2014; Ji y Park, 2016). Sin embargo, esta puede reincorporarse a las reglas adoptando una aproximación similar, y tomando en cuenta como se redistribuye el *light-field* en ella. A efectos prácticos, se puede reformular \mathbf{W} como una matriz \mathbf{W}' al igual que se ha hecho con \mathbf{L} . Intuitivamente, los pesos correspondientes a cada vista se agrupan en las filas de \mathbf{W}' . Esto resulta en las siguientes expresiones para cada capa:

$$\mathbf{F}_j \leftarrow \mathbf{F}_j \times \left(\frac{\sum_{i=1}^{X \times Y} [\mathbf{W}'_{i,j} \mathbf{L}'_{i,j}] \mathbf{G}_{j+offset(i)}}{\sum_{i=1}^{X \times Y} [\mathbf{W}'_{i,j} \mathbf{L}'_{i,j}] \mathbf{F}_{j+offset(i)}} \right) \quad (2.23)$$

$$\mathbf{G}_j \leftarrow \mathbf{G}_j \times \left(\frac{\sum_{i=1}^{X \times Y} [\mathbf{W}'_{i,j} \mathbf{L}'_{i,j}] \mathbf{F}_{j+offset(i)}}{\sum_{i=1}^{X \times Y} [\mathbf{W}'_{i,j} \mathbf{L}'_{i,j}] \mathbf{G}_{j+offset(i)}} \right) \quad (2.24)$$

En el Anexo A se muestra la implementación realizada en Matlab para 1 y M instantes de tiempo respectivamente.

2.6. Resultados e implementación en modelo experimental

En las secciones previas se han introducido los métodos clásicos para la síntesis tomográfica del *light-field* en el sistema multicapa. Estos métodos computacionales permiten trasladar la información contenida en dicha función de forma computacional. En esta Sección se presenta la implementación de los resultados obtenidos sobre un modelo experimental.

2.6.1. Modelo propuesto

Debido a la estructura multicapa del sistema, es posible la aparición de artefactos y efectos no deseables en la imagen. Esto es debido, en mayor medida, a que cada una de las capas puede acometer un cierto grado de tolerancia en su posición y rotación relativa al sistema. De ello surge un desalineamiento entre píxeles que puede limitar la definición propuesta en la Ecuación 2.8.

De manera alternativa a trabajos previos, se plantea como objetivo el diseño de un sistema tal que, dichos errores puedan ser manualmente ajustados en función de los requerimientos impuestos por el contenido a representar. Para ello, se propone un sistema compuesto por tres capas, siendo la frontal y trasera configurables en seis grados de libertad con respecto a la central. En la Figura 2.12 se muestra una representación tridimensional del modelo propuesto. La capa central actúa como referencia y soporte para el resto del sistema. Sobre el marco de esta, se disponen unos raíles con cierto grado de flexibilidad, que actúan como guía para las capas exteriores. Las capas frontal y trasera están compuestas por dos marcos, uno interno, sobre el que se dispone la lámina LCD, y otro externo, que actúa como soporte de esta última. Dicho marco incluye dos micrómetros en la parte inferior, y uno en su lateral, siendo estos los encargados de controlar el *roll* relativo del marco interior, así como los desplazamientos

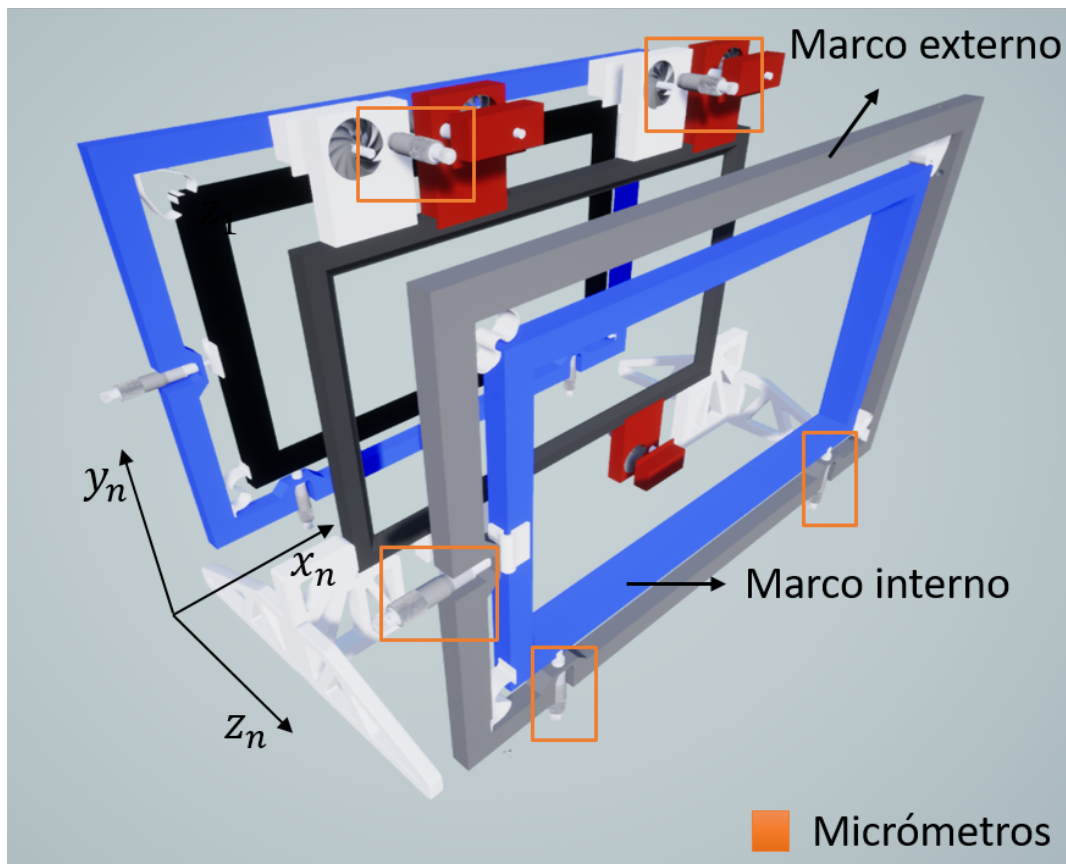


Figura 2.12 Modelo del sistema ajustable de tres capas propuesto.

verticales y horizontales. El control de desplazamiento en profundidad y las rotaciones de *yaw* y *pitch* se regulan por medio de unos micrómetros dispuestos en sentido contrario a las guías.

En colaboración con el Departamento de Ingeniería Mecánica de la Universidad de La Laguna, se ha ensamblado un modelo basado en el diseño propuesto. La totalidad del cuerpo del dispositivo se ha fabricado por medio de la técnica de impresión 3D. En contraposición a los micrómetros, se ha añadido un resorte. Esto sirve como contrapunto, y obliga a las pantallas a estar en una posición fija. Al soporte se le han adherido tres pantallas láminas de cristal líquido (LC) extraídas de monitores *BENQ XL2411 144Hz*. Se ha elegido este modelo de pantalla debido a la elevada tasa de refresco. Esto permite alcanzar valores superiores de multiplexado temporal y, por tanto, obtener un mejor rendimiento en el dispositivo. Asimismo, la electrónica, incluyendo controladores y alimentación de las pantallas, se ha acondicionado en la parte superior de cada marco, correspondiéndose con su capa de manera solidaria.

Finalmente, se han introducido un par de polarizadores en la pantalla delantera y trasera, respectivamente, desfasados 90 grados. De la misma manera, se introduce un difusor en la pantalla delantera con el fin de mitigar el patrón de *moiré* debido a la multiplicidad de rejillas Bell et al. (2007); Hirsch et al.. Como iluminación trasera, se ha hecho uso de la propia fuente *LED* procedente del modelo de monitor empleado. En la Figura 2.13 se muestra el

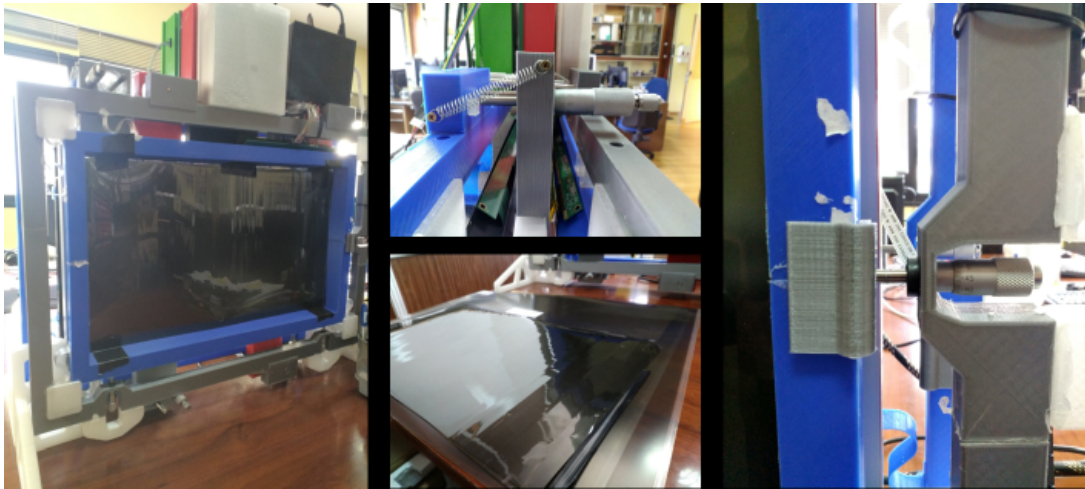


Figura 2.13 Modelo experimental del sistema multicapa ajustable para tres pantallas. A la izquierda se muestra el prototipo final. A la derecha, se muestran las diferentes partes que constituyen el dispositivo.

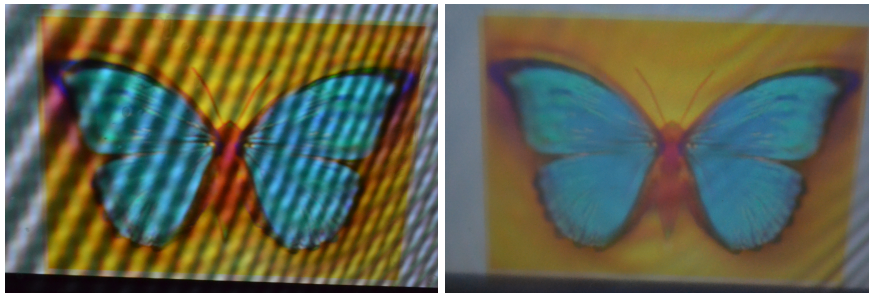


Figura 2.14 Comparativa entre el sistema sin difusor (izquierda) y con difusor (derecha)

resultado obtenido. A la izquierda de la Figura se encuentra el prototipo final. De manera más concreta, se ilustra a la derecha con varios componentes del dispositivo, siendo estos los pares de micrómetros resortes que actúan como guías, los polarizadores y difusor, y el apoyo lateral del marco externo de una de las capas.

Un defecto inherente a la técnica es la aparición de patrones de *moiré*. Estos surgen a raíz de la superposición de rejillas con un cierto desfase relativo (Glassner, 1997). No obstante, y de manera similar a otros trabajos (Hirsch et al.), hemos introducido un difusor en la capa delantera con el fin de mitigar este defecto. En la Figura 2.14 se muestra una comparativa entre el sistema previo y posterior a la introducción de este. El objetivo tras esta disposición se encuentra en limitar las frecuencias que se representan en el sistema, y en consecuencia, evitar dentro de lo posible la aparición de este fenómeno.

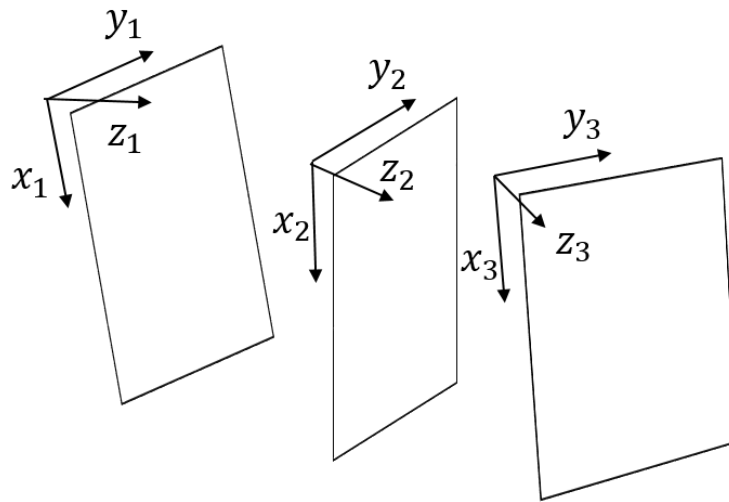


Figura 2.15 Orientación arbitraria de planos para el sistema de tres capas propuesto

2.6.2. Calibración del sistema

La calibración del sistema se realiza a partir de la medida de un estado no deseado, donde se asume que cada uno de los planos se encuentra orientado de forma distinta con respecto al centro. Más concretamente, se busca alinear las pantallas de tal forma que, idealmente, los centroides de cada pixel queden alineados con su vecindad en las demás capas asumiendo que son proyectados a infinito. Se asume por tanto una orientación inicial de las pantallas tal como la que se muestra en la Figura 2.15. Se definen a partir de cada plano unos ejes de coordenadas cartesianos (x_n, y_n, z_n) , estando cada capa n contenida en el plano definido por (x_n, y_n) .

Se definen los cambios relativos de los ejes de coordenadas con respecto al plano central, condicionando las transformaciones de los planos externos a dicho origen. Con el fin de estimar los parámetros de ajuste del sistema, se parte del uso de un patrón de calibración conocido. Considerando el comportamiento transmisor de las pantallas LCD, es posible representar el mismo patrón individualmente en cada capa, de tal manera que las dos restantes no afecten sobre dicho contenido. Un patrón comúnmente empleado es el *tablero de ajedrez* o *checkerboard*, el cual se define por una rejilla de cuadrados blancos y negros alternos. En la Figura 2.16 se muestra un ejemplo del patrón de calibrado empleado.

El proceso de obtención de dichos parámetros se basa en la estimación de la transformación sufrida por el patrón tras un cambio en el sistema de coordenadas inicial, y es conocido como homografía (Semple J. and Kneebone G, 1979). Se define como matriz de homografía, aquella que contiene la relación entre un sistema de coordenadas y otro. Más concretamente, esto se puede expresar de la siguiente manera:

$$\mathbf{H}^* \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} \quad (2.25)$$

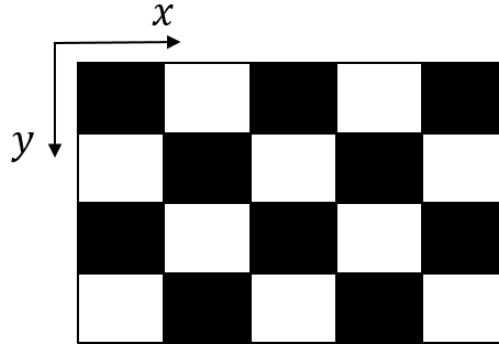


Figura 2.16 Patrón de calibración

donde \mathbf{H} es la matriz de homografía y se define como:

$$\mathbf{H} = \begin{bmatrix} s_x \cos \theta & s_y \cos \theta & t_x \\ s_y \cos \theta & s_x \cos \theta & t_y \\ 0 & 0 & 1 \end{bmatrix}, \quad (2.26)$$

siendo θ el ángulo de rotación, t_x, t_y la traslación, y s_x, s_y el factor de escala. La matriz presentada en la Ecuación 2.26 se corresponde con una transformación de similaridad: solo tiene en cuenta traslación, rotación y escala.

A partir de la definición de matriz realizada, es posible extraer individualmente los parámetros obtenidos como sigue:

$$\begin{aligned} t_x &= a_{1,3} \\ t_y &= a_{2,3} \\ s_x &= \sqrt{a_{1,1}^2 + a_{2,1}^2} \\ s_y &= \sqrt{a_{1,3}^2 + a_{2,2}^2} \\ \theta &= \arctan -\frac{a_{2,1}}{a_{1,1}}; \theta = \arctan -\frac{a_{2,1}}{a_{1,1}} \end{aligned} \quad (2.27)$$

En el caso particular de la calibración propuesta sobre el sistema implementado, los pasos a seguir son los siguientes:

1. Captura del plano n con el patrón especificado. Para lograr esto, las demás capas deben tomar el valor 0 (no interfieren modificando el ángulo de la polarización).
2. Tomando como referencia el plano intermedio, se computa la matriz de homografía.
3. Se ajustan los micrómetros de cada capa de acuerdo a los parámetros obtenidos.

La estimación de la matriz de homografía se puede realizar por medio de una descomposición en valores singulares (SVD) (Semple J. and Kneebone G, 1979). La implementación de tal método se ha realizado sobre Matlab2016R, y se basa en el código propuesto por Young



Figura 2.17 Montaje para la calibración del dispositivo multicapa

(2008). Para la adquisición de los patrones representados en el sistema se ha hecho uso de una cámara *SONY α6000* con un campo de visión de 50 grados y situada a 90cm de la capa delantera. En la Figura 2.17 se muestra el montaje realizado.

En la Tabla 2.3 están contenidas las estimaciones realizadas para el sistema propuesto, tomando en cuenta tanto la capa delantera como la trasera. La primera columna muestra los valores obtenidos para el sistema sin calibrar. Nótese que tanto capa trasera como delantera tienen un cierto valor de escala acorde a su distancia con respecto al centro. De manera similar, en la segunda columna están contenidos los valores tras realizar el ajuste manual del sistema. Salvando la rotación, los demás parámetros se adhieren a un resultado similar. Esta discrepancia tiene origen en el factor de escala obtenido.

Con fines ilustrativos se muestra en la Figura 2.18 una comparativa entre el sistema previo (izda.) y posterior (dcha.) a la calibración. Como se observa, ha sido posible alcanzar un cierto grado de alineamiento entre capas. No obstante, con el fin de evaluar el error residual del ajuste, es necesario, de manera computacional, escalar las imágenes con acuerdo a los valores extraídos en la segunda columna de la Tabla 2.3. Haciendo esto, y tras computar nuevamente la matriz de homografía, se obtienen los parámetros mostrados en la tercera columna. En este caso si se ha reducido el error, quedando limitado a una transformación rígida: rotación y traslación. La Figura 2.19 muestra el resultado obtenido una vez realizada la corrección de escala. A la izquierda se muestra la imagen del prototipo, mientras que a la derecha se le ha realizado la transformación de manera computacional con fines ilustrativos.

2.6.3. Capturas del sistema

Realizada la calibración del sistema, con el fin de demostrar la viabilidad del dispositivo diseñado, se han representado un conjunto de *light-fields*. Para la obtención de las vistas del

Cuadro 2.3 Parámetros extraídos de la matriz de homografía.

Capa delantera	Sin calibrar	Tras calibración	Corrección de escala
$t_x(u)$	16,778	16,687	3,237
$t_y(u)$	-12,106	9,543	3,496
s_x	0,986	0,983	0,994
s_y	0,984	0,980	0,994
$\theta(deg)$	0,309	0,413	0,030
Capa trasera	Sin calibrar	Tras calibración	Corrección de escala
$t_x(u)$	-24,778	-22,754	2,492
$t_y(u)$	-11,012	-11,826	1,472
s_x	1,035	1,034	1,003
s_y	1,033	1,031	1,003
$\theta(deg)$	-0,927	0,127	0,122

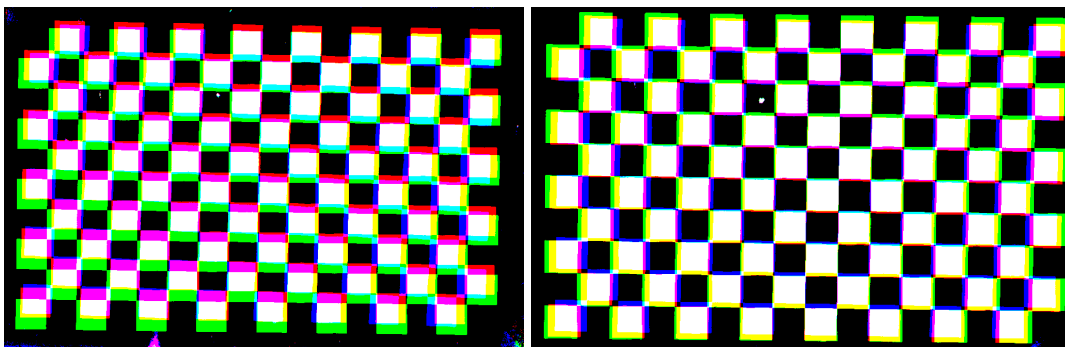


Figura 2.18 Comparación entre el sistema sin calibración y tras ser calibrado respectivamente. La imagen se ha representado en formato RGB, donde cada canal de color está asociado a una capa, siendo rojo para la capa media, verde para la delantera y azul para la trasera.

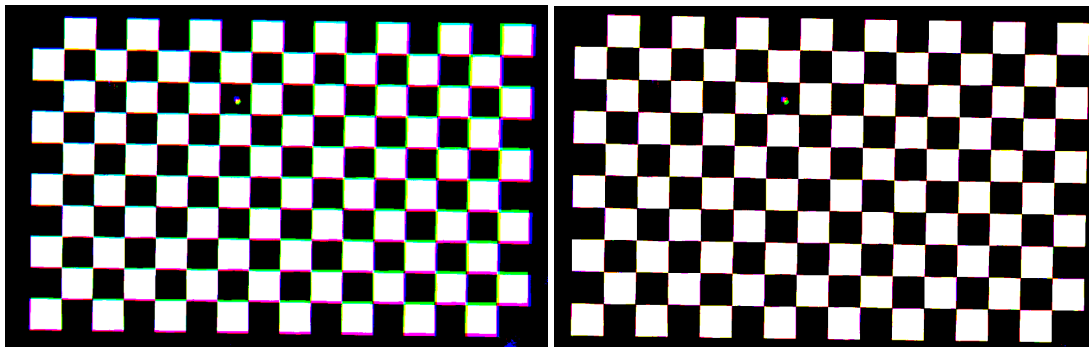


Figura 2.19 Comparación entre el sistema calibrado con corrección de escala, y tras alineamiento en simulación por medio del error residual. La imagen se ha representado en formato RGB, donde cada canal de color está asociado a una capa, siendo rojo para la capa media, verde para la delantera y azul para la trasera.

sistema, se ha hecho uso del mismo sistema de captura empleado en la calibración. En la Figura 2.20 se muestran los resultados para varios *light-field* sintéticos. Los tres primeros se corresponden con escenas extraídas del repositorio de *light-field* del MIT. La última se ha renderizado por medio del *software Blender* (Blender Foundation, 2018). En todos los casos, los parámetros del *light-field* coinciden con los empleados en la Sección 2.4.

2.7. Sumario y Conclusiones

Las limitaciones propias de los sistemas de imagen integral motivan el surgimiento de alternativas como el sistema multicapa, que busca solucionar dichos inconvenientes. En este Capítulo se ha realizado un recorrido sobre las implementaciones más comunes de este tipo de dispositivos.

Se ha presentado una metodología basada en los algoritmos presentes en la bibliografía (Wetzstein et al., 2012c) para mejorar en términos de calidad la representación del *light-field* sobre el sistema multicapa. Dicha metodología se basa en la búsqueda de una solución inicial basada en la profundidad de la escena a representar, y su implementación sobre el WNMF. Se ha realizado un análisis cuantitativo de los resultados y se ha comparado con otros métodos propuestos en la literatura. Comparativamente, la propuesta basada en profundidad supera a los demás casos, si bien se ha observado un cierto grado de dependencia con la escena analizada.

Debido a los productos intermedios de las reglas multiplicativas empleadas para la descomposición del WNMF, se ha presentado una alternativa basada en una reordenación de los términos. Esta metodología permite realizar la misma operación, limitando el uso de memoria al número de elementos del *light-field* y las capas. Asimismo, este reordenamiento permitiría teóricamente una implementación en *hardware* de forma paralela.

Finalmente, se ha implementado un sistema propio, similar a los ya mostrados en la literatura. No obstante, este se ha diseñado de tal forma que las láminas posterior y anterior del

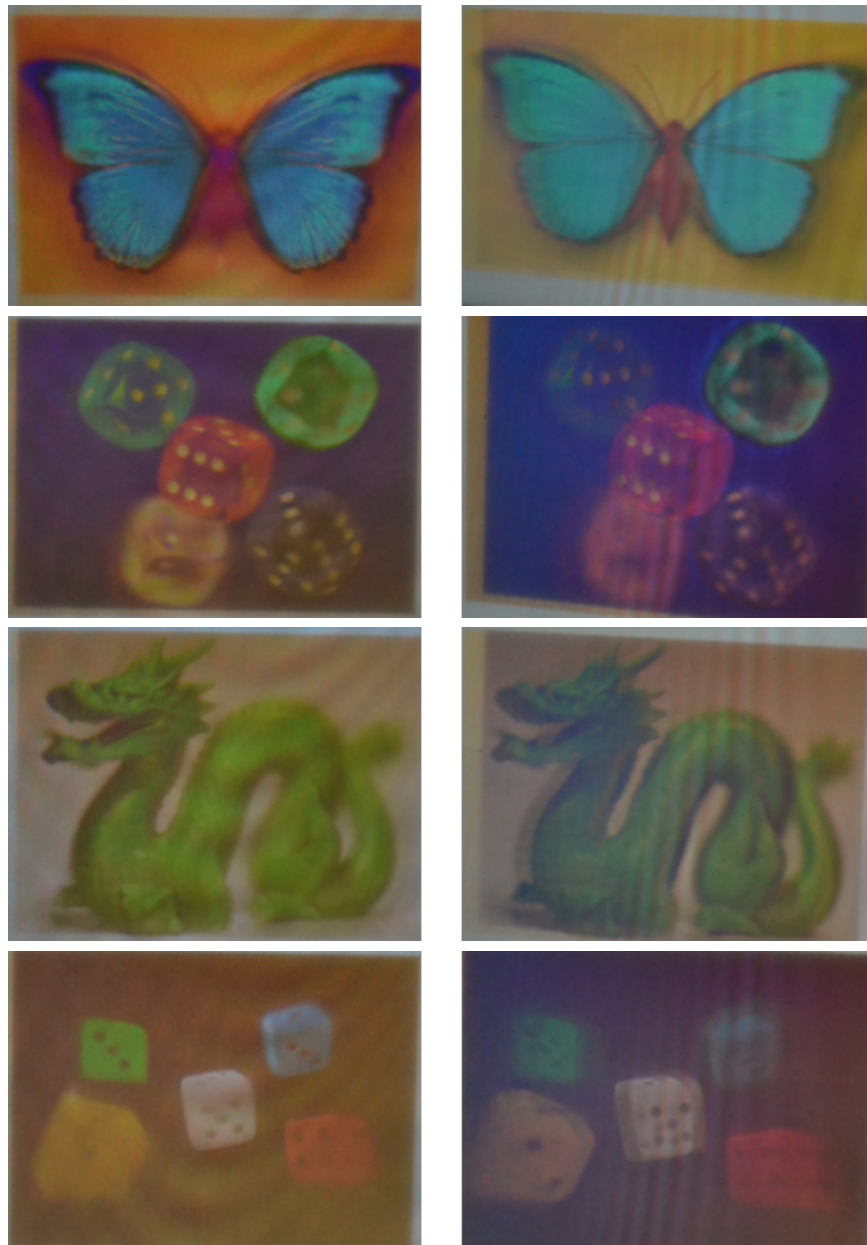


Figura 2.20 De izquierda a derecha, vista frontal y lateral obtenidas a partir de la representación del *light-field* sintéticos sobre el sistema propuesto. Las imágenes se han tomado con 0 y 10grados con respecto a la normal del sistema.

sistema cuentan con un movimiento de 6 grados de libertad. Esto ha permitido diseñar una metodología simplificada de calibrado con el fin de alinear las láminas apropiadamente. De manera alternativa, la topología planteada puede ser empleada para otros fines, como añadir un desplazamiento relativo entre pantallas. Sobre el sistema diseñado se han presentado un conjunto de *light-field* tanto sintéticos como reales, validando la metodología propuesta.

Capítulo 3

Sistema multicapa para la prevención de *shoulder-surfing*

No nos hagamos pues ilusiones: si la vida de un hombre basta para saber algo de todas las disciplinas humanas, apenas es suficiente para dominar hasta el detalle una o dos de ellas.

– Santiago Ramón y Cajal

En este Capítulo se estudia la aplicación del sistema multicapa en el ámbito de la privacidad. A diferencia de lo definido anteriormente, el objetivo no se encuentra en la representación de información tridimensional, sino en el aprovechamiento del ancho de banda de un sistema de estas características para discriminar entre observadores y en definitiva limitar el acceso al contenido mostrado desde ciertos puntos de vista.

Publicaciones relacionadas

Los resultados presentados en el presente Capítulo se corresponden con las comunicaciones en los congresos *SPIE Defense + Commercial Sensing 2019* y *ESODM2019* respectivamente.

- David Carmona-Ballester, Juan M Trujillo-Sevilla, Lara Díaz-García, Daniel Walo, Ángela Hernández-Delgado, Juan J Fernández-Valdivia, Óscar Casanova-González, Óscar Gómez-Cárdenes, y José M Rodríguez-Ramos. Privacy-enabled displays. volume 10666, page 1066609, 2018b. URL <https://doi.org/10.1117/12.2306029>
- David Carmona-Ballester, Juan M. TrujilloSevilla, Sabato Ceruso, y Oscar Casanova-Gonzalez. Shoulder-surfing prevention on multilayered light-field displays. In *ESODM-2019*, 2019c

3.1. Protección de privacidad y *Shoulder-surfing*

Recientemente ha surgido el término de *shoulder-surfing* en la literatura científica (Ali et al., 2014; Eiband et al., 2017). Este hace referencia a la observación furtiva de un contenido representado en un sistema de visualización y cuyo destinatario está acotado a un observador en particular. En la actualidad, más aún debido a la prevalencia de los dispositivos inteligentes, es prácticamente inevitable que un usuario se vea afectado por una situación de este tipo, que en muchos casos puede ser puramente accidental. Precisamente por ello, surge la necesidad de evaluar técnicas que permitan garantizar la privacidad de un contenido y la representación de este de una manera selectiva y segura.

Una de las preocupaciones esenciales que subyacen a esta cuestión es la adquisición de información sensible como son las contraseñas. Algunos trabajos han abordado esto desde un punto de vista de detección, esto es, si se conoce de la presencia de un observador no autorizado, los contenidos distribuidos por el sistema de visualización quedan bloqueados para todos los sujetos o sencillamente indicando un aviso (Zhou et al., 2016). Otra posibilidad se encuentra en el enmascaramiento de la pantalla en torno a las zonas en las que el sujeto haya centrado su vista, y por tanto limitando el contenido mostrado a observadores no autorizados (Khamis et al., 2018). Estas soluciones, si bien efectivas dentro de un término razonable, no buscan resolver el problema de raíz. Asimismo, otras técnicas incluyen la adición de sistemas secundarios encargados de mostrar parcialmente la información al observador (Winkler et al., 2015).

Alternativamente, este problema se ha enfocado tomando en consideración la interacción de las personas con un dispositivo común. A partir de múltiples trabajos, Vogel y Balakrishnan (2004) definen diferentes representaciones sobre el mismo sistema dirigidas a usuarios en función de distancias, interés o contenido de la información representada.

Los métodos anteriormente expuestos son capaces de cumplir con el objetivo de controlar el contenido mostrado. No obstante, estas metodologías están limitadas en cuanto a que su modo de operación es destructivo. Esto quiere decir que el observador debe condicionar la calidad o incluso el contenido de la información representada a cambio de garantizar un umbral de privacidad. Una manera de reducir este intercambio se encuentra en el uso de sistemas de *light-field*. Esto se debe, a que con este tipo de sistemas, se añade una componente angular, y en consecuencia, es posible controlar dentro de unos márgenes el contenido distribuido en ciertas direcciones del espacio. Se puede considerar que la primera vez que se hace uso de este tipo de sistemas con esta finalidad se encuentra en el *Doble Retrato del Rey Frederick IV y la Reina Louise*. Esta obra, del pintor Gaspar Antoine de Bois-Clair (Figura 3.1) emplea una técnica similar a las barreras de paralaje para mostrar selectivamente a cada monarca de acuerdo a la posición del observador (Shickman, 1977).

Particularmente, Mun-Kyu et al. (2016) propone una metodología para limitar la adquisición de la clave de acceso a un terminal con barreras de paralaje (Ives, 1903). Paralelamente, se propone el uso de un sistema de imagen integral para la encriptación de información sensible por medio de inversión de la función plenóptica (Yiqun, 2018).



Figura 3.1 Doble Retrato del Rey Frederick IV y la Reina Louise

La aplicación de un sistema basado en barreras de paralaje limita en gran medida el contenido que se puede mostrar, debido al ancho de banda que estos imponen Chai et al. (2000); Zwicker et al. (2006). Esto no debe ser necesariamente un problema si el contenido a representar no sobrepasa dicho límite. Sin embargo, si la información representada incluye frecuencias superiores a éste, la solución no sería válida. Es por ello que la aplicación del sistema multicapa, el cual excede en gran medida esta limitación, puede ser una alternativa.

3.2. Sistema multicapa para privacidad

A pesar de que el sistema multicapa surge con el objetivo de representar información tridimensional (Gotoda, 2010), es precisamente la aparición de una discretización angular la que permite trasladarlo a otros ámbitos como el de la privacidad. Este tema ha sido previamente estudiado en la bibliografía de sistemas de representación de *light-field*, donde es posible representar información no relacionada para cada ángulo discreto del sistema.

En el caso del sistema multicapa, existen pocas referencias al tema en cuestión, encontrándose estas principalmente en los trabajos más relevantes Wetzstein et al. (2011, 2012b). Esto puede en cierta manera venir motivado por la necesidad de una correlación aparente entre las vistas del sistema. En los trabajos anteriormente citados sobre imagen integral, la representación de cada vista es, idealmente independiente del resto, ya que para cada vista definida en el plano (u, v) del *light-field*, se corresponde un subconjunto de píxeles contenidos en el plano (x, y) , independientes unos de otros. Sin embargo, en el caso de multicapa, la representación de la función de *light-field* resulta por naturaleza más compleja, ya que los píxeles de cada capa $f_m^{(n)}$ subtienden a las capas anteriores y posteriores, y a su vez son afectados por sus vecindades.

Esta Sección evalúa la técnica de una forma alternativa. En este sentido, se busca modificar la señal de entrada con el objetivo de que el *light-field* a representar guarde un criterio de correlación, pero a su vez permita definir sus proyecciones como bloqueadas y no bloqueadas.

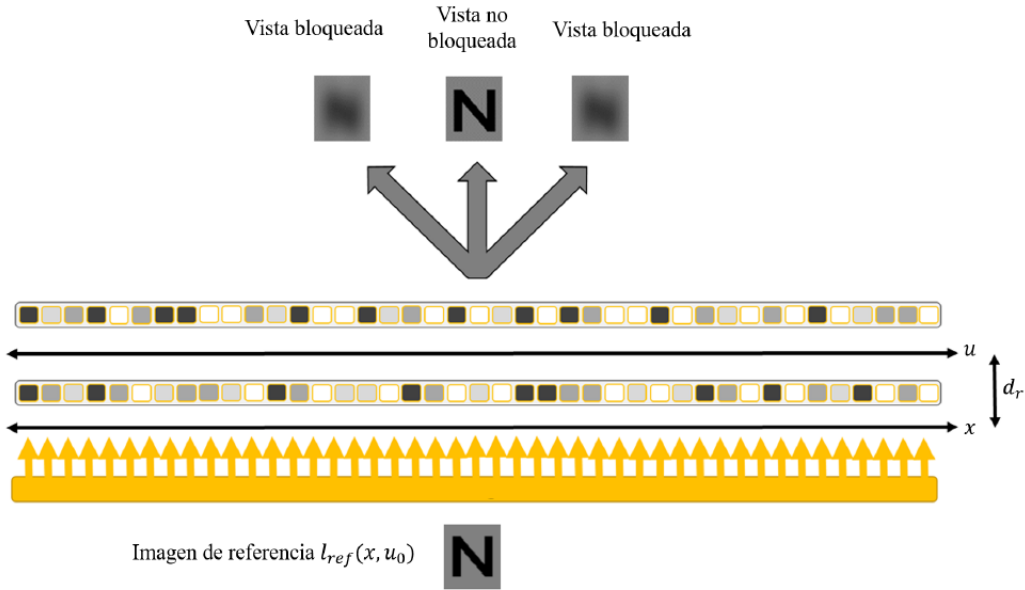


Figura 3.2 Referencia del sistema multicapa aplicado a privacidad

3.2.1. Criterio de privacidad

Siguiendo la representación del *light-field* en dos planos indicada en el Capítulo 2, se define el siguiente criterio de privacidad:

$$l_{bloqueado}(x, u_0) = \begin{cases} l_{ref}(x, u_0) & \text{si } (u_0) \text{ no está bloqueado,} \\ l_{ref}(x, u_0) * h(x, u_0) & \text{si } (u_0) \text{ está bloqueado} \end{cases}, \quad (3.1)$$

donde l_{ref} es el *light-field* de referencia, mientras que $l_{blocked}$ es dicho *light-field* una vez se ha aplicado una función $h(x, u_0)$ que distinga entre las vistas bloqueadas y las no bloqueadas. En la Figura 3.2 se muestra una representación de un sistema compuesto por dos capas siguiendo la definición anterior.

Es necesario, por tanto, definir una función que modifique el *light-field* de tal manera que un observador que perciba una vista bloqueada, no pueda entender su contenido. Para realizar una selección del filtro, se ha partido de estudios previos en relación a los límites de percepción de un observador. Existe un campo de estudio amplio en relación al estudio determinístico de los límites en la percepción del sistema visual humano. Algunos estudios, como los propuestos por Alexander et al. (1994) y Christman et al. (1997) analizan como afectan diferentes tipos de filtros en el dominio de la frecuencia cuando se presenta a un observador un conjunto de caracteres. Más concretamente, Kwon y Legge (2012, 2011, 2013), evalúan de forma estadística el límite en frecuencia en el que dicho observador deja de ser capaz de distinguir una letra de otra. En dichos estudios, se parte de un filtro de *Butterworth* de orden 3, determinando un valor mínimo de 0.9ciclos/letra (CPL). En el presente caso, se sigue una metodología idéntica, eligiendo dicho filtro como la función $h(x, u)$, permitiendo de esta

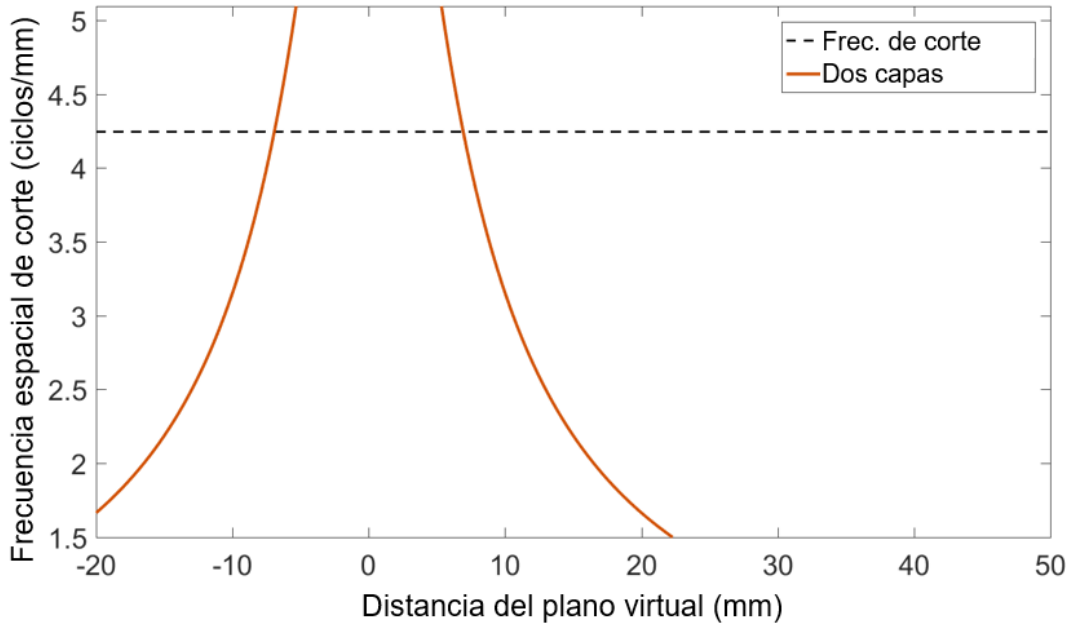


Figura 3.3 Límite superior en la profundidad de campo obtenida para un sistema de dos capas. La línea discontinua representa la frecuencia de corte de cada una de las capas que componen el sistema.

forma, garantizar –en relación a dichos estudios– la discriminación entre vistas bloqueadas y no bloqueadas. La función del filtro se puede expresar de la siguiente forma:

$$H_{butterworth} = \frac{1}{1 + \left(\frac{d}{r}\right)^{2n}}, \quad (3.2)$$

siendo r la frecuencia radial de corte, $d = \sqrt{x^2 + y^2}$ y n el orden del filtro. Como en los trabajos anteriormente mencionados, se parte de $n = 3$. Se busca que, al sintetizar las capas del sistema, las proyecciones bloqueadas mantengan dicho filtro, y por tanto, nieguen a un observador puntual su percepción.

Precisamente, se propone el uso de un filtro pasa-bajas en detrimento de otros debido a la naturaleza de la representación en un sistema de imagen tridimensional. A continuación se profundizará sobre esta idea, así como se presentarán las consideraciones esenciales a la hora de definir el *light-field* en el ámbito de la privacidad.

3.2.2. Profundidad de campo del sistema de dos capas

En un sistema multicapas se define la profundidad de campo como la máxima frecuencia espacial alcanzable en un plano a una cierta distancia del sistema Wetzstein et al. (2012a). Esta definición permite, más concretamente, tener en cuenta el tipo de filtro que debe aplicarse sobre un *light-field* previo a su representación en un sistema multicapa. En relación al objetivo de la privacidad,

$$|f_c(d_0)| \leq \frac{1}{p} \sqrt{\frac{3\Delta d^2}{3\Delta d^2 + 12d_0^2}}, \quad (3.3)$$

siendo p el tamaño del píxel de cada capa, y Δd la separación entre estas, y d_0 como la distancia al plano virtual. A pesar de que la Ecuación 3.3 da lugar a un valor de frecuencia $f_c = 1/p$, a su vez, existe una limitación dada por el propio tamaño del píxel, y que siguiendo el teorema de *Shanon-Nyquist* es equivalente a $1/2p$. En la Figura 3.3, se muestra la curva definida para un sistema de dos capas con una separación de píxeles de $118\mu\text{m}$. Como se observa, existe un área debajo de la curva en la que se conserva la frecuencia de muestreo impuesta por cada pantalla. Sin embargo, al alejarse del punto central, se experimenta un decrecimiento en las frecuencias que pueden representarse. Este comportamiento no es único en los sistemas multicapa, sino que se extiende a otros sistemas de imagen tridimensional, como son el caso de la imagen integral (Chai et al., 2000; Zwicker et al., 2006), y tiene origen esencialmente en el intercambio de frecuencias espacio-angulares. Sin embargo, y en comparación con la imagen integral, los sistemas multicapa muestran un equilibrio entre resolución angular y espacial superior (Wetzstein et al., 2012c).

Reordenando la Ecuación 3.3 en función de la separación de capas, se obtiene que:

$$\Delta d(f_{\text{ext}}, d_0) \geq 2pf_{\text{ext}}(|d_0| + 2\zeta) \sqrt{\frac{1}{1 - (pf_{\text{ext}})^2}}, \quad (3.4)$$

siendo ζ el límite físico impuesto por el grosor de las placas del sistema y f_{ext} como la frecuencia de corte del texto a representar. Particularmente, en este caso, se introduce un coeficiente de valor 2 en relación al número de pantallas. Si se elige un $f_{\text{ext}} = 1/2p$ de tal manera que se alcance el límite total impuesto por el tamaño de píxel de la pantalla, se puede obtener un valor de separación óptimo para el sistema multicapa:

$$\Delta d(d_0) \geq \frac{2}{\sqrt{3}}(|d_0| + 2\zeta) \quad (3.5)$$

Siguiendo el criterio definido en la Expresión 3.3, el punto de máxima frecuencia se obtiene cuando el plano virtual se sitúa en el centro del sistema, esto es, a una distancia $d_0 = 0$. La Expresión 3.5 define el límite de proximidad al que deben situarse las capas en función de la distancia a la cual se proyecta el plano virtual si se quiere conservar la totalidad de las frecuencias representadas para un texto dado.

3.3. Evaluación de la técnica

Es necesario definir una métrica que permita discernir entre una imagen percibida correctamente o no por un observador dado. Como se ha introducido anteriormente, se parte de las conclusiones obtenidas en los trabajos de Kwon y Legge (2011). Sin embargo, se asume que al introducir el filtro en frecuencia en las vistas bloqueadas garantiza su presencia una vez se ha

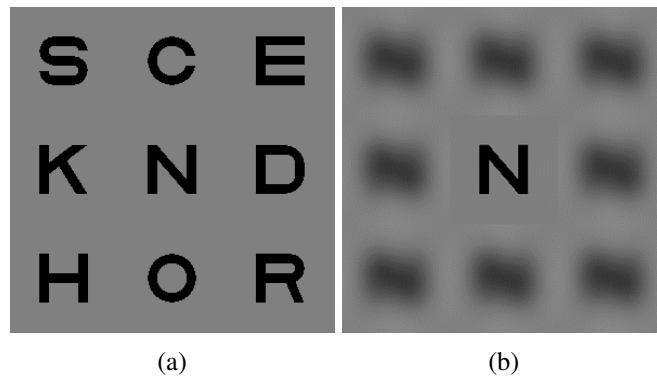


Figura 3.4 *Light-field* de entrada; (3.4a) se corresponde con el caso NC mientras (3.4b) es para el caso FL. La vista central se define como no bloqueado, mientras que las laterales se definen como bloqueadas.

reconstruido el *light-field* a través de la síntesis en capas. La validación de la técnica parte de evaluar el grado de conservación del filtro a la vez que se garantiza la reconstrucción de las vistas no bloqueadas.

3.3.1. Recuperación del filtro impuesto como función de mérito

La metodología propuesta encuentra como métrica acorde a los resultados la recuperación del filtro estimado introducido en el *light-field* de entrada. Esta operación puede definirse como sigue:

$$H_{butterworth} \approx H_{rec} = \frac{\mathcal{F}\{l_o\}}{\mathcal{F}\{l_{ref}\}}, \quad (3.6)$$

siendo $\mathcal{F}\{\}$ es el operador de transformada de Fourier.

Este planteamiento ya se menciona previamente en la bibliografía como método para evaluar la conservación de un cierto filtro una vez aplicada la función de transferencia del sistema (Goodman, 2005). En este caso se considera que la función de transferencia del sistema contiene el proceso de síntesis de las capas y su posterior proyección.

3.4. Resultados experimentales

Con el fin de analizar el comportamiento de la metodología propuesta sobre el sistema multicapa, se ha evaluado un conjunto de letras generadas siguiendo un planteamiento similar a Kwon y Legge (2013). Se ha generado un conjunto de letras de *Sloan* (Sloan, 1959). A cada letra se le ha asignado un tamaño de 45×45 píxeles. A su vez, cada letra se ha solapado sobre un fondo gris (valor de 127 sobre 255) de 250×250 píxeles. Se han generado dos *light-field* diferentes de 3×3 vistas y campo de visión de 20° . Para el primero, se ha asignado una letra diferente para cada vista, de tal manera que no haya correlación entre estas. Alternativamente, las vistas del segundo *light-field* se corresponden con el mismo carácter, con la diferencia de que las

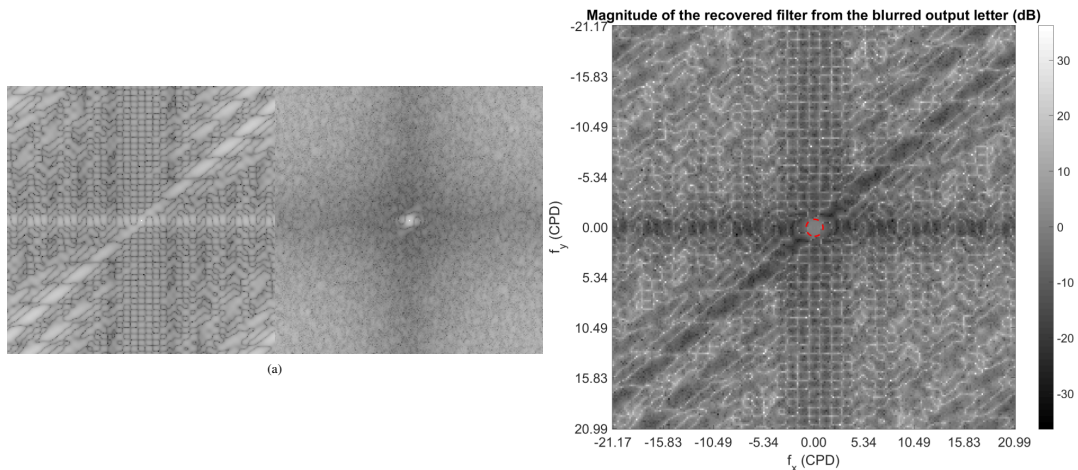


Figura 3.5 Magnitud espectral tras sintetizar el *light-field* tras aplicar el filtro correspondiente. A la izquierda se muestra una comparativa del espectro de las vistas recuperadas para letra bloqueada y no bloqueada. A la derecha, relación de magnitudes obtenida entre los espectros de letra bloqueada y no bloqueada respectivamente de acuerdo a la Ecuación 3.6

vistas laterales se les ha aplicado el filtro de *Butterworth* (véase Ecuación 3.2) siguiendo los requisitos expuestos en la bibliografía para garantizar el criterio de privacidad, es decir, que el filtro introducido se conserve en la imagen percibida. Una representación de ambos casos se puede observar en la Figura 3.4. Para simplificar, denominaremos NC al caso no correlado y FL al caso filtrado.

La metodología propuesta se ha estudiado sobre un sistema de dos capas. La síntesis de las capas se ha realizado por medio del algoritmo WNMF. Se ha aplicado la métrica propuesta sobre las proyecciones de las capas en simulación. Asimismo, con el fin de mantener la claridad, los resultados mostrados para el caso LF serán específicamente para la letra *n*. No obstante, el comportamiento obtenido es equivalente para cada uno de los caracteres estudiados, siendo estos *s*, *c*, *e*, *k*, *d*, *h*, *o*, *r*.

En la Figura 3.5 se muestra el filtro recuperado para el caso FL sobre una de las vistas bloqueadas. A la izquierda se representa la magnitud de la vista no bloqueada y bloqueada respectivamente. Junto a estas, se encuentra el resultado tras aplicar la Ecuación 3.6. Se ha trazado una circunferencia en torno a la frecuencia de corte del filtro. Como se puede observar, en esta zona la magnitud del filtro se mantiene igual con respecto al filtro de *Butterworth*. Sin embargo, surge información nueva en frecuencias diferentes, por lo cual es necesario evaluar si éstas realmente se corresponden con información que pudiera permitir al observador no autorizado discernir el carácter. En la Figura 3.6 se muestra el promedio radial de la magnitud recuperada. Se puede observar que la magnitud del filtro hasta la frecuencia de corte se conserva completamente. No obstante, las contribuciones a lo largo del espectro más allá de estas frecuencias excede dicho valor.

De forma ejemplar se muestran en la Figura 3.7 los resultados obtenidos para la letra *n*. La primera fila, de izquierda a derecha contiene la imagen de referencia, la vista no bloqueada y

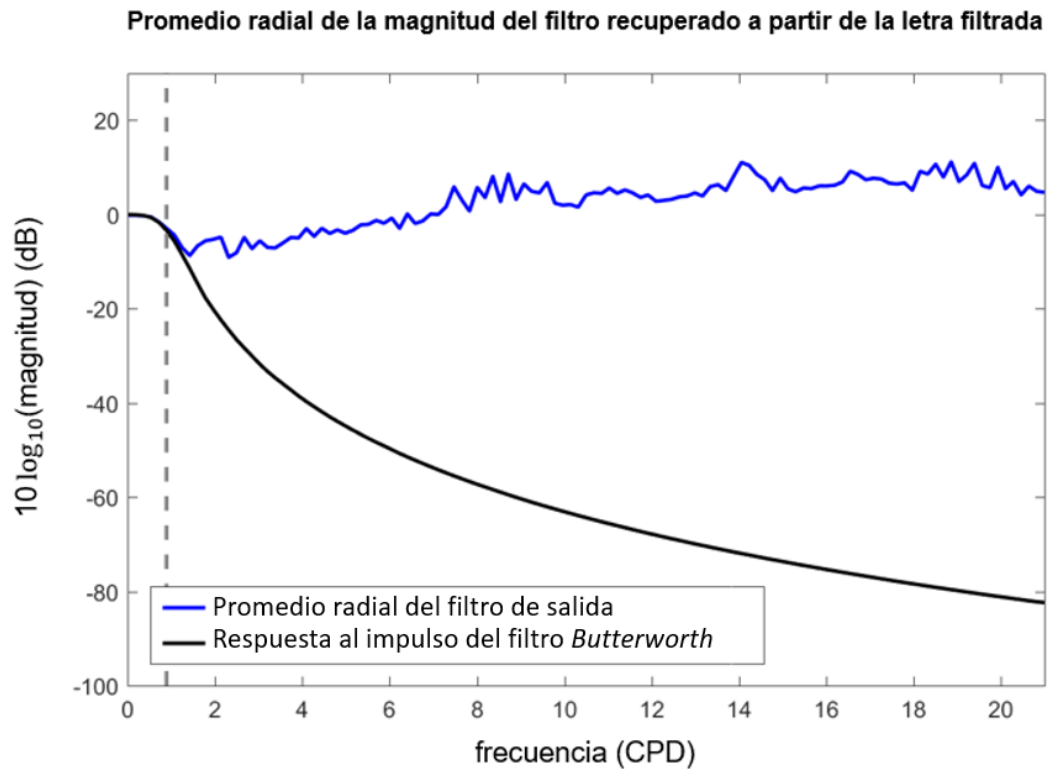


Figura 3.6 Promedio radial de la magnitud en decibelios del filtro recuperado a partir de la letra desenfocada en relación al filtro de *Butterworth*. La línea discontinua representa la frecuencia de corte de la letra introducida.

bloqueada respectivamente. Un primer vistazo parece confirmar que el criterio de privacidad se cumple, siguiendo las conclusiones obtenidas previamente. Para identificar en el dominio espacial al cual se corresponden las frecuencias espurias observadas, se ha aplicado un filtro paso alto sobre la letra bloqueada, correspondiente a la inversa del filtro de *Butterworth* propuesto. El resultado de esta operación se muestra en la segunda fila de la Figura 3.7. De izquierda a derecha se muestra la letra de referencia, la vista no bloqueada, y la vista bloqueada. Estas dos primeras imágenes permiten fácilmente identificar el carácter propuesto. Sin embargo, en la tercera imagen se observan varias fuentes de ruido en torno al área de interés. De estas fuentes podemos asociar el ruido de fondo a la descomposición multimodal del WNMf, siendo más prominente éste en la letra bloqueada. La fuente de ruido central se corresponde con la aproximación realizada por el sistema a partir de los valores de inicialización para ajustar un desenfoque en la zona de interés.

3.4.1. Interdependencia entre filtro y número de *frames*

Una vez determinada la viabilidad de la técnica sobre la topología propuesta, resta evaluar el comportamiento del sistema para diferentes puntos de operación. En el caso de estudio presentado, se plantean dos puntos esenciales a analizar; el grado en el que se limita la

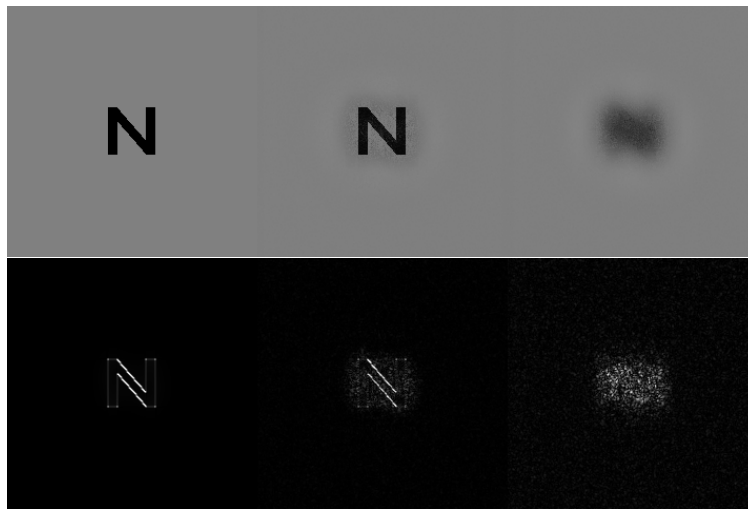


Figura 3.7 Resultado obtenido para la vista bloqueada y no bloqueada en el caso FL con un filtro de 0,9CPL y $M = 4$. La fila superior muestra de izquierda a derecha el carácter de referencia, la vista no bloqueada y bloqueada respectivamente. La fila inferior contiene el resultado de aplicar la inversa del filtro empleado sobre cada uno de los casos anteriores.

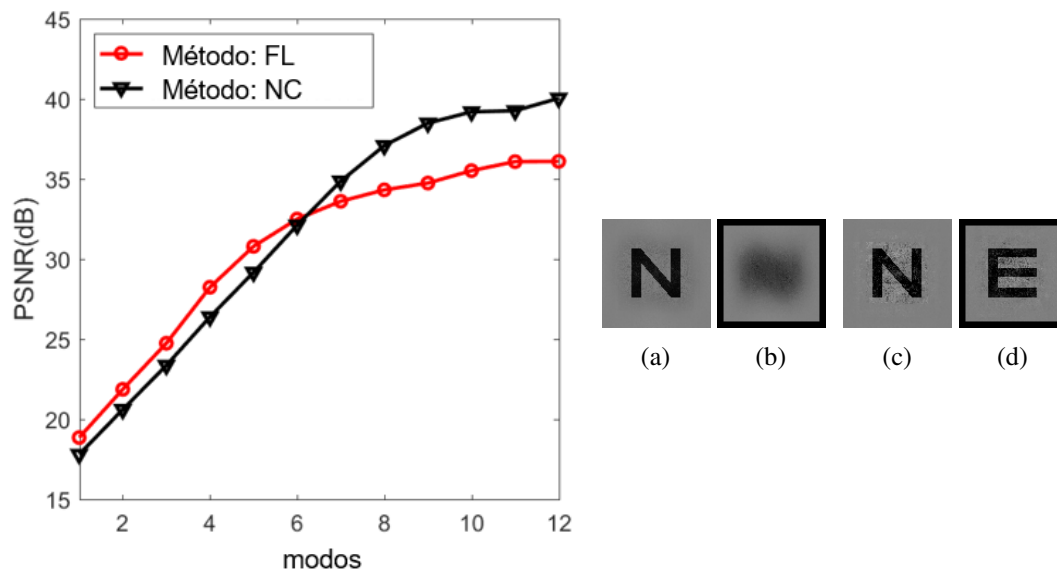


Figura 3.8 Tendencia del PSNR para ambas técnicas propuestas tras incrementar el rango de la descomposición. A la derecha de la imagen se muestran los resultados obtenidos para los casos FL (3.8a-3.8b) y NC (3.8c-3.8d) respectivamente. Estos resultados se corresponden concretamente con la descomposición en cuatro instantes de tiempo.

percepción de las vistas bloqueadas, y la calidad en la representación de aquellas no bloqueadas. La primera cuestión ya se ha evaluado por medio de la obtención del filtro impuesto como función de mérito. No obstante, no se ha tomado la calidad de la reconstrucción en términos generales. Como se observaba en la Figura 3.7, incluso la vista no bloqueada adolece de un cierto grado de ruido o imperfecciones. Por ello, se evaluará la tendencia de la calidad tanto para vistas bloqueadas como no bloqueadas en términos del PSNR, significando esto la cercanía al *light-field* de referencia.

Siguiendo la bibliografía previa, (Lanman et al., 2010; Wetzstein et al., 2012a), el modos M de la descomposición afecta positivamente al error cometido en la descomposición del tensor, o en este caso, la matriz de estudio. Se hace necesario evaluar la progresión en términos del error para las vistas no bloqueadas. Por simplicidad, se han reducido las dimensiones de las vistas de 250×250 a 100×100 . En la Figura 3.8 se muestra la progresión de esta métrica en función del rango de la descomposición. Tanto para el caso NC como FL se observa una tendencia creciente, siendo este un comportamiento esperable. No obstante, la alternativa del filtro muestra un incremento del PSNR al menos para los primeros 6 modos, si bien a partir de este punto se ve superada por el caso NC. De manera ejemplar, se muestran junto a la gráfica las imágenes obtenidas para cada caso de vista bloqueada y no bloqueada. Notablemente, se puede indicar que ambas metodologías alcanzan el criterio de privacidad, si bien efectivamente lo mostrado por el caso FL da lugar a una representación más limpia del carácter estudiado. Se teoriza que este comportamiento puede deberse precisamente a la similitud en espectro del *light-field* en baja frecuencia para FL mientras que en el caso NC existe una diferencia superior, y de ahí que la descomposición tienda a equilibrar la representación de ambas letras.

Finalmente se ha evaluado de manera general la calidad de la reconstrucción a partir de las variables libres del sistema, esto es, el radio del filtro aplicado y el rango de la descomposición. Esto se ha hecho obteniendo nuevamente el PSNR para el caso de LF en vistas bloqueadas y no bloqueadas. En la Figura 3.9 se muestran los mapas obtenidos. A la izquierda se evalúa el valor de PSNR en función de dichas variables libres. Con el fin de ilustrar los resultados obtenidos, se han seleccionado valores representativos sobre estos mapas, representados a la derecha de la Figura 3.9. Se puede observar la tendencia de mejora según aumenta el rango de la descomposición (primer bloque a la derecha), donde para valores bajos, el carácter es reconocible llegando incluso a generar imágenes duplicadas. Contrastando los resultados observados en la Figura 3.8, a partir de $M = 4$ se comienza a obtener una representación correcta de la letra filtrada. A pesar de que incrementar este valor da lugar a una mejoría observable, es posible considerar este como el punto de operación en el que el sistema propuesto alcanza el criterio de privacidad. Por comparación, el valor de PSNR de la vista no bloqueada se encuentra en el rango de los 40db, considerado en la bibliografía un valor aceptable en términos de calidad para una compresión con pérdidas Welstead (2009).

De manera similar, en el segundo bloque a la derecha de la Figura 3.9 se muestra la influencia del cambio de filtro sobre el resultado obtenido. A pesar de que esta evaluación no tiene como objeto solucionar el problema de la privacidad, sirve para observar como la

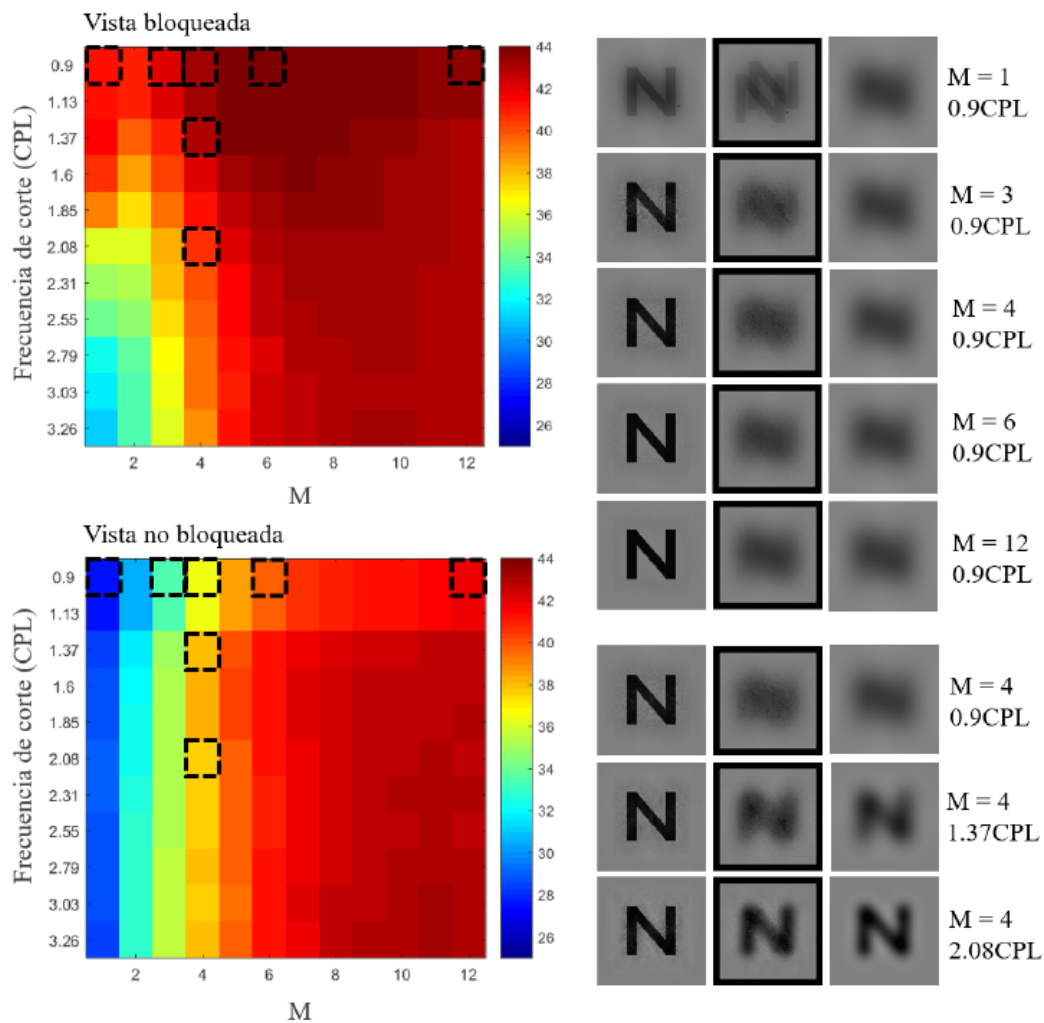


Figura 3.9 Tendencia del PSNR para las vistas bloqueadas y no bloqueadas en el caso LF. La columna izquierda contiene los mapas obtenidos en función del rango de la descomposición y la frecuencia de corte del filtro impuesto. Sobre los mapas se han resaltado varios casos puntuales. A la derecha se muestran estos casos. De izquierda a derecha, vista no bloqueada, vista bloqueada, y letra filtrada como referencia. A su vez, se agrupan dos bloques, en base a la variable controlada, esto es, M o frecuencia de corte.

calidad de la reconstrucción se reduce según el filtro aplicado permite pasar un mayor número de frecuencias. Este comportamiento encuentra sentido en las premisas iniciales, donde se asumía que precisamente lo que permite reducir los criterios técnicos del sistema, esto es, el rango de la descomposición, es la correlación espacial entre las frecuencias bajas de la imagen. De manera más general, este comportamiento se puede observar en la simetría inversa entre vista bloqueada y no bloqueada en el eje de frecuencia.

3.4.2. Implementación sobre modelo

Para demostrar la viabilidad de la metodología estudiada, se ha realizado una implementación de la misma sobre el dispositivo desarrollado en el Capítulo 2. Debido a las limitaciones del dispositivo, se ha representado sobre éste el caso particular de privacidad sobre dos capas y una descomposición en cuatro modos o instantes de tiempo. Para actuar sobre el sistema como si estuviera compuesto por solo dos capas, es necesario introducir una imagen con todos los valores a cero en la capa intermedia, ya que de esta manera introduce una rotación de 0 grados sobre la luz que la atraviesa. En la Figura 3.10 se muestra el resultado capturado haciendo uso del mismo montaje empleado en el Capítulo 2. Asimismo, se ha hecho uso de un tiempo de exposición en la cámara suficientemente elevado como para evitar efectos indeseados, esto es 1s. La primera fila se corresponde con las imágenes observadas para el caso FL, para el caso bloqueado y sin bloquear respectivamente. Como se observa, los resultados se asemejan a lo obtenido en simulación, si bien existe una reducción en contraste debida a la propia implementación del sistema. Las dos filas inferiores se corresponden con los instantes de tiempo para cada capa, siendo la superior la capa delantera, y la inferior la trasera.

Con fines ilustrativos, se ha elaborado un vídeo del sistema funcionando en el caso anteriormente expuesto. En este se simula el movimiento de un observador entre las posiciones bloqueadas y no bloqueada, donde se puede apreciar el efecto de privacidad sobre las mismas. El acceso a este contenido puede realizarse por medio del siguiente enlace:

- http://www.cafadis.ull.es/videos/PrivacidadMulticapa_LetraN.mp4

3.5. Sumario y Conclusiones

Una alternativa a la representación tridimensional sobre un sistema de *light-field* se puede encontrar en la discriminación de observadores para limitar los efectos del conocido *shoulder-surfing*. La naturaleza angular de estos dispositivos permite controlar en un sentido práctico la información distribuida en diferentes direcciones. El trabajo realizado en este Capítulo se ha centrado en evaluar los límites técnicos de un sistema multicapa siendo éste empleado para dicho fin. Asimismo, se ha presentado una metodología que permite reducir los requisitos técnicos a nivel de *hardware* en comparación a metodologías más tradicionales. Este análisis se ha llevado a cabo en simulación, lo que ha permitido evaluar el comportamiento para diferentes

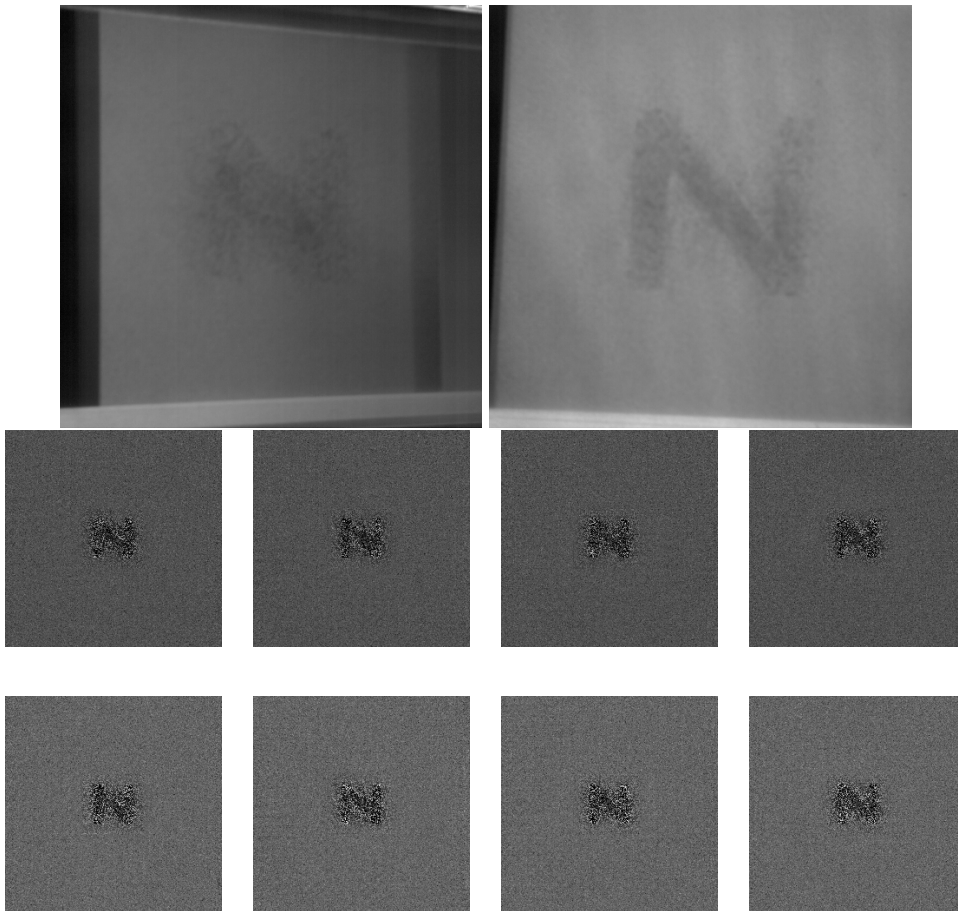


Figura 3.10 Captura del método de privacidad sobre el sistema implementado. La fila superior se corresponde con las vistas bloqueada (izquierda) y no bloqueada (derecha). Las dos filas inferiores se corresponden con los cuatro instantes de tiempo de la capa delantera y trasera del sistema respectivamente, siendo estos correspondientes de forma vertical.

casos de estudio desde el punto de vista de calidad en la reconstrucción como cumplimiento de un criterio de privacidad establecido.

El método propuesto se basa en la introducción de un filtro de frecuencia pasa-bajas en la entrada del sistema, de tal manera que las vistas percibidas por un observador no autorizado sean irreconocibles. Debe entenderse lo anterior como una primera prueba de concepto que permita abrir esta dirección en futuros trabajos. No obstante, el análisis realizado se ha limitado a un conjunto de estudio basado en caracteres predefinidos a fin de evaluar de la manera más apropiada un comportamiento sujeto a la percepción humana. Estos resultados se limitan a texto, dejando abierta una vía de estudio a un campo más amplio como es el de la imagen. Existe un amplio campo de investigación centrado en los mecanismos cognitivos del ser humano asociados a la percepción y discriminación entre imágenes (DiCarlo et al., 2012; Riesenhuber y Poggio, 1999). Precisamente, a partir de estos modelos, surgen implementaciones basadas en estas ideas con el fin de simular dichos procesos (Fan et al., 2014; Serre et al., 2005, 2007). Una línea futura de investigación podría tomar como punto de partida la evaluación adaptativa de las imágenes mostradas, dando pie a búsquedas heurísticas de la función filtro. Alternativamente, se podría también seguir un modelo basado en generación de imágenes por medio de redes neuronales generativas adversarias o *generative adversarial networks*.

Capítulo 4

Adaptación de WNTF a tomografía atmosférica

Menester es, pues, fiar algo a la casualidad, provocándola mediante una serie reiterada de tanteos, en los cuales no podemos ser guiados más que por la intuición auxiliada por el conocimiento, todo lo profundo y preciso posible, de los reactivos y procederes técnicos recién introducidos en la Química y la Industria.

– Santiago Ramón y Cajal

Como aplicación alternativa, en este Capítulo se presenta el método WNTF empleado en los Capítulos 2 y 3 en el campo de la astronomía, más concretamente el de la óptica adaptativa –del inglés *adaptive optics* (AO)–. En primer lugar, se dará un marco teórico sobre el área de aplicación y los motivos que llevan a trasladar una técnica como el WNTF a éste. A continuación, se mostrarán los pasos necesarios para adaptar dicho método a la síntesis de los gradientes de fase normalmente empleados en AO y se comparará con un método tradicionalmente usado como es el *Kaczmarz iterativo*. Para finalizar, se realizará un análisis exhaustivo de la metodología aplicada para validar la técnica en simulación.

Publicaciones

Los resultados del presente Capítulo se corresponden con la publicación realizada en la revista *Astronomy and Astrophysics*.

- David Carmona-Ballester, Juan M. Trujillo-Sevilla, Sergio Bonaque-González, Óscar Gómez-Cárdenas, y José M. Rodríguez-Ramos. Weighted nonnegative tensor factorization for atmospheric tomography reconstruction. *Astronomy and Astrophysics*, 614 (A41), 2018a. doi: <https://doi.org/10.1051/0004-6361/201832597>

4.1. Marco teórico

La AO consiste en un conjunto de técnicas que buscan mejorar la imagen obtenida por telescopios (Roddier, 1999). En esencia, la degradación de la imagen astronómica suele proceder de diferentes fuentes, siendo una de las más importantes la influencia de la atmósfera. La presencia de aberraciones dificulta el análisis y estudio de los objetos, siendo por tanto necesaria su corrección. La medida de este tipo de aberraciones se realiza por medio de un sensor de frente de onda, también denominado en inglés como *wavefront sensor* (WFS). Partiendo del conocimiento de dicha información, se puede realizar la implementación de un sistema de control –bien en lazo abierto o cerrado– por medio de la introducción, entre otros de un dispositivo capaz de compensar dichas aberraciones, siendo por ejemplo uno de estos el espejo deformable (Roddier, 1999).

Particularmente, la aplicación de estas técnicas se extiende a los futuros Telescopios Extremadamente Grandes o *Extremely Large Telescopes* (ELT), los cuales son capaces de obtener una imagen mucho más detallada del cielo observado. En estos sistemas donde la apertura es superior a otros más convencionales, la aplicación de la AO resulta en una corrección espacialmente más acotada. De manera alternativa, existen un conjunto de técnicas agrupadas en la llamada óptica adaptativa de campo extenso o *wide-field adaptive optics* (WFAO). Con esto se intenta incrementar el área efectiva del cielo en la cual se pueden corregir de manera apropiada los efectos de la atmósfera (Le Louarn et al., 2000). En este sentido, es necesario evaluar de forma computacional la información remitida por el WFS para diferentes ángulos y obtenida a partir de estrellas artificiales de referencia. De esta manera es posible obtener un perfil tridimensional de la atmósfera en estudio, y entonces corregir de manera más efectiva sus efectos en la imagen muestreada. A este procedimiento se le conoce en la bibliografía como tomografía atmosférica (AT) Ramlau y Rosensteiner (2012); Ramlau et al. (2014).

Es sencillo encontrar en el marco del presente trabajo la similitud con los sistemas de representación tridimensional. Si bien en este caso no se busca la representación de dicho perfil sobre un sistema de visualización, su obtención parte de principios muy similares. No en vano, en su gran mayoría los métodos sobre los que se asienta la AT toman como base la teoría de tomografía computerizada. En esencia, la emisión de ondas electromagnéticas en diferentes ángulos y su posterior medición al atravesar un cuerpo, da cuenta en forma de atenuación del sistema que intermedia (Beckmann, 2006).

Comúnmente, las técnicas más socorridas en la reconstrucción tomográfica de la turbulencia atmosférica enfocan el problema desde un punto de vista algebraico, definiendo en muchos casos un sistema de ecuaciones que toma en cuenta cada una de estas proyecciones (Ellerbroek, 2002). Precisamente es la dispersión de los datos de estudio la que permite un análisis desde esta perspectiva. Comparativamente, lo anteriormente introducido muestra una evidente relación con los métodos propuestos por Gotoda (2010) y Wetzstein et al. (2011) posteriormente. Por ello, se puede asumir que una técnica favorable en uno de estos ámbitos podría trasladarse de forma eficaz al otro y viceversa.

4.2. Recuperación de fase

En los Capítulos previos se ha trabajado principalmente sobre la propagación de la luz considerando solo la intensidad como elemento para definir el *light-field*. Esto se hace precisamente asumiendo que dicha componente es la que tiene prevalencia a la hora de percibir la información tridimensional. Sin embargo, un haz de luz monocromático es una onda electromagnética, y por tanto un fasor que puede definirse de la siguiente manera:

$$U(P) = A(r) \exp(-jk\phi(r)), \quad (4.1)$$

siendo $A(r)$ y $\phi(r)$ las componentes de amplitud y fase del frente de onda en una posición r del espacio y j como el número complejo. En esta expresión k se corresponde con el número de onda, relacionado a su vez con la longitud de onda de la señal como $k = 2\pi/\lambda$.

Podemos entender el concepto de fase como el desplazamiento o adelanto local de la onda de estudio. Precisamente, esta definición nos permite entender como la intensidad emitida por un punto en el espacio se distribuye y propaga a lo largo del camino óptico. Esta definición resulta de suma importancia al realizar observaciones de los astros a través de un sistema óptico como es el caso de los telescopios, ya que la atmósfera introduce un defecto óptico o aberración que impide la correcta evaluación de estos. En el caso de la tomografía atmosférica, a diferencia de lo que se ha visto para el sistema multicapa, se parte de un conjunto de derivadas parciales de la fase observada por un cierto sensor. En la práctica, estas derivadas de fase se obtienen por medio de sensores de fase de frente de onda empleados en los propios sistemas de AO. De manera excepcional, partimos de las definiciones propuestas por Tallon y Foy (1990), donde la tomografía se divide en los gradientes de fase vertical y horizontal.

4.2.1. Kaczmarz iterativo

La representación del problema de la tomografía atmosférica suele abordarse desde un punto de vista algebraico, construyendo un sistema de ecuaciones que dé cuenta de como cada rayo, o de forma más general, el campo, se transfiere desde el objeto en estudio hasta la pupila de entrada del sistema. El origen de los algoritmos más extendidos actualmente en este campo surge del análisis realizado por el matemático polaco Stefan Kaczmarz (Kaczmarz, 1937). El modelo de la tomografía atmosférica asume que las capas que componen la atmósfera pueden ser entendidas como una función discreta de tal forma que la proyección final consiste en las contribuciones aditivas de cada uno de estos en cada ángulo muestreado (Cormack, 1964; Hounsfield, 1973). De manera similar al caso del sistema multicapa, es posible representar la atmósfera – en cuanto al problema de la tomografía– como un conjunto discreto de capas que alteran la proyección de los rayos procedentes de los astros. A título orientativo, se muestra en la Figura 4.1 un esquema del problema en cuestión. Para un conjunto de astros de referencia, se obtiene en la pupila del telescopio la proyección de cada uno de éstos. Las aportaciones de fase del volumen atmosférico contribuyen a la fase del campo observado para cada astro, dando de esta manera información acerca del perfil atmosférico.

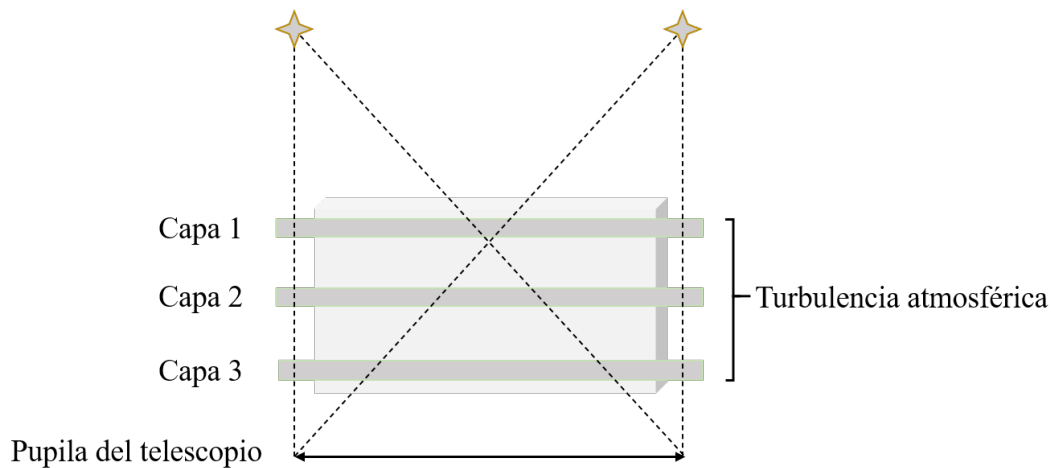


Figura 4.1 Esquema simplificado de la tomografía atmosférica. Se representa un caso particular para la tomografía atmosférica discretizando en un número de tres capas y por medio de la observación de dos objetos de estudio.

No obstante, es precisamente este razonamiento el que da lugar a un problema mal condicionado, y obliga a emplear algoritmos iterativos con el objetivo de mejorar la solución obtenida (Gordon et al., 1970; Ramlau y Rosensteiner, 2012; Ramlau et al., 2014). Como este tipo de sistemas indeterminados no cuentan con una solución única, tienen más ecuaciones que incógnitas, existen numerosas técnicas en la bibliografía basadas en la minimización de la norma euclídea.

El algoritmo de Kaczmarz se basa en la minimización del error cuadrático de las proyecciones, partiendo de los gradientes de fase de la atmósfera. Por ello, y de manera similar a lo presentado en el Capítulo 2, es posible definir un sistema de ecuaciones lineales, de tal manera que, minimizando el error, se obtenga un compromiso en relación al perfil atmosférico buscado. Esta relación es descrita en Tallon y Foy (1990), y es como sigue:

$$\arg \min \|\mathbf{A}d - Hs\|_2^2, \quad (4.2)$$

siendo A la matriz de proyecciones para los rayos de luz propagados a través de la atmósfera, s como la derivada de la fase obtenida desde cada ángulo en la entrada de pupila, H el vector de alturas correspondiente a cada capa, y d el vector de soluciones. La Ecuación 4.2 solo hace en este caso referencia a uno de los gradientes de fase. Sin embargo, el razonamiento para obtener el otro gradiente es exactamente el mismo.

Se entiende por método de *Kaczmarz aleatorizado* como aquella técnica basada en el algoritmo de *Kaczmarz* pero donde las columnas de la matriz A son alternadas de forma aleatoria con una probabilidad proporcional al cuadrado de la norma euclídea. Se demostró en (Strohmer y Vershynin, 2009) que dicha aproximación mejoraba la razón de convergencia del algoritmo, y estando este basado en los trabajos de Herman y Meyer (1993); Natterer (2001);

Vempala et al. (1998). Las reglas de actualización de la solución para esta técnica se definen de la siguiente manera:

$$\mathbf{d}_{k+1} = \mathbf{d}_k + \frac{Hs - \langle \mathbf{A}_{r(i,:)}, \mathbf{d}_k \rangle}{\|\mathbf{A}_{r(i,:)}\|_2^2}, \quad (4.3)$$

donde k se corresponde con la iteración actual y $\mathbf{A}_{r(i,:)}$ es una columna aleatoria de la matriz \mathbf{A} . La proyección de s en el hiperplano $\langle \mathbf{A}_{r(i,:)}, \mathbf{d}_k \rangle$ da lugar a una cierta razón de convergencia, en comparación al *Kaczmarz* clásico.

4.2.2. WNTF y Gradientes de fase

El problema de la AT es similar a la descomposición del *light-field* sobre un conjunto de capas. En ambos casos, se trata de una tomografía de ángulo limitado, siendo por tanto un problema mal condicionado. Es por ello que, se plantea la evaluación del WNTF como alternativa a otros métodos en la bibliografía. Concretamente, y a diferencia de lo propuesto anteriormente, el problema debe dividirse en torno a la recuperación de los gradientes de fase a partir de las derivadas obtenidas por el sensor. Siguiendo la formulación del problema de la WNTF, sea \mathbf{P} un tensor N -dimensional y de tamaño $I_1 \times I_2 \times \dots \times I_N$, con N como el número de capas de la descomposición. De esta manera puede definirse el problema de la síntesis tomográfica como:

$$\mathbf{P} = \sum_{m=1}^M \prod_{n=1}^N \mathbf{d}_{n_m}, \quad (4.4)$$

siendo \mathbf{d}_{n_m} los vectores de fase en los que se descompone el tensor. En la Figura 4.2 se muestra un esquema con la descomposición del tensor \mathbf{P} . Nuevamente, esto se consigue por medio de las reglas multiplicativas propuestas por (Lee y Seung, 2001) para cada uno de ellos. Más concretamente, el sistema a minimizar sería:

$$\arg \min_{\mathbf{D}_n} \|\mathbf{P} - \mathbf{W} \otimes [\mathbf{D}_1, \mathbf{D}_2 \cdots \mathbf{D}_N]\|_2^2, \quad (4.5)$$

donde $\mathbf{D}_n = [\mathbf{d}_{n_1}, \mathbf{d}_{n_2}, \dots, \mathbf{d}_{n_M}]$, siendo m el rango de la descomposición.

A diferencia de lo visto en el Capítulo 2, el problema de la tomografía atmosférica no queda bien representado en torno a las reglas multiplicativas propuestas por Lee y Seung (1999). Se debe a que las contribuciones de fase de cada una de las capas de la atmósfera son aditivas. Si bien esto no es esencialmente un problema, es conveniente adaptar el modelo de tal manera que la descomposición se ajuste lo máximo posible a la realidad. Podemos definir las contribuciones de fase globalmente como:

$$\phi(r) = \sum_{n=1}^N \phi_n(r), \quad (4.6)$$

donde N es el número de capas de la atmósfera para el sistema de estudio y ϕ_n la fase aportada por cada una de las capas. Siguiendo la definición del WNTF, la relación debería ser

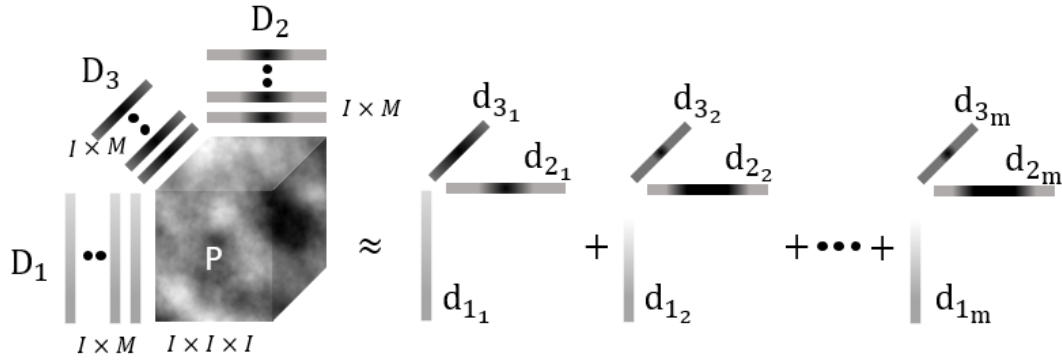


Figura 4.2 Descomposición de un tensor de tres dimensiones en sus componentes para un rango M arbitrario.

multiplicativa en su lugar. Es posible recondicionar la Ecuación 4.6 en base a la exponencial de la siguiente manera:

$$e^{\phi(r)} = e^{\sum_{n=1}^N \phi_n(r)}, \quad (4.7)$$

y por propiedades de la exponencial:

$$e^{\phi(r)} = \prod_{n=1}^N e^{\phi_n(r)} \quad (4.8)$$

De esta manera, hemos establecido una relación multiplicativa entre las contribuciones de fase de la atmósfera y la fase total percibida por el WFS. Partiendo de esta nueva definición, se introducirán en la matriz los valores referentes a la exponencial y no a la fase. Significa pues que los productos obtenidos serán también exponenciales de la contribución de fase de cada capa. Asimismo, se garantiza que los elementos del tensor son siempre positivos, manteniendo el requisito de no-negatividad que garantiza la convergencia del WNTF (Paatero y Tapper, 1994). Para obtener los valores de fase se debe invertir la exponencial.

En tanto a la representación de dicha información en el tensor P , seguiremos la misma aproximación que en el sistema multicapa. Conociendo los ángulos de proyección de cada una de las contribuciones, se dispone el resultado de la Ecuación 4.8 siguiendo los índices obtenidos atendiendo a la siguiente expresión:

$$\chi_n = x_0 + h_n \tan(\theta_x), \quad (4.9)$$

donde h_n se corresponde a la altura de cada capa, θ_x el ángulo de proyección, y x_0 la coordenada espacial en la pupila del sistema. Dependiendo de la dimensionalidad del problema, se obtendrá un conjunto de índices $\chi_n = \{\chi_1, \chi_2, \dots, \chi_N\}$.

4.2.3. Error en la descomposición multimodal

Si bien lo anteriormente expuesto abarca la adaptación del algoritmo WNTF a la tomografía atmosférica, existe una limitación intrínseca a la técnica. Hasta ahora se ha observado que al aumentar M , esto es, el rango de la descomposición y por tanto de las matrices resultantes, es posible alcanzar una representación más acertada de la matriz o tensor de estudio. Recordemos que la factorización es en esencia una aproximación y no debe entenderse como una operación exacta. De hecho, muchas de las propiedades vistas en la bibliografía surgen de fijar un cierto rango que ponga de manifiesto información embebida dentro de la propia matriz (Cichocki et al., 2009; Langville y Meyer, 2005; Lee y Seung, 2001; Wang y Zhang, 2013; Welling y Weber, 2001).

En el caso del sistema multicapa era posible asociar dicho rango a un conjunto de instantes de tiempo, tomando como referencias el sistema visual humano y el umbral de integración del ojo. Sin embargo, en la tomografía atmosférica, esta representación temporal no tiene un sentido físico. A pesar de ello, se puede considerar igualmente si ésta permite mejorar los resultados. Desafortunadamente, este hecho acarrea un error que debe ser tomado en cuenta. Si revisamos la Figura 4.2, veremos que la descomposición del tensor para $M > 1$ implica un producto de capas previo a la suma. Debido a que nuestra metodología se basa en la factorización de dos tensores de manera simultánea, que se corresponden a los gradientes horizontal y vertical de la fase, es necesario poder acceder a cada uno de estos gradientes por capa para obtener la fase a cada altura elegida. Como se observa, si se hiciera dicha aproximación, se estaría cometiendo un error. Formalmente, se diría que el gradiente de fase vendría dado por:

$$\mathbf{D}_n^X = \sum_{m=1}^M \mathbf{D}_{n,m}^X \quad (4.10)$$

Esto implica el siguiente valor de error en el RMS:

$$e_{RMS} = \text{RMS} \left(\prod_{n=1}^N \sum_{m=1}^M \mathbf{d}_{n,m} - \sum_{m=1}^M \prod_{n=1}^N \mathbf{d}_{n,m} \right), \quad (4.11)$$

siendo RMS el valor eficaz. Nótese que, este error crece a razón de M y N . Sin embargo, veremos más adelante que dicha aproximación resulta favorable a pesar del error cometido.

4.3. Evaluación de la técnica

Tomando como punto de partida las definiciones de los dos métodos propuestos, se ha realizado un conjunto de simulaciones que permita validar la viabilidad de la aplicación del WNTF para la obtención de la tomografía atmosférica.

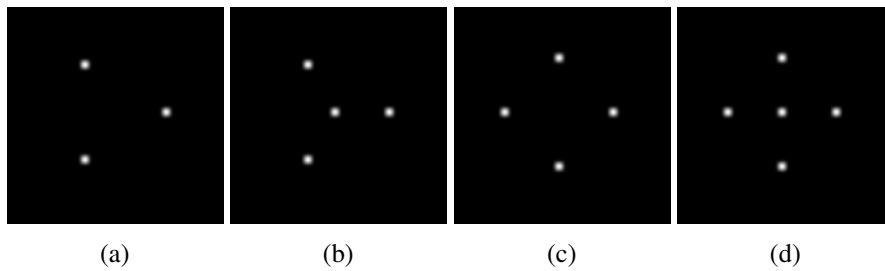


Figura 4.3 Distribuciones propuestas de constelaciones de estrellas láser guía (LGS). De izquierda a derecha: (a) distribución en triángulo (TS), (b) distribución en triángulo *on-axis* (TSO), (c) distribución en cuadrado (SS) y (d) distribución en cuadrado *on-axis* (SSO). Cada cuadrado representa una cobertura del cielo de 97.8arcsec.

4.3.1. Modelo atmosférico

A pesar de la analogía seguida con el sistema multicapa a lo largo de este Capítulo, la atmósfera es, a diferencia del mismo, un medio continuo. La propagación de la luz a través de dicho medio resulta en una desviación de los rayos debido a los cambios de índice de refracción, en definitiva, una modificación de la fase del campo electromagnético. No obstante, a efectos de evaluar la técnica tomográfica, es habitual reducir dicho volumen continuo en un conjunto de capas que representen de forma cercana el comportamiento del sistema. No es objeto de este estudio la estimación del número de capas necesarias para modelar el sistema atmosférico, si bien es cierto que posteriormente se evaluará de manera cuantitativa el efecto esta discretización a la hora de obtener capas no definidas.

Las presentes simulaciones se basan en un modelo cuantitativo de la atmósfera obtenido a partir de las medidas realizadas por el telescopio *Gemini South* (La Serena, Chile) y que puede ser consultado en Cortes et al. (2012). El modelo se compone de un conjunto de siete capas de turbulencias con un parámetro global de *Fried* de 15.4cm, y una resolución de 256×256 píxeles por capa. En la Tabla 4.1 se recogen los valores C_n^2 del perfil de turbulencia del modelo. Estos se corresponden con la ponderación de las capas dentro del modelo y están normalizados de tal manera que su sumatorio equivalga a la unidad. Dicho modelo se ha empleado previamente (Trujillo-Sevilla, 2017), dando lugar a unos resultados que se ajustan a la realidad. Cada capa se ha simulado asumiendo un espectro de Kolmogorov (Kolmogorov, 1941) y de manera independiente a partir de su C_n^2 , asumiendo la presencia de subarmónicos sobre la pantalla de fase (Lane et al., 1992).

4.3.2. Sistema de estudio

Con el fin de realizar una comparación estadísticamente válida entre los modelos propuestos, se han generado un total de 1000 atmósferas independientes, siendo estas posteriormente empleadas para el cómputo de la tomografía. Si bien es cierto que el número de capas empleadas en la reconstrucción es todavía un campo de estudio ampliamente revisado (Auzinger, 2015; Auzinger et al., 2015; Le Louarn et al., 2000; Saxenhuber D, Auzinger G, Louarn ML, 2017),

Cuadro 4.1 Parametrización del modelo atmosférico

Capa	Altura (m)	Perfil de turbulencia (C_n^2)
1	0	0,615
2	1 800	0,126
3	3 750	0,094
4	4 500	0,110
5	6 250	0,031
6	7 900	0,008
7	9 000	0,016

se ha estudiado para dos y tres alturas respectivamente. En cada uno de estos casos, las alturas elegidas han sido de 0 y 4500m para dos capas y 0, 4500m y 9000m para tres capas.

A pesar de que en muchos casos se hace uso de los propios objetos celestes para evaluar la fase aportada por la atmósfera, no es raro emplear estrellas artificiales como fuentes puntuales (Duarte, 2001; Primmerman et al., 1991). Esto se consigue proyectando un haz de luz con un láser de alta potencia que atraviesa la atmósfera e ioniza los átomos de la capa de sodio, por tanto mostrando una proyección similar a la de un cuerpo celeste. De manera general, es posible agrupar estrellas láser de guía o *light guiding stars* (LGS) en constelaciones con el fin de obtener un mejor condicionamiento del sistema. En las simulaciones propuestas se propone el uso de cuatro constelaciones de LGS, observándose éstas en la Figura 4.3. Cada estrella se considera puntual y las respectivas constelaciones subtienden un ángulo de 96.68arcsec, estando posicionadas de forma teórica a una altura de 90km.

Asimismo, se ha simulado un telescopio con una pupila de 8m. La propagación del campo inicial procedente de cada estrella y sus subsecuentes propagaciones a través de cada capa de la atmósfera y del telescopio forman la imagen de cada una de las fuentes emisoras. El cono correspondiente a la propagación de cada una de estas fuentes se asume como detectado de forma completa por un sensor de fase ideal de resolución 32x32. Para completar el lazo del sistema, se ha introducido un espejo deformable ideal encargado de realizar las correcciones estimadas sobre las aberraciones detectadas. Finalmente, la fase teórica percibida por el sensor se ha obtenido sustrayendo la fase de la turbulencia atmosférica original.

4.3.3. Razón de Strehl

La evaluación de la técnica en tomografía atmosférica se suele realizar por medio de diferentes métricas, siendo la razón de Strehl una de las más extendidas. Esta razón se define como la relación entre la intensidad de imagen obtenida de una fuente puntual considerando las aberraciones del sistema y la intensidad máxima que se puede obtener partiendo de un sistema

no afectado por aberraciones y limitado solo por difracción (Strehl, 1895). Siguiendo la aproximación de Maréchal (Maréchal, 1948), para un sistema con una apertura de diámetro D :

$$S = e^{-\sigma^2}, \quad (4.12)$$

con σ^2 como:

$$\sigma^2 = \mu \left(\frac{D}{Lr_0} \right), \quad (4.13)$$

con r_0 como el parámetro de *Fried*, y siendo μ un valor comprendido en el rango 0,2 y 1, condicionado al número efectivo de actuadores del espejo deformable implementado sobre el sistema de AO (Roddier, 1999).

A pesar de que numéricamente sería posible aceptar que la razón de Strehl llega a tomar la unidad como valor máximo, es necesario tener en cuenta los límites ópticos impuestos por el sistema. Si se sustituye L en la Ecuación 4.13 por el número de subaperturas en el sensor de fase, y D como el diámetro de la pupila del telescopio, el valor máximo obtenido para la razón de Strehl es de 0,6386.

Es posible obtener una aproximación de manera computacional de esta razón límite. Para demostrar la validez del cálculo previo, se realizaron 1000 simulaciones del sistema, corrigiendo precisamente con los valores de fase generados. Esto dio lugar a un valor medio de 0,6361 a lo largo del campo de visión, con una desviación típica de 0,0414. Por lo tanto, se valida que la estimación realizada de forma teórica coincide con el modelo implementado.

Comúnmente, y siguiendo la idea previa, se suele presentar no únicamente una razón, sino un mapa de estos valores a lo largo del campo de estudio. Esto se conoce como mapa de Strehl, y será empleado posteriormente para evaluar el desempeño de la técnica.

4.4. Resultados

La implementación de los algoritmos *Kaczmarz* y *WNTF* así como todas las simulaciones se han realizado sobre MatlabR2016a (MATLAB, 2016). El algoritmo de *Kaczmarz* se ha ajustado a 20 iteraciones, ya que se ha observado experimentalmente que para el presente caso de estudio existe una convergencia en torno a dicho valor. De una manera similar a como se ha trabajado en los Capítulos anteriores, se ha partido de la descomposición del tensor P en el número de capas correspondiente. Además, siguiendo criterios similares a los ya indicados en la bibliografía, el límite de iteraciones para este último método se ha situado en 100, mientras que el número de modos de la descomposición se ha fijado en 32. En ninguno de los dos casos se ha detenido la síntesis de las capas una vez alcanzado un cierto valor de tolerancia.

En el Capítulo 2 se determinó que la elección de una solución inicial apropiada podía recaer en una mejora sustancial de la calidad del resultado obtenido (Carmona-Ballester et al., 2017; Langville et al., 2006). No obstante, en la implementación empleada en la síntesis tomográfica de la atmósfera se ha preferido optar por una aproximación más simple,

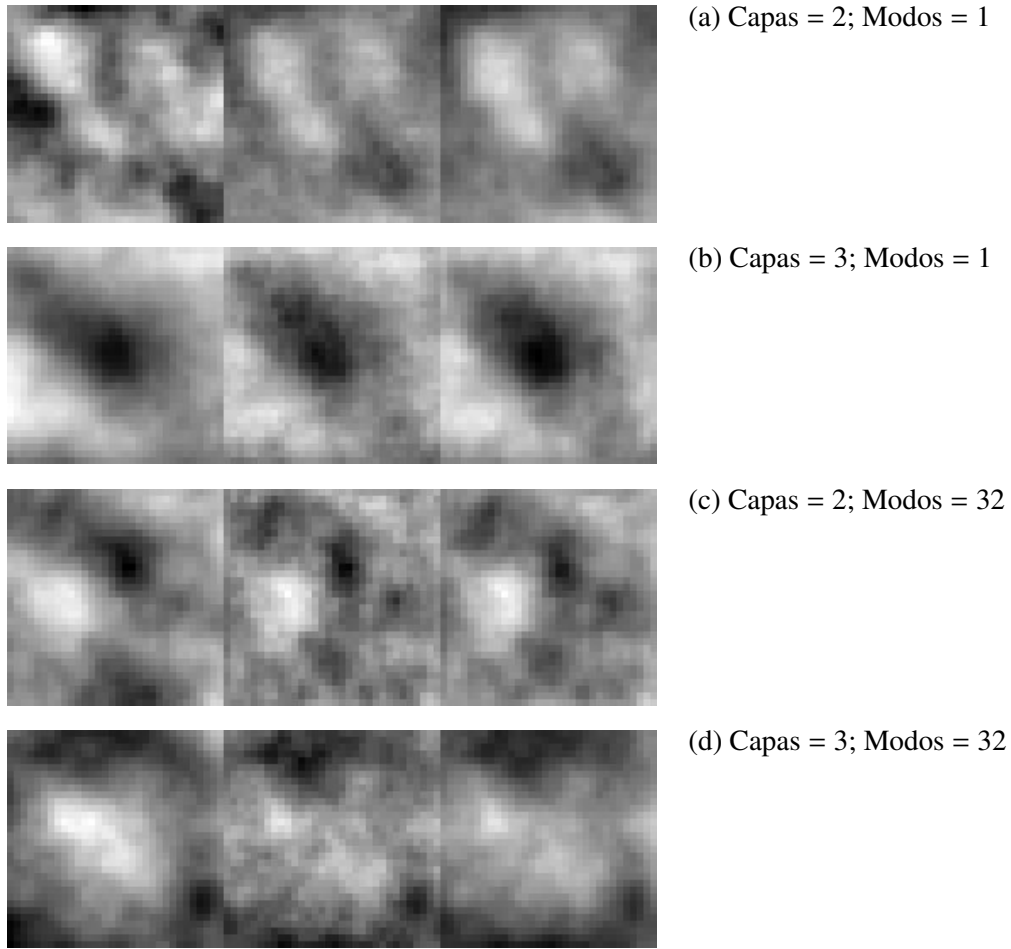


Figura 4.4 Ejemplo comparativo entre las fases recuperadas para cada uno de los métodos propuestos. De izquierda a derecha, fase de referencia, recuperación con *Kaczmarz* y WNTF respectivamente.

evitando una comparación injusta con el método base. Por ello, cada una de las capas \mathbf{D}_n se ha inicializado con ruido *gaussiano* de media $\mu = 0$ y desviación típica $\sigma = 1$. Posteriormente, los valores resultantes se han limitado al rango de $[0, 1]$.

Para mantener una correspondencia entre los valores de fase obtenidos, se ha aplicado una corrección de estas a partir de un coeficiente resultado de la minimización del RMS con respecto a la fase teórica. Esta aproximación se realiza para cada una de las técnicas, y el objetivo es permitir evaluar el desempeño de ambas de manera objetiva. Esto se puede expresar de la siguiente manera:

$$\arg \min \|\text{RMS}[\phi_n(r) - k\phi'_n(r) + c]\|_2^2, \quad (4.14)$$

donde $\phi'_n(r)$ se corresponde con la fase aproximada obtenida por cada uno de los dos métodos.

Adicionalmente, partiendo de las simulaciones, se ha realizado un contraste estadístico *t* de *Student* de los promedios radiales de las razones de Strehl obtenidas. Este se ha realizado

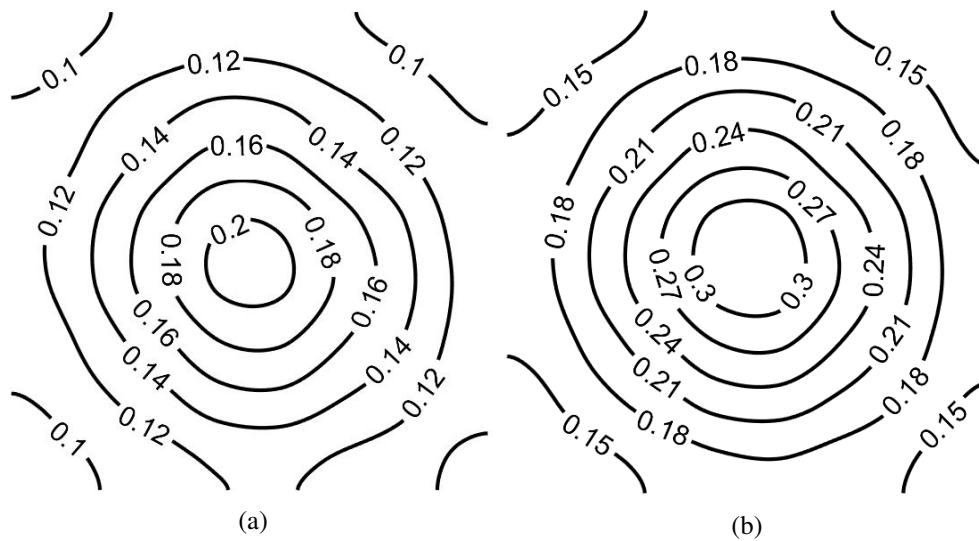


Figura 4.5 Promedio radial de la razón de Strehl a lo largo del campo de visión para un diámetro de pupila de 8m tras 1000 simulaciones para los métodos de (4.5a) *Kaczmarz* iterativo y (4.5b) WNTF. Las condiciones de simulación se corresponden con tres capas atmosféricas, 96.68arcsec, la distribución de LGS SSO.

en intervalos de 0.77arcsec (desde 0 a 48.34arcsec), y se ha realizado para cada combinación posible de situaciones, es decir, número de capas, métodos y constelaciones. Se considera una muestra estadísticamente significativa, y por tanto una diferencia válida, si el p-valor es menor o igual a 0,05.

De manera ejemplar y con fines ilustrativos, se muestra en la Figura 4.4 una comparativa para el caso de tres capas extraído de las 1000 simulaciones realizadas para dos y tres capas respectivamente haciendo uso de la topología SSO. De izquierda a derecha se encuentran la fase original, la recuperación por medio del algoritmo de *Kaczmarz* iterativo y por WNTF. Se ha incluido además el resultado de la descomposición con el WNTF para 1 y 32 modos respectivamente. Una primera observación permite concluir que los resultados obtenidos para el método propuesto mejoran en cierto grado los obtenidos por el algoritmo base. En línea con el análisis realizado en la Sección 4.2, se ha estudiado el impacto de la descomposición multimodal en los valores de Strehl obtenidos con el fin de entender el impacto del error generado al realizar la aproximación correspondiente. Se observa que, para una síntesis de rango 1, el valor de pico decae en torno a un cuarto con respecto al caso de 32 modos. Por ello, en adelante las simulaciones se han realizado a partir de este último valor.

En la Figura 4.5 se muestran lado a lado los mapas de Strehl obtenidos tras promediar 1000 recuperaciones para la topología SSO. De manera global, el método WNTF da lugar a un incremento de aproximadamente el 50 % en la razón de Strehl. Este incremento se mantiene constante a lo largo del campo.

Se ha realizado posteriormente el promedio radial de los mapas de Strehl obtenidos. La Figura 4.6 contiene dichos resultados. Nótese que el rango de representación del campo de visión es de la mitad ya que la operación se aplica desde el eje central del campo. Cada gráfica

contiene la curva de Strehl para cada método, así como un valor de referencia. Esta referencia es el valor promedio obtenido si no se aplica ninguna corrección al sistema. Además, se ha representado el valor de la varianza en varias posiciones de la curva por medio de barras horizontales. Para todos los casos de estudio se ha obtenido una razón de Strehl superior haciendo uso del caso de WNTF frente al *Kaczmarz* siendo estos resultados estadísticamente significativos. Notablemente, se observa que la metodología propuesta tiene una varianza muy superior en cualquiera de las topologías. Estos valores además se reducen conforme nos alejamos del eje del sistema. Como cabía esperar, los valores de reconstrucción para cualquiera de los dos métodos son superiores cuando se evalúa un número superior de capas. Ya se indicó previamente que, para un mayor número de capas, el sistema estaría mejor condicionado, y por ello, se derivan estos resultados (Le Louarn et al., 2000; Saxenhuber D, Auzinger G, Louarn ML, 2017; Wetzstein et al., 2012a).

Es posible que al recuperar la fase de capas definidas, la corrección se vea numéricamente afectada de forma positiva. Por ello, se ha repetido la recuperación para alturas cuyas capas no estén definidas. Se han elegido para esta simulación los valores de altura 1000m, 3000m y 8000m. La evaluación de la técnica se ha realizado de manera idéntica a la metodología previa, pero por simplicidad se ha limitado a la configuración SSO. La gráfica resultante está contenida en la Figura 4.7. Como se puede observar, de manera general ha habido una reducción en la razón de Strehl para ambas curvas. Sin embargo, la técnica propuesta continúa produciendo de media una mejora considerable con respecto al método base.

Al final de esta memoria se ha añadido el Anexo B en el que se muestra la Figura B.1. Esta contiene la relación estadística entre cada uno de los métodos aplicados a la reconstrucción con cada una de las configuraciones de LGSs. A pesar de que los resultados presentados en la Figura 4.6 demuestran una mejoría considerable en todos los casos, estadísticamente el método de *Kaczmarz* es superior al WNTF en algunos de estos ($p\text{-valor} \leq 0,05$). Las celdas resaltadas en negro indican un mejor resultado para el WNTF, mientras que las blancas determinan lo opuesto. Se ha indicado además en algunos de los casos los rangos en los que la comparación estadística es válida. Las celdas resaltadas en gris no son concluyentes en ninguno de los casos.

4.5. Sumario y Conclusiones

La similitud entre el sistema multicapa y la tomografía atmosférica han dado lugar a la implementación de una técnica cuya aplicación no es común en el campo de la astronomía. Como resultado de esto, se ha podido concluir, de manera provisional y teórica, que es posible mejorar considerablemente los valores obtenidos por métodos más tradicionales. Más allá de esto, la novedad de lo presentado en este Capítulo radica en la evaluación de la técnica considerando la información obtenida por un sensor de fase como el *SH* a modo de *light-field*. Precisamente por ello, ha sido posible concluir que la aplicación de la técnica, dentro de unos límites definidos, podría ser viable en un caso práctico.

Intuitivamente, y siguiendo lo estudiado en Capítulos previos, se teoriza acerca de que la mejora observada es producto de la generación implícita de información no conocida a raíz de la descomposición. Más concretamente, se considera que la síntesis de las capas atmosféricas a partir de la retopologización de la información extraída del sensor, entra dentro de la categoría de *Compressed Sensing*. En trabajos previos ya se ha estudiado como la descomposición de una matriz se hace basada en patrones, y en cierta manera, es posible extraer información subyacente al elemento de estudio sin necesidad de ser explícitamente representada (Chien y Bao, 2018; Jain y Oh, 2014; Lee et al., 2007; Nickel et al., 2016). Precisamente, Trujillo-Sevilla et al. (2016) ya presentaba previamente, desde un punto de vista práctico, la generación de información a partir de un *light-field* incompleto partiendo de aproximaciones similares.

Anteriormente se ha demostrado que la aplicación de las reglas multiplicativas no se ajusta al problema en cuestión, derivándose un término de error. No obstante, este tiene un efecto reducido sobre el resultado final. Se teoriza que esto es debido mayoritariamente a la prevalencia de las propiedades que subyacen a los casos de estudio, esto es, componentes vertical y horizontal de los gradientes, y la capacidad del método de tomar ventaja de éstas.

Para trabajo futuro queda la necesidad del estudio de la viabilidad en tiempo real del algoritmo. Ya se veía en el Capítulo 2 que es posible plantear un reacondicionamiento del método WNMF tanto para reducir sus requisitos computacionales, como para adaptarlo a GPU o FPGA. Comparativamente, el método de *Kaczmarz* iterativo tiene una complejidad $O(n)$ o polinómica, lo cual, para los sistemas de AO es más que suficiente. Sin embargo, el algoritmo WNTF para el caso de las reglas multiplicativas, corresponde a un tiempo polinomial no determinístico, esto es, se trata de un problema NP-hard (Vavasis, 2010). Por ello, el crecimiento del sistema afecta de manera excesivamente negativa a los tiempos y requisitos computacionales. Este es uno de los motivos por los que se ha limitado la resolución del sistema a dos y tres capas respectivamente.

En definitiva, los avances presentados en este Capítulo deben entenderse como una prueba de concepto de la aplicación de la técnica a un sistema de AO. Se han realizado múltiples aproximaciones que, en el caso de un telescopio, no serían completamente válidas. Este es el caso por ejemplo de las constelaciones de LGS, asumidas como puntuales y la corrección aplicada por el sistema (espejo deformable ideal). Existen en la bibliografía otras simulaciones más ajustadas que podrían arrojar algo más de luz al método en cuestión.

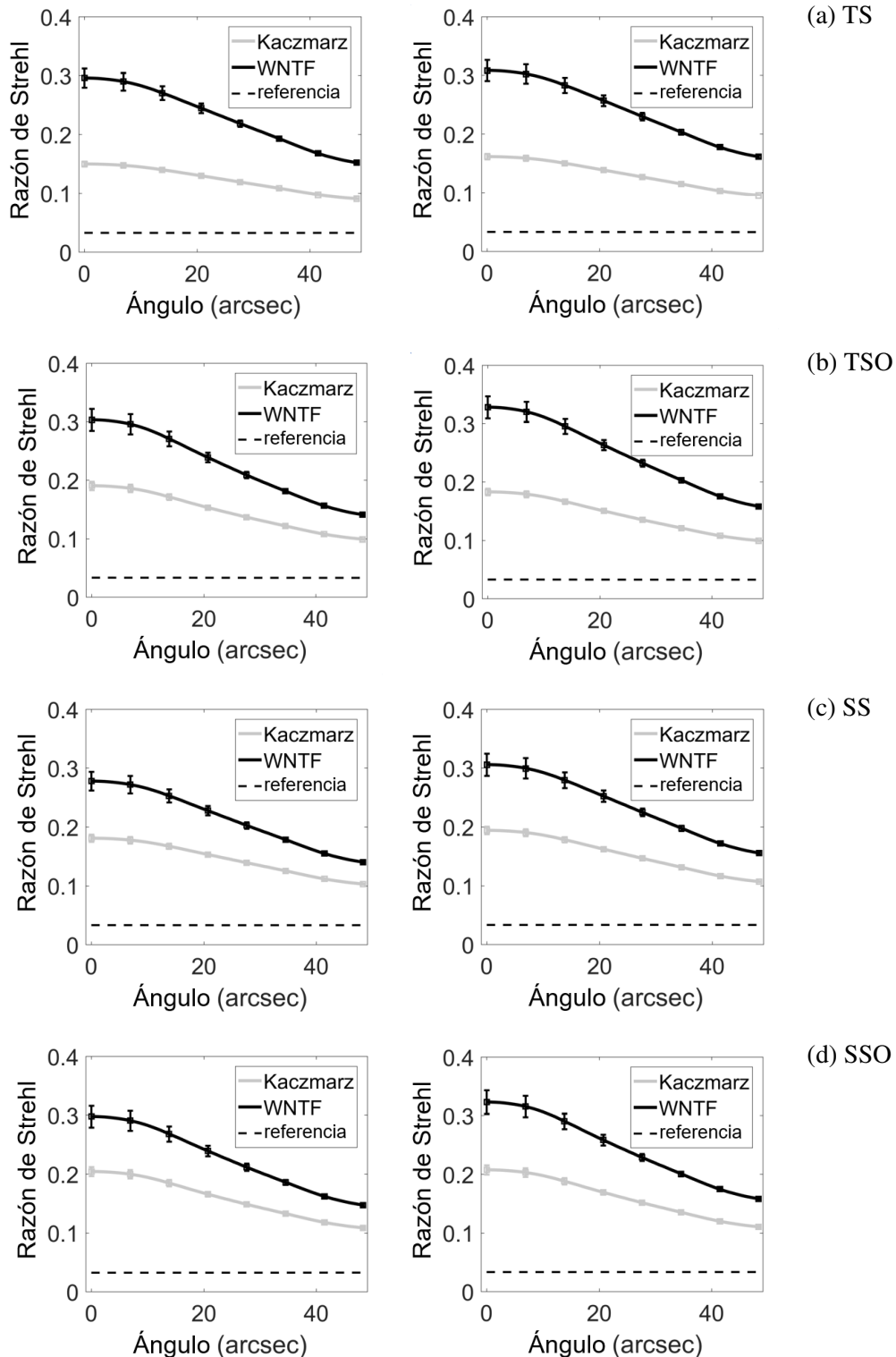


Figura 4.6 Promedio radial de la razón de Strehl a lo largo del campo de visión obtenido a partir de 1000 simulaciones de atmósfera diferentes. Cada fila tiene la siguiente equivalencia con las constelaciones definidas: (a) TS, (b) TSO, (c) SS y (d) SSO. La columna izquierda y derecha se corresponden a la descomposición en 2 y 3 capas respectivamente. Se ha representado además la varianza de algunos puntos como barras de error verticales. La línea discontinua indicada como *referencia* representa la razón de Strehl del telescopio sin ninguna corrección. El valor máximo alcanzable es de 0.6386 para esta simulación (véase Ecuación 4.13).

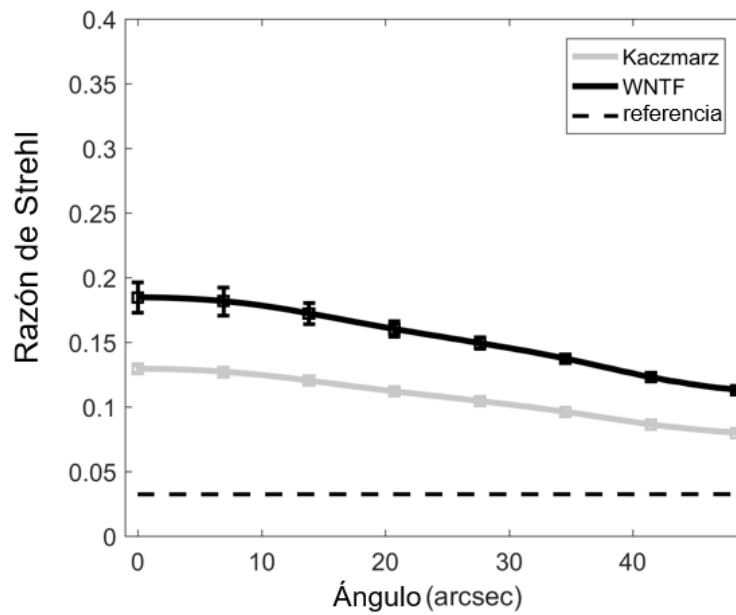


Figura 4.7 Relación de Strehl radial promedio a lo largo del campo de visión tras realizar la descomposición en tres capas atmosféricas a alturas no correspondientes con las definidas por defecto (1000, 3000 and 8000m). La simulación se corresponde con la constelación de LGS cuadrada en eje.

Capítulo 5

Mejora de la resolución lateral de *display* bidimensional por introducción de desenfoque

Inevitables son, por consecuencia, las coincidencias y las contiendas de prioridad. Y es que, lanzada al público una idea, entra a formar parte de ese ambiente intelectual donde todos nutrimos nuestro espíritu; (...) En todo caso, si alguien se nos adelanta, haremos mal en desalentarnos. Continuaremos impertérritos la labor, que al final llegará nuestro turno.

– Santiago Ramón y Cajal

En los Capítulos anteriores se ha estudiado el uso de un sistema multicapa para la generación de contenido tridimensional como parte de un primer bloque en el marco de los soportes de visualización. Este Capítulo forma parte del segundo bloque de contenido de esta memoria, y los conceptos que se estudian a en éste se plantean desde una perspectiva bi-dimensional, es decir, que su aplicabilidad está orientada a un sistema tradicional de visualización. Sin embargo, y como se verá al final del Capítulo, se teoriza sobre la posibilidad de incorporar dichas metodologías a los contenidos y algoritmos anteriormente expuestos en el marco de imagen tridimensional.

Más concretamente, el método que se presenta busca superar la limitación en ancho de banda que un sistema de visualización tiene por el hecho de estar constituido por píxeles discretos. Esto se consigue introduciendo de forma controlada un desenfoque que genere una cierta interacción entre éstos, y posteriormente pre-compensar las imágenes representadas teniendo en cuenta esta información.

5.1. Antecedentes

La idea de pre-compensado o deconvolución de la imagen ya se ha analizado previamente en la literatura con diferentes fines. En esencia, lo que se busca es corregir con base en una función que el sistema de observación introduce a la imagen de estudio. Una de las aplicaciones más comunes es la de corrección de defectos visuales (Huang et al., 2011; Miguel y Barreto, 2003; Miguel et al., 2005). No obstante, esta no es la única aplicación que esta metodología encuentra. De manera paralela, la introducción del pre-compensado busca en otros casos la mejora sobre la resolución de un dispositivo. Dentro de estas aproximaciones se pueden encontrar metodologías basadas en el tiempo de integración del ojo, como propone Didyk et al. (2010), tomando en consideración el movimiento relativo dentro de una secuencia de imágenes. El resultado de esta operación permite la integración natural de las imágenes y por tanto mejora la percepción del observador. En muchos otros casos se proponen sistemas dinámicos, considerando las perturbaciones introducidas sobre ellos (Berthouzoz y Fattal, 2012). Otros trabajos se aproximan a esta cuestión desde el punto de vista de la percepción humana. Por ejemplo, Shen et al. (2015) proponen pre-compensar a razón de conservar la consistencia en términos perceptuales al redimensionar imágenes, más concretamente, tomando como base el círculo de confusión en relación al tamaño y la distancia al sistema.

A pesar de que estas técnicas se generalizan para sistemas más tradicionales como monitores o televisiones, existe también una evaluación similar para sistemas de proyección. Damera-Venkata y Chang (2007) propone un método basado en el solapamiento de múltiples proyectores, de tal manera que al evaluar las áreas intersectadas por las imágenes resultantes, es posible superar el límite de resolución impuesto de manera individual por los mismos. Posteriormente, Damera-Venkata y Chang (2009) continúan con una evaluación más detallada de lo que pasa a denominarse el efecto de *alias-cancelling* o supresión de alias, que no es otra cosa que lo anteriormente citado: la obtención de frecuencias espaciales superiores al límite impuesto por el ancho de banda del sistema. Otros autores avanzan de manera similar, apoyándose en las cuestiones desarrolladas por los trabajos previos para extender la proyección a superficies no planas (Aliaga et al., 2012).

Paralelamente, autores como Zhang y Nayar (2006) comienzan a presentar metodologías de pre-compensado adaptativas. En este caso particular, se propone deconvolucionar partiendo de un desenfoque ajustado a la distancia de proyección. Esto permite representar con una resolución muy similar los contenidos, garantizando uniformidad en el campo proyectado. Un efecto derivado de esta técnica es la supresión del llamado efecto *screen-door* o rejilla, teniendo este origen en la visualización de los intersticios de los píxeles. En el trabajo de Sajadi et al. (2012) se propone algo similar, pero añadiendo un conjunto de moduladores espaciales de luz o *spatial light modulators* (SLM)¹, con una cierta similitud al concepto propuesto por Lanman et al. (2014). Por medio de esta topología se demuestra que es posible incrementar la

¹Un SLM consiste en una placa que permite modular la forma de un haz de luz. Además, ocasionalmente se puede hacer uso de éstos para controlar el retardo o fase de una onda con diferentes fines prácticos. Suele tratarse de láminas LCD, y son de uso común en proyectores (Kreis, 2012).

resolución del sistema si se descompone la imagen en un componente de borde y una imagen de baja resolución, siendo ambas representadas sobre el proyector. A pesar de ser una solución sobresaliente, la introducción de SLM complica el sistema, además de provocar una reducción en la cantidad efectiva de luz del 50 %.

Una evolución del trabajo anterior se propone en Sajadi et al. (2013), donde ahora los SLM no forman parte del diseño. En su lugar, superponen al sistema un tipo de objetivo consistente en múltiples lentes, de tal manera que la proyección realizada da lugar a múltiples copias o imágenes fantasma. Con un planteamiento similar a Damera-Venkata y Chang (2007), se realiza un pre-compensado de la imagen que permite mejorar en cierto grado la resolución efectiva del sistema.

Son pocos los trabajos que analizan o plantean el uso de un desenfoque controlado para la mejora de la resolución de un sistema de visualización *a priori* conocido. Específicamente, los trabajos de Zhang y Nayar (2006) y Sajadi et al. (2012) que sí llegan a plantear esto, tratan el tema de forma bastante simplificada al no terminar de dar una respuesta a preguntas esenciales como el rango de operación óptimo, dependencia de la calidad con el desenfoque y efectos intrínsecos de la técnica. En este último bloque de la memoria se busca dar respuestas a algunas de estas preguntas, tomando como referencia una aproximación basada en la idea del solapamiento entre píxeles, y mostrando cómo las nuevas relaciones establecen un margen de operación que permite de forma teórica sobrepasar el límite en frecuencia impuesto por los sistemas de visualización.

5.2. Estimación de la función de solapamiento

En esta Sección se estudiará la introducción de una función de desenfoque como método para generar un conjunto de interdependencias en forma de solapamientos en una rejilla de píxeles *a priori* definida. Para ello, se diferenciará el análisis para dos tipos diferentes de topología. La primera, a la que denominaremos *Solapamiento Puntual* (SP) partirá de una rejilla en la que las fuentes de luz –o píxeles– serán infinitesimales. Alternativamente, se estudiará el mismo efecto sobre un sistema con rejilla cuadrada, similar a las estructuras de píxeles más comunes, y lo denominaremos *Solapamiento Continuo* (SC).

5.2.1. Solapamiento puntual como caso de estudio

Sea $I(x, y)$ la función intensidad producida por un sistema de visualización. Definamos a su vez una rejilla bi-dimensional de fuentes emisoras $\delta(x, y)$ separadas una cantidad fija entre ellas (x_0, y_0) . Estas distancias pueden entenderse como el tamaño de cada píxel cuyo conjunto forma el sistema de visualización propuesto. De acuerdo con el teorema de *Nyquist-Shanon* o de muestreo, la resolución de nuestro sistema está limitada tanto horizontalmente como verticalmente por el número de fuentes (K, M) .

Partiendo de esta definición, la resolución del sistema es fija, siendo cada punto una unidad de información. Es necesario aplicar un operador sobre dicha función que dé lugar a nuevos

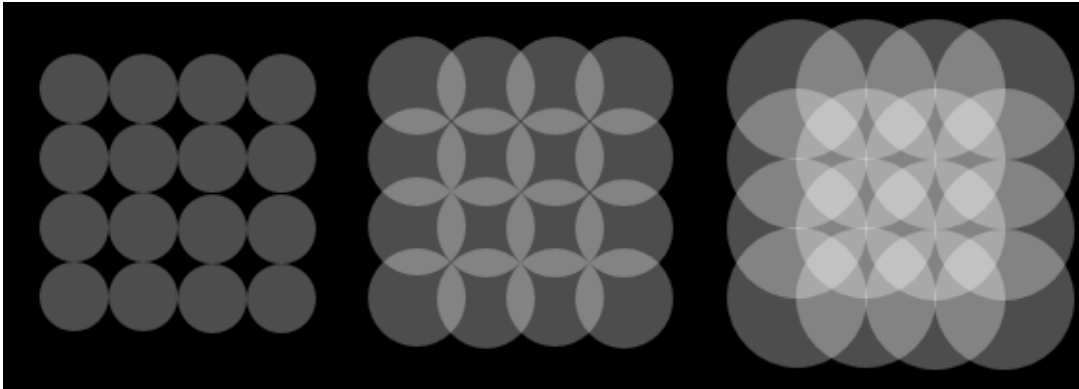


Figura 5.1 Representación teórica de la función de intensidad I_c tras convolucionar con una PSF de forma circular con diferentes radios.

puntos cuyos valores sean accesibles a la imagen que se busca representar. Asumamos que el sistema de visualización propuesto está afectado por una función de dispersión del punto, comúnmente referida en la literatura como *point spread function* (PSF). Este término se puede introducir en el sistema anteriormente indicado por medio de una convolución:

$$I_c(x, y) = \sum_{\substack{k=0 \\ m=0}}^{K, M} [\delta(x - kx_0, y - my_0) \otimes PSF(x, y)], \quad (5.1)$$

donde \otimes es el operador de convolución e $I_c(x, y)$ es la función intensidad convolucionada. Siguiendo el razonamiento anterior, la PSF puede ser ajustada de tal forma que se generen unas interacciones entre los píxeles de manera controlada, produciendo con ello produciendo áreas de interacción entre los mismos y, por tanto, áreas de intensidad gobernadas por el valor de un conjunto finito de éstos. De manera ilustrativa, en la Figura 5.1 se muestran varios ejemplos tras convolucionar con una PSF circular de diferentes radios. Como se observa, el resultado consiste en una réplica de dicha PSF en las posiciones de cada píxel. Las interacciones entre cada círculo están resaltadas con un tono más claro. Cuanto mayor es el radio, mayor es el número de solapamientos presentes en la imagen, y por tanto, mayor dependencia generada en torno a ese área.

Cabe mencionar que la PSF no tiene por qué ser necesariamente un círculo, sino que podría ser cualquier otra función que diera lugar a un efecto similar de solapamiento entre los píxeles. Sin embargo, se ha elegido así porque equivale a introducir un desenfoque en el sistema, mitigando las complejidades derivadas en una implementación física del modelo. Siguiendo la metodología anteriormente expuesta, se define la PSF circular como sigue:

$$PSF(x, y) = \frac{1}{\pi r^2} circ(x, y), \quad (5.2)$$

con radio r . La función $circ(x, y)$ representa un pistón, siendo 1 en el interior de la circunferencia, y 0 en su exterior. En la expresión anterior, se normaliza el valor de la función por un factor $\frac{1}{\pi r^2}$, a fin de garantizar que la energía total del punto se conserve tras la convolución.

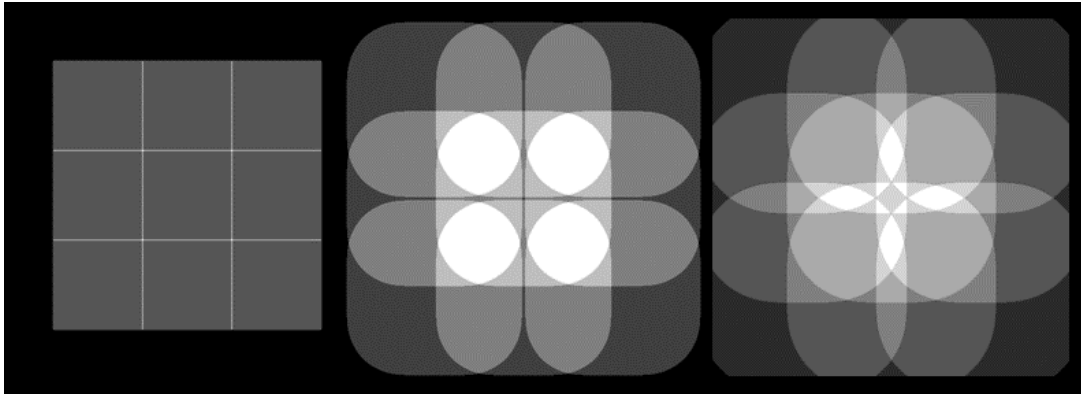


Figura 5.2 Representación teórica de la función de intensidad I_c tras convolucionar con una PSF en forma circular con diferentes radios y asumiendo una rejilla de píxeles cuadrados.

Finalmente, se puede expresar la Ecuación 5.1 de la siguiente manera:

$$I_c(x, y) = \sum_{\substack{k=0 \\ m=0}}^{K,M} \frac{1}{\pi r^2} \text{circ}(x - kx_0, y - my_0) \quad (5.3)$$

Esta expresión se corresponde con la Figura 5.1, variando el valor de r con el objetivo de obtener una serie de solapamientos.

5.2.2. Solapamiento continuo y generalización

Previamente se ha definido un metodología que permita generar interacción entre los píxeles que conforman un sistema de visualización con el fin último de usar esta circunstancia a favor de la generación de nuevas áreas de representación controlables. Sin embargo, lo presentado se limita al caso de un sistema en el que los píxeles están definidos como elementos infinitesimales en un área de visualización específica. En la mayoría de los casos prácticos, esto no es así. Normalmente, un sistema de visualización está compuesto por unidades finitas. Específicamente, la función δ puede ser sustituida por aquella que defina la estructura real de los píxeles. A nivel práctico, y obviando por el momento la cromaticidad, cada unidad se puede definir como un cuadrado por medio de la función $\text{rect}(\frac{x}{W}, \frac{y}{W})$, donde W define el ancho y el alto. Si se sustituye en la Ecuación 5.1, se obtiene la siguiente expresión:

$$I_c(x, y) = \sum_{\substack{k=0 \\ m=0}}^{K,M} \left[\text{rect}\left(\frac{x - kx_0}{W}, \frac{y - my_0}{W}\right) \otimes \text{PSF}(x, y) \right] \quad (5.4)$$

En la Figura 5.2 se muestra de forma teórica I_c tras la convolución con la PSF. Nótese que, de manera similar al caso infinitesimal, se producen una serie de solapamientos. Sin embargo, y como se verá más adelante, los solapamientos producidos en este caso no serán visibles en la imagen final, como sí ocurrirá con la topología anterior.

5.2.3. Puntos de operación

Se han definido las funciones de intensidad para dos disposiciones de píxeles diferentes; en un caso, se ha asumido una estructura de fuentes infinitesimales de luz (SP), y en otro, una rejilla formada por cuadrados contiguos (SC). En ambos, una función PSF circular –equivalente a un desenfoque variable– es introducida con el fin de generar un conjunto de solapamientos en ciertas zonas de la función de intensidad del sistema de visualización. Sin embargo, resta dar respuesta a cómo debe tratarse el desenfoque en el sistema, y qué radio debe asignarse para maximizar la mejora obtenida.

Partamos nuevamente del SP. Sea r'/r la relación entre el radio r de una circunferencia unitaria, y asumiendo que se encuentra justo en el límite previo a producir un solapamiento con su vecindad (véase Figura 5.1) y un r' que defina un caso arbitrario de solapamiento. Un incremento de esta relación supone un aumento de los solapamientos entre vecindades, lo cual lleva de forma unívoca a un mayor número de interdependencias entre los valores de intensidad controlables. Nótese que estos solapamientos se suceden de forma discreta, esto es, hay ciertos valores para r'/r que definen intervalos en los que el número de interacciones es fijo.

Si asumimos que la PSF es circular, es posible obtener los rangos de solapamiento calculando la distancia mínima desde un círculo de referencia a las vecindades. Esto puede expresarse de la siguiente manera:

$$d_{k,m} = \sqrt{(kx_0)^2 + (my_0)^2}, \quad \forall(k, m) \in \mathbb{N} \quad (5.5)$$

donde los valores que se obtienen para la relación r'/r vienen dados en función de la mitad del diámetro en el rango obtenido, esto es:

$$r'/r = \frac{d_{k,m}}{2r} \quad (5.6)$$

En la Figura 5.3 se representa de manera general un caso de 3x3 fuentes ya convolucionadas con una PSF cuyo $r'/r = 1$. Como se observa, es posible obtener las distancias a los diferentes centros por medio del teorema de Pitágoras. A la derecha, se muestra una versión esquematizada para los casos de $r'/r = 1$ y $r'/r = \sqrt{2}$. Más concretamente, los círculos más pequeños se corresponden con el primer orden de desenfoque, esto es, $r'/r = 1$, en el cual no se establece ningún solapamiento. El otro par de círculos, algo mayores, se corresponde con el $r'/r = \sqrt{2}$. En este último caso, debido al incremento en el radio de la función de desenfoque, se puede observar un cierto solapamiento. A pesar de que aquí solo mostramos dos desenfoques diferentes, es posible aplicar cualquier valor arbitrario de r'/r obtenido en la Ecuación 5.5. Ordenando estos valores se pueden obtener los rangos de solapamiento de esta técnica. En la Tabla 5.1 se recogen los primeros cinco rangos.

Es importante destacar que el SP da lugar a solapamientos en los que la intensidad total es la suma de las fuentes que afectan a dicho punto, produciendo un efecto de rejilla similar al que se puede observar en la Figura 5.1. Esto podría inducir a pensar que en el caso de SC – en presencia de píxeles cuadrados– tendría lugar un comportamiento similar como se

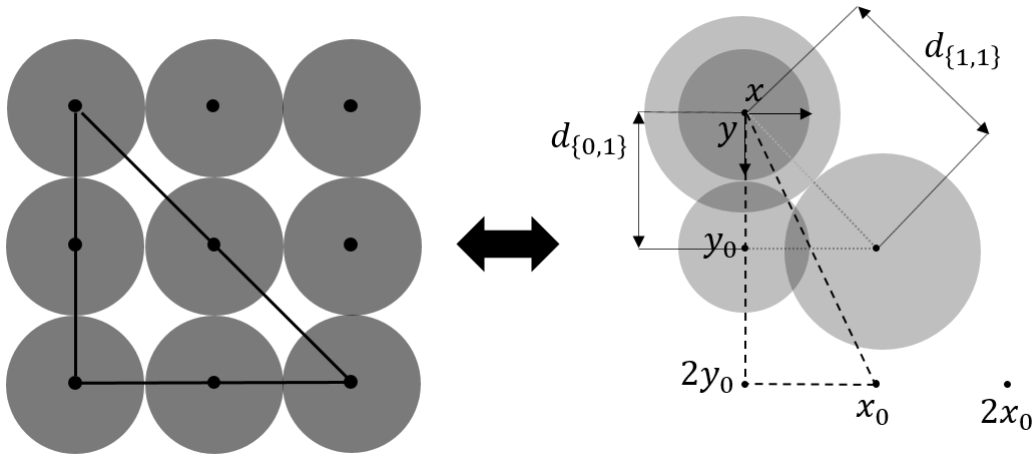


Figura 5.3 Representación visual de la estimación de los rangos de operación del SP. A la izquierda se muestra una versión esquematizada de las fuentes puntuales tras ser convolucionadas con una circunferencia de radio unitario. Se representa sobre éstas un triángulo. A la derecha se ilustran dos pares de píxeles en las coordenadas (0, 0), (0, y₀) y (0, 0), (x₀, y₀) respectivamente. El primer par conserva el círculo de desenfoque correspondiente a una circunferencia de radio unitario, mientras que al segundo par se le ha aplicado una función con radio $\sqrt{2}$.

Cuadro 5.1 Primeros cinco rangos de solapamientos en la técnica de SP obtenidos a partir de la Ecuación 5.5.

r'/r	Coordenadas de intersección del círculo (k,m)
$[1, \sqrt{2})$	(0,1)
$[\sqrt{2}, 2)$	(1,1)
$[2, \sqrt{5})$	(0,2)
$[\sqrt{5}, \sqrt{8})$	(1,2)
$[\sqrt{8}, 3)$	(2,2)

muestra en la Figura 5.2. Sin embargo, es posible demostrar que las zonas de solapamiento entre dichos píxeles no presentarían este tipo de estructura. Partiendo del razonamiento seguido en el desarrollo previo, si asumimos que $x_0 = W$:

$$\sum_{\substack{k=0 \\ m=0}}^{K,M} \text{rect}\left(\frac{x - kx_0}{W}, \frac{y - my_0}{W}\right) = 1, \quad (5.7)$$

y, por tanto:

$$I_c(x, y) = \int_0^{Kx_0} \int_0^{My_0} \text{PSF}(X - x, Y - y) dXdY = 1 \quad (5.8)$$

Ahora, tomando el razonamiento previo como base, asumamos que cada píxel toma un valor arbitrario. Debido a que la PSF está normalizada al área del círculo definido, se garantiza que, el valor de cada punto correspondiente a la integral anterior siempre va a ser menor o igual a 1 en el caso previo, y se corresponderá con el valor promedio de cada una de las vecindades, produciendo una curva suave.

A pesar de ello, la ausencia de dicha estructura no limita la interacción resultante entre los píxeles vecinos. Es posible ajustar los rangos en este caso asumiendo que el resultado de la convolución ocupa un espacio equivalente a la suma de la PSF y el píxel, con lo que se introduciría en la relación r'/r una corrección que tuviera en cuenta este hecho, ajustándose nuevamente a lo visto en la Tabla 5.1. Esto se puede expresar como sigue:

$$d'_{k,m} = 2d_{k,m} - w \quad (5.9)$$

Sin embargo, hay que tener en cuenta que esto no es totalmente cierto. Debido a la forma resultante tras la convolución, existe un rango de operación en este último caso en el que la respuesta del SC no se ajusta a la del SP. Intuitivamente, esto tiene lugar en el tramo en el que el radio de la PSF es la mitad y el doble del lado del píxel respectivamente. Previo a éste, se puede aproximar por un cuadrado, y posteriormente por un círculo. En otras palabras, dada una función de desenfoque suficientemente grande, los píxeles a convolucionar pueden ser eventualmente aproximados como fuentes infinitesimales. De la misma manera, aún partiendo de píxeles cuadrados, si la distancia entre estos es suficientemente amplia, sería posible analizar el caso como SP en lugar de SC.

5.3. Consideraciones y limitaciones de la técnica

En esta Sección se busca dar respuesta a algunas de las consideraciones previas surgidas a raíz de la implementación de las metodologías presentadas anteriormente. Éstas se derivan esencialmente de la introducción del desenfoque como función que relaciona píxeles en un sistema *a priori* independientes.

5.3.1. Relación contraste-resolución

La introducción de una función de desenfoque en el sistema plantea un intercambio con el contraste efectivo de la imagen. Desde el punto de vista del SP, se obtiene que la energía total de la fuente puntual es distribuida en el espacio a razón del radio r de la PSF, como puede observarse en la Ecuación 5.3. Por lo tanto, el máximo valor de intensidad alcanzable en este caso viene dado por el mayor número de solapamientos simultáneos, esto es:

$$I_{max} = \frac{1}{\pi r^2} \times \text{máx}\{n(x, y)\}, \quad (5.10)$$

donde n se corresponde con el número de solapamientos máximo en dicha imagen. De esta manera, el rango total que el sistema de visualización es capaz de ofrecer es $[0, \frac{n_{max}}{\pi r^2}]$ normalizado en el rango $[0, 1]$. De todo ello se deriva que la mejora del contraste del sistema se encuentra limitada por el factor r , a su vez condicionado por la distancia entre las fuentes puntuales, es decir, (x_0, y_0) . Limitando esta distancia, es posible reducir el tamaño del círculo de confusión necesario para alcanzar el mismo resultado, a expensas de mantener un rango de intensidades adecuado.

Asimismo, y como se indicó en la Sección 5.2, las zonas en las que se genera solapamiento dan lugar a áreas de mayor intensidad, por lo tanto, produciendo un efecto de *rejilla* sobre las imágenes representadas. Esto podría causar un cambio sobre el valor medio de intensidad de dicha imagen, y en consecuencia, afectar al valor de contraste eficaz, también denominado comúnmente en la literatura *root-mean-square* (RMS). Más concretamente:

$$I_{RMS} = \sqrt{\frac{1}{KMx_0y_0} \int_0^{Kx_0} \int_0^{My_0} I_c(x, y)^2 dx dy} \quad (5.11)$$

Finalmente, y en relación con ambas metodologías –SP y SC–, existe una reducción de contraste debida a la propia deconvolución (Ye et al., 2018). En esencia, introduce artefactos de mayor o menor intensidad con el objetivo de que, una vez se aplique el filtro –la PSF en el caso presente– se obtengan los valores deseados en las zonas de solapamiento. Por ello, el rango dinámico del sistema queda ligeramente reducido y condicionado al contenido de la imagen a representar.

5.3.2. Sistemas de visualización en color y estructuras alternativas

La metodología presentada se ha condicionado a dos topologías de píxeles diferentes. Por un lado, se ha analizado la introducción del desenfoque en un sistema cuyas fuentes de luz son puntuales (SP), dando lugar al solapamiento de la propia PSF con las vecindades de éstas. De forma alternativa, se ha introducido una estructura con un píxel cuadrado (SC). La imagen resultante, en definitiva, es dependiente de la estructura planteada. Existen actualmente varios tipos de ordenación de los píxeles que componen a un sistema de visualización, cuyo origen se encuentra en diferentes motivos, siendo uno de ellos la representación de color. Normalmente, aunque no de forma exclusiva, la estructura más clásica de color se obtiene por medio de tres

canales; rojo, verde y azul, comúnmente referidos en inglés como *RGB*. La adaptación de la metodología propuesta pasaría por la modelización de esta estructura y su introducción dentro de la definición de I_c , teniendo en cuenta forma y separación de las fuentes. En cualquier caso, daría lugar a una cierta cromaticidad en la imagen final, de manera similar a lo obtenido en Sajadi et al. (2013). Sin embargo, no es objeto de este Capítulo la evaluación del sistema de visualización en sí mismo, sino la aplicación de la técnica y su evaluación sin consideración previa a estructuras notablemente diferentes. Estas cuestiones se analizarán más adelante en el Capítulo 6.

5.3.3. Círculo de confusión

En la Sección 5.2 se ha asumido que el círculo de confusión de un desenfoque, es decir, la PSF propuesta, coincide con una función pistón. Sin embargo, esta aproximación tiene origen en la óptica geométrica y podría no considerarse correcta en ciertas situaciones prácticas. Por ello, partamos de la definición de la luz como campo electromagnético y definamos el campo producido por una fuente infinitesimal como $U(x, y)$ en un plano. La función de desenfoque se puede modelar como una fase aplicada sobre el frente de onda de la señal y dependiente a la focal f de la lente:

$$U(x, y) = S(x, y)U_i(x, y) \exp \left[-i \frac{2\pi}{\lambda} \left(\frac{x^2 + y^2}{2f} \right) \right], \quad (5.12)$$

siendo U_i el campo electromagnético plano producido por una fuente distante, y con λ como su longitud de onda. Definimos la función de pupila del sistema como circular debido a su uso extendido en sistemas de imagen. La Ecuación 5.12 representa la propagación de un campo U hasta la pupila del sistema. Si la propagación fuera considerada geométrica, se perdería la representación de los patrones difractivos debido a la distancia con respecto a la pupila. Si no se considera óptica de *Fourier*, se puede decir que la propagación de dicho campo será únicamente limitada por el tamaño de la pupila, produciendo un círculo perfecto en intensidad $|U(x, y)|^2$. De forma alternativa, se puede hacer uso de otras aproximaciones en la propagación del campo para tener en cuenta este tipo de efectos. Por ejemplo, es posible considerar la propagación de *Fresnel* para describir la traslación de $U(x, y)$ desde la fuente a una pupila a distancia z . Esto se consigue introduciendo en la siguiente fase en la expresión anterior:

$$U'(x, y) = \mathcal{F}^{-1} \left\{ \mathcal{F} \{ U(x, y) \} \exp \left[-i\pi\lambda z (f_x^2 + f_y^2) \right] \right\}, \quad (5.13)$$

siendo \mathcal{F} y \mathcal{F}^{-1} las transformadas directa e inversa de *Fourier*, situando (f_x, f_y) como las coordenadas en frecuencia dentro del dominio de *Fourier*. De forma ilustrativa, se representa en la Figura 5.4 un corte lateral de la función resultante tras la propagación descrita en la Ecuación 5.13. En la gráfica, se muestran tanto el resultado de la propagación geométrica como de *Fresnel*. En el caso de *Fourier*, se puede observar cómo se genera un patrón de anillos típico de este tipo de propagación. Si se considera el sistema de muestreo humano, el sistema visual, es posible concluir la validez de la aproximación. Se debe evaluar la relación entre la agudeza

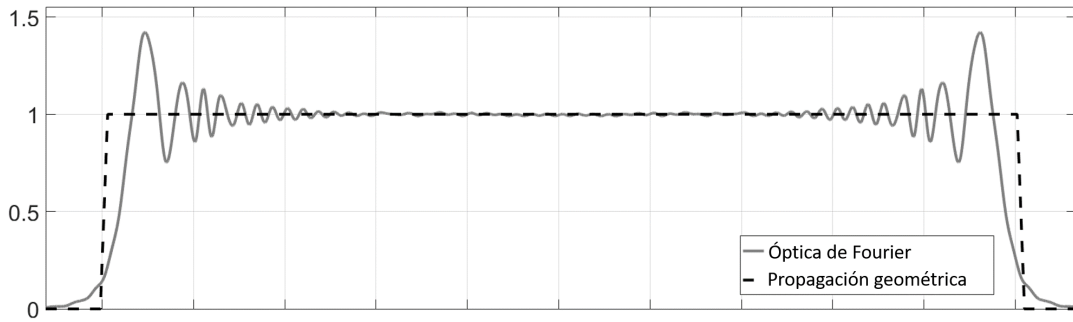


Figura 5.4 Comparación entre un corte radial del círculo de confusión generado por una propagación geométrica y de Fresnel respectivamente.

visual y el tamaño de los anillos producidos por la propagación. Se entiende por agudeza visual a la frecuencia más alta que un sistema visual humano puede percibir. Normalmente definido como un ángulo visual de un arco-minuto, sería necesario que el observador se posicionara a una distancia de $2/\tan(1/60)mm \approx 120mm$ del sistema para poder distinguir un píxel de ancho $1mm$. Si asumimos que los artefactos generados son un orden de magnitud menor que el tamaño del píxel, es posible concluir que la distancia necesaria para percibirlos no sería razonable, y por tanto, se puede aceptar la aproximación en primera instancia.

5.4. Deconvolución como pre-compensado de la imagen

A partir del razonamiento previo, en esta Sección se introducirán los métodos numéricos empleados a fin de obtener una imagen adaptada al desenfoque introducido que supere las limitaciones de ancho de banda impuestas por el sistema de visualización previo a dicha operación. Debe tenerse en cuenta que, indistintamente de si se considera la técnica SP o SC, el aparato matemático que se indica a continuación es aplicable.

5.4.1. Inversión de la PSF

Una vez se ha introducido la PSF en el sistema, se busca pre-compensar la imagen tomando como referencia las interacciones producidas. Estos solapamientos pueden ser considerados como píxeles virtuales en el espacio imagen. Sea $\rho = [\rho_{0,0}, \rho_{0,1} \dots \rho_{K,M}]$ un vector de soluciones correspondiente la imagen pre-compensada, es decir, a los píxeles que componen la pantalla. Dicho vector puede ser añadido a las Ecuaciones 5.1 y 5.5 de forma generalizada. Sea $D(x, y)$ una nueva imagen de intensidad formada a razón de la combinación lineal de dichos coeficientes:

$$D(x, y) = \sum_{\substack{k=0 \\ m=0}}^{K,M} \rho_{k,m} [f(x - kx_0, y - my_0) \otimes \text{PSF}(x, y)] \quad (5.14)$$

A partir de la Ecuación 5.14 es posible obtener un conjunto de ecuaciones como parte de un sistema lineal. Dichas ecuaciones corresponden a una imagen D , con una resolución de $P \times Q$.

$$\mathbf{D} = \mathbf{A}\rho, \quad (5.15)$$

siendo \mathbf{A} una matriz dispersa que tiene en cuenta la extensión e intensidad de cada una de las fuentes convolucionadas (kx_0, my_0) . Como se ha indicado anteriormente, el objetivo es encontrar un conjunto de soluciones ρ que satisfaga dicha relación. Esta operación, por tanto, debe entenderse como el pre-compensado de la imagen del sistema de visualización y corresponde con una deconvolución. Con esto, se busca obtener una imagen D con mayor resolución, aplicando las características de la PSF definida en la Ecuación 5.2. Para ello, se debe minimizar la expresión anterior:

$$\begin{aligned} & \underset{\rho}{\text{minimizar}} \quad \|\mathbf{A}\rho - \mathbf{D}\|_2^2 + \alpha^2 \|\rho\|_2^2, \\ & \text{sujeto a} \quad \quad \quad 0 \leq \rho_{k,m} \leq 1 \end{aligned} \quad (5.16)$$

con $\alpha \|\rho\|_2^2$ como un término de regularización de Tikhonov (Tikhonov, 1963) y α siendo el parámetro de ajuste. El sistema de ecuaciones propuesto representa un problema mal condicionado siempre que $P \geq K$ y $Q \geq M$. Esto quiere decir que no tiene una única solución en ρ , sino que existen múltiples valores de para este vector que minimizan hasta cierto punto el error. La matriz \mathbf{A} se representa como sigue a continuación:

$$\mathbf{A} = \begin{bmatrix} I_{c,k=0,m=0}(0,0) & I_{c,k=0,m=1}(0,0) & \cdots & I_{c,k=K,m=M}(0,0) \\ I_{c,k=0,m=0}(0,1) & I_{c,k=0,m=1}(0,1) & \cdots & I_{c,k=K,m=M}(0,1) \\ \vdots & \vdots & \ddots & \vdots \\ I_{c,k=0,m=0}(P,Q) & I_{c,k=0,m=1}(P,Q) & \cdots & I_{c,k=K,m=M}(P,Q) \end{bmatrix}, \quad (5.17)$$

con $I_c(k, m)$ como la imagen teórica cuando la fuente situada en (kx_0, my_0) tiene valor máximo y el resto es 0, indistintamente de la topología de estudio. A pesar de que la operación de minimización se corresponde con una deconvolución partiendo de una PSF que afecta por igual a toda la imagen, podría ser posible introducir de forma local diferentes PSFs con el fin de llevar a cabo una deconvolución zonal. Cabe destacar que, incrementar el tamaño de la PSF reduciría la dispersión de los datos en la matriz \mathbf{A} , en consecuencia, dando lugar a un problema peor condicionado.

5.4.2. Inversión multimodal

La operación de inversión presentada anteriormente se basa en la imposición de una única PSF sobre el sistema de visualización. Este planteamiento tiene su sentido experimental en el uso de una lente o un elemento óptico que produzca un cierto desenfoque. Sin embargo, teóricamente, en lugar de usar una sola PSF, es posible descomponer I_c en un subconjunto de

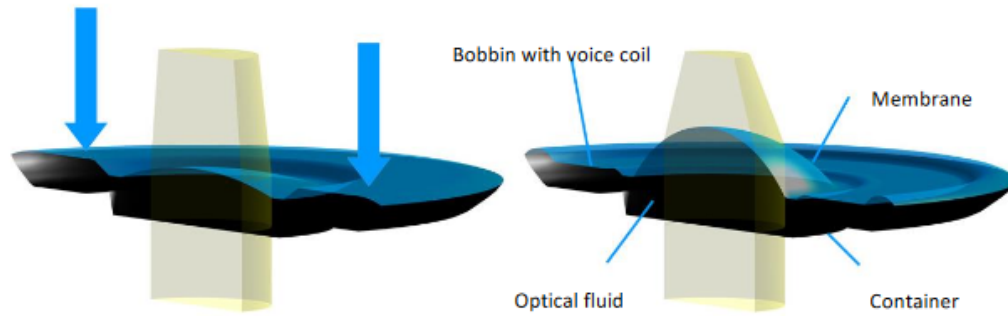


Figura 5.5 Esquema de lente líquida EL10-30d el fabricante Optotune. Imagen extraída de <https://www.optotune.com/products/focus-tunable-lenses/electrical-lens-el-10-30> a 05/01/2020

imágenes afectadas individualmente por diferentes PSFs. Llamaremos modo a cada uno de los pares imagen-PSF, y a la operación en conjunto como inversión multimodal.

De forma intuitiva, y comparativamente a otros trabajos (Lanman et al., 2010; Wetzstein et al., 2012c), esto puede entenderse como una descomposición temporal si se tiene en cuenta la persistencia retinal del sistema visual humano. Nótese que esto implicaría la presencia de un elemento óptico cuya PSF fuera variable con el tiempo. Existen en la actualidad varios dispositivos que permiten realizar este tipo de tarea. Entre ellos se encuentran las lentes líquidas, siendo estas un líquido encerrado en un espacio reducido y conectado a dos terminales entre los que se aplica un cierto valor de tensión. Esto afecta a la forma del líquido, simulando en definitiva una lente con focal variable. Dependiendo del modelo, se pueden encontrar lentes líquidas cuya respuesta oscila en torno a los 150-200Hz, superando el límite de 24Hz impuesto por el ojo (Brettel et al., 2006). En la Figura 5.5 se muestra una lente líquida del fabricante *Optotune*.

Partiendo del desarrollo previo para la inversión de la PSF, sea ρ un nuevo vector de soluciones de tamaño $n \times K \times M$, donde n es el número de modos de la inversión. Reformulando la Ecuación 5.14 en torno a la nueva definición:

$$D(x, y) = \sum_{n=0}^n \sum_{\substack{k=0 \\ m=0}}^{K, M} \rho_{k, m}^{[n]} [f(x - kx_0, y - my_0) \otimes \text{PSF}_{[n]}(x, y)] \quad (5.18)$$

Debido a lo anteriormente expuesto, la matriz A debe dar cuenta de dichos modos. Esto queda expresado de la siguiente manera:

$$\mathbf{A} = \begin{bmatrix} I'_{c,k=0,m=0}(0,0) & I'_{c,k=0,m=1}(0,0) & \cdots & I'_{c,k=K,m=M}(0,0) \\ I'_{c,k=0,m=0}(0,1) & I'_{c,k=0,m=1}(0,1) & \cdots & I'_{c,k=K,m=M}(0,1) \\ \vdots & \vdots & \ddots & \vdots \\ I'_{c,k=0,m=0}(P,Q) & I'_{c,k=0,m=1}(P,Q) & \cdots & I'_{c,k=K,m=M}(P,Q) \end{bmatrix}$$

$$\begin{bmatrix} I''_{c,k=0,m=0}(0,0) & I''_{c,k=0,m=1}(0,0) & \cdots & I''_{c,k=K,m=M}(0,0) \\ I''_{c,k=0,m=0}(0,1) & I''_{c,k=0,m=1}(0,1) & \cdots & I''_{c,k=K,m=M}(0,1) \\ \vdots & \vdots & \ddots & \vdots \\ I''_{c,k=0,m=0}(P,Q) & I''_{c,k=0,m=1}(P,Q) & \cdots & I''_{c,k=K,m=M}(P,Q) \end{bmatrix} \quad (5.19)$$

donde $[I', I'', \dots, I' \dots']$ son los diferentes pares imagen-PSF. Nótese que en la Ecuación 5.15, el tamaño de D se conserva, permitiendo realizar la operación sin necesidad de definir previamente la descomposición en sub-imágenes.

Es también importante destacar que, debido a la descomposición multimodal y el carácter aditivo de esta aproximación, existe una pérdida de intensidad global en la imagen. Esto se debe a que no es posible garantizar que cada instante de tiempo haga uso del rango dinámico completo de la pantalla. Un efecto similar se observa en Chang et al. (2018), donde proponen alternativamente introducir un panel OLED para mitigar este efecto.

5.5. Simulaciones

En esta Sección se muestran los resultados obtenidos en simulaciones para los métodos propuestos de pre-compensado. Así mismo, se estudian de forma práctica los límites y puntos de operación óptimos de la técnica para las dos topologías propuestas –SP y SC–.

5.5.1. Pre-compensado simple o con desenfoque único

Como se indica en la Sección 5.4, se busca obtener una imagen de resolución $K \times M$, tal que, cuando se representa en un sistema de visualización afectado por una cierta PSF, el observador pueda percibir una imagen de resolución $P \times Q$, cumpliéndose que $P > K$ y $Q > M$. Para este fin, es necesario llevar a cabo una minimización de la Ecuación 5.16 sujeta a la serie de restricciones indicadas anteriormente. En este caso particular, se ha hecho uso de MatlabR2016 (Matwhworks Inc.), en el que se ha implementado una solución basada en un algoritmo de minimización conocido como reflexivo de región de confianza (Coleman y Li, 1996). Concretamente, la ejecución de este programa se llevó a cabo en una computadora con un procesador Intel i7-7700 3.60GHz y 64GB de memoria RAM.

Las simulaciones se han realizado en la SP y SC respectivamente. Los casos de estudio parten de las siguientes premisas:

- La resolución base del sistema (RB) es de 128×128 píxeles ($K \times M$).
- La resolución objetivo (RO) es de 256×256 píxeles ($P \times Q$).
- Con el fin de evitar errores inducidos por sub-muestreo, la resolución de las imágenes de intensidad I_c es cuatro veces el valor de la resolución objetivo, es decir, 1024×1024 . Denominaremos a esto la resolución de muestreo del sistema (RM).

En la Sección 5.4 se obtuvo un conjunto de rangos sobre los que la técnica varía su comportamiento. Sin embargo, queda por determinar el efecto sobre la imagen de estos rangos. Por lo tanto, es necesario obtener de forma empírica el punto o puntos de operación óptimos del método para el caso SP y SC. Se ha evaluado, siguiendo el planteamiento anterior, un rango de valores dentro de la relación r'/r correspondiente a $[1.0, 4.0]$. Como función de mérito, se ha empleado el PSNR y el índice de semejanza estructural –del inglés *structural similarity index measure* (SSIM)–. Dichas métricas se han realizado sobre la imagen teórica percibida por el observador y se ha comparado con la imagen a la resolución objetivo.

En la Tabla 5.2 se muestran los resultados obtenidos en las métricas para diferentes imágenes de muestra. La parte superior de la tabla se corresponde con el caso de SP, mientras que la parte inferior hace referencia al caso de SC. Con el objetivo de comparar, se ha obtenido el PSNR y SSIM de la imagen en resolución base. Asimismo, se muestra específicamente el mejor resultado asociado al círculo de desenfoque. Como se observa, en todos los casos se ha superado a la imagen de referencia. Notablemente, los valores óptimos se han encontrado en torno a círculos de desenfoque pequeños, siendo estos los que solo causan un solapamiento.

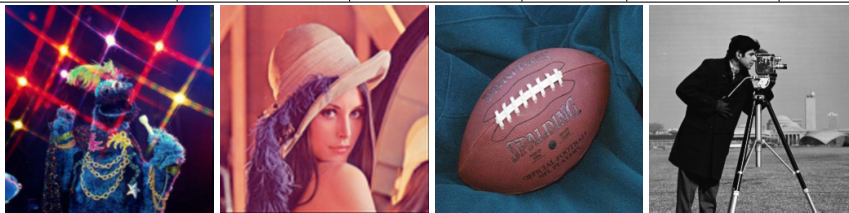
Cabe destacar que, como se indicó en la Sección 5.3, en el caso del SP se genera una rejilla con mayor intensidad superpuesta sobre la imagen. Esta estructura puede ser eliminada con motivo de la evaluación estructural del resultado, ya que su presencia no afecta directamente a la resolución obtenida, sino que se trata de un artefacto que condiciona el contraste de la imagen. Ponderando los valores obtenidos sobre el número máximo de solapamientos en cada área podemos neutralizar este efecto. De forma ejemplar, se muestra en la Figura 5.6 como, una imagen original (primera fila), da lugar a una rejilla superpuesta (segunda fila). Siguiendo el razonamiento previo, la ponderación de estas áreas de solapamiento permite distinguir la imagen resultante sin dicha rejilla.

La dependencia entre el tamaño del círculo de confusión y el incremento en resolución se muestra en la Figura 5.7, donde se ha obtenido la variación de las métricas empleadas (PSNR y SSIM) frente al cambio de radio de la PSF. Esta simulación se corresponde con el caso práctico de la imagen *Lena*. En ambas gráficas, se observa de forma generalizada que la técnica aplicada en ambas topologías muestra un comportamiento similar, encontrando un máximo en los límites de operación estimados en la Sección 5.3. Específicamente, el caso del SP encuentra su mejor resultado en torno a $r'/r = \sqrt{2}$, mientras que en el caso de SC, el valor óptimo tiende a encontrarse en las cercanías de $r'/r = 2$. Cada caso muestra una tendencia similar, con la salvedad del grado de mejora obtenido. Mientras *Lena* ha resultado con un margen de mejora elevado en torno a su máximo, *Cameraman* y *Cookie* han obtenido un rango

Cuadro 5.2 Valores de PSNR y SSIM del punto de operación óptimo obtenidos para cada una de las topologías planteadas. La parte superior de la tabla se corresponde con el caso de SP, mientras que la segunda pertenece al de SC. En la última columna se indica el radio del círculo de confusión al que pertenecen los resultados. Debajo de ésta se muestran las imágenes de estudio *Cookie*, *Lena*, *Football* y *Cameraman* de izquierda a derecha.

Muestra	Ref. PSNR (dB)	Precomp. PSNR (dB)	Ref. SSIM	Precomp. SSIM	$\frac{r'}{r}$
<i>Cookie</i>	23.66	24.15	0.90	0.92	1.42
<i>Lena</i>	29.65	31.72	0.97	0.98	1.42
<i>Football</i>	29.74	30.00	0.94	0.95	1.42
<i>Cameraman</i>	24.47	25.00	0.80	0.82	1.42

Muestra	Ref. PSNR (dB)	Precomp. PSNR (dB)	Ref. SSIM	Precomp. SSIM	$\frac{r'}{r}$
<i>Cookie</i>	23.66	24.93	0.90	0.93	2.04
<i>Lena</i>	29.65	36.20	0.86	0.99	2.00
<i>Football</i>	29.74	30.82	0.94	0.95	2.04
<i>Cameraman</i>	24.47	25.55	0.80	0.83	2.04



for to his dying day he
would not tell. He knew
for to his dying day he
would not tell. He knew
for to his dying day he
would not tell. He knew

Figura 5.6 Texto de muestra sobre el que se ha realizado el precompensado asumiendo fuentes puntuales, siendo $r'/r = \sqrt{2}$. La primera fila es la imagen a la resolución base; la segunda fila muestra el resultado tras la minimización; y la tercera fila tras ponderar con los pesos de los solapamientos.

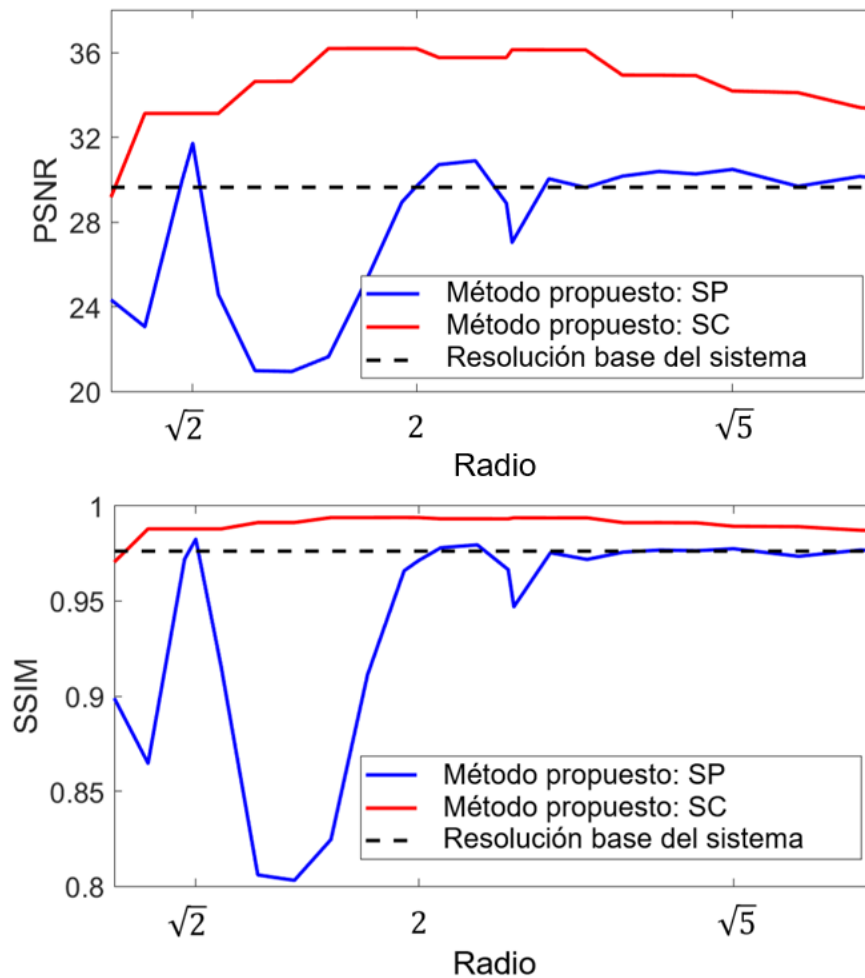


Figura 5.7 Gráfica comparativa del PSNR y SSIM para el caso SP y SC en relación a la variación del círculo de confusión

más limitado. En todas las imágenes estudiadas se ha encontrado un máximo en los valores anteriormente especificados, siempre mejorando con respecto a la resolución base. Es posible que esta discrepancia tenga que ver con la disposición de las zonas de altas frecuencias que se desea representar en cada caso y la estructura obtenida en los solapamientos (Sajadi et al., 2013).

Con el fin de ofrecer en mayor detalle cómo afecta el cambio de radio sobre el resultado final, se muestra en la Figura 5.8 una imagen que representa una carta de *Sloan*, comúnmente empleada para medir la agudeza visual de una persona (Sloan, 1959). Sobre esta imagen se ha realizado un pre-compensado con círculos de confusión cuya relación r'/r se corresponde con los límites de los primeros rangos de solapamiento en el caso del SP, esto es $\sqrt{2}$, 2 y $\sqrt{5}$. En consonancia con los resultados mostrados anteriormente, se puede observar que el incremento en resolución se maximiza en torno al primer valor, mientras que según aumenta, la imagen comienza a degradarse. Cabe destacar que, en cualquiera de los tres casos, existe un cierto margen de mejora con respecto a la resolución base.

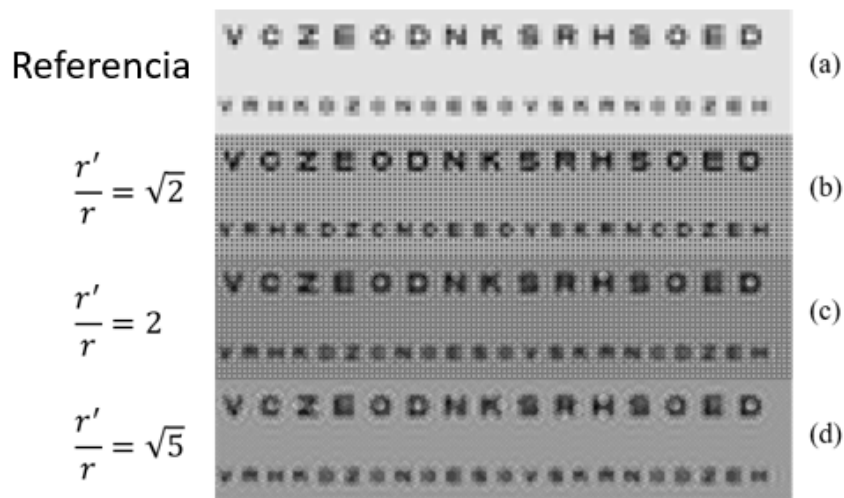


Figura 5.8 Comparativa de imágenes resultantes en el caso de SP tras pre-compensar con diferentes círculos de desenfoque (véase Ecuación 5.16). Las imágenes se corresponden con (a) la resolución base, y el resultado tras pre-compensar con (b) $r'/r = \sqrt{2}$, (c) $r'/r = 2$ y (d) $r'/r = \sqrt{5}$ respectivamente.

En la Figura 5.9 se muestran los resultados obtenidos en simulación para el SC. De acuerdo con lo expuesto en la Sección 5.3, se ha asumido que la posición de los píxeles correspondientes a cada canal de color es la misma, y ocupa el espacio de un píxel completo. Específicamente se muestran dos imágenes (*Cookie* en la fila superior y *Lena* en la fila inferior). La columna izquierda contiene la imagen representada en la resolución base como referencia. En la segunda columna se muestra el resultado tras el pre-compensado. Finalmente, se muestra también una última columna incluyendo ampliaciones de ciertos detalles con fines ilustrativos. Se han añadido los valores de PSNR y SSIM respectivamente debajo de sendos casos, siendo estos superiores en comparación con la imagen de referencia.

De forma alternativa se ha realizado un estudio de la MTF teórica del sistema en simulación. Para ello se ha asumido que la PSF del sistema de visualización se corresponde con la forma del píxel en nuestro caso: forma puntual en SP y forma cuadrada en SC respectivamente. Se ha computado la transformada de *Fourier* sobre una imagen radial compuesta por señales cuadradas de frecuencia creciente en relación a la proximidad del centro, y a la vez radializando el espectro obtenido. Asimismo, se ha normalizado la energía de la transformada con respecto a la PSF establecida anteriormente. En la Figura 5.10 se muestran los diferentes perfiles extraídos. Se ha representado con fines comparativos el resultado obtenido tanto para un sistema en resolución base como en resolución objetivo. De manera similar a lo indicado en el caso de SP, se ha realizado esta evaluación tanto en la imagen sin ponderar como en la imagen ponderada. Como se observa, ambas metodologías propuestas son capaces de superar el límite teórico del sistema en resolución base; no obstante quedando por debajo de la resolución objetivo. Cabe destacar que, en el caso en el que no se ha promediado la intensidad de la imagen

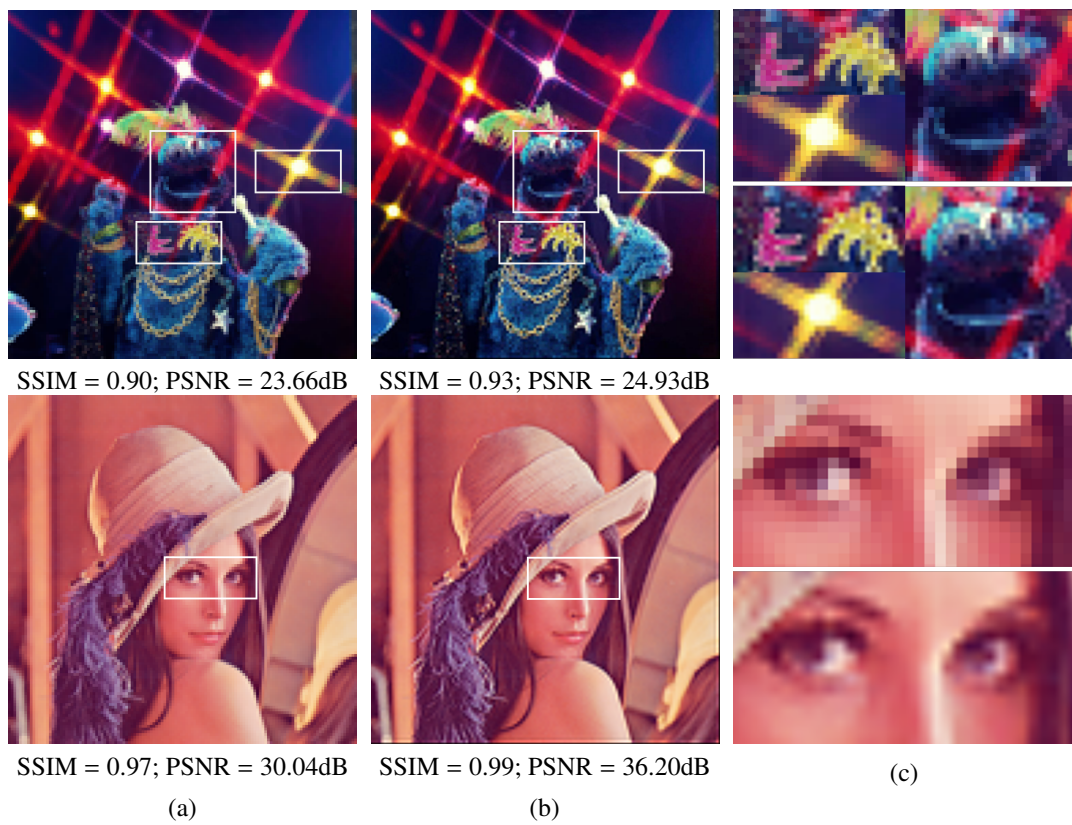


Figura 5.9 Resultados obtenidos en simulación para la resolución base en el caso de SC. Dos imágenes de muestra fueron empleadas, *Cookie Monster* (primera fila) y *Lenna* (segunda fila). La primera columna se corresponde con la imagen en resolución base. La segunda columna es el resultado obtenido después de mostrar la imagen pre-compensada. En la última columna se muestra una ampliación de algunas zonas clave de éstas.

Cuadro 5.3 Rangos de desenfoque para el pre-compensado multimodo en el caso SC

Rangos de desenfoque r'/r
$(\sqrt{2}, 2, \sqrt{5})$
$(\sqrt{2}, \sqrt{5}, 3)$
$(\sqrt{5}, 2\sqrt{2}, 3)$
$(\sqrt{2}, 1,6197, 1,8251, 2,0306, \sqrt{5})$

para la topología de SP, existe una zona de la MTF que es también inferior a la resolución base. Se hipotetiza que esto se debe precisamente a la reducción de contraste ocasionada por la rejilla impuesta sobre la imagen resultado.

Para una mejor comprensión de los resultados indicados, se muestra justo debajo de la gráfica en la Figura 5.10 el conjunto de imágenes obtenidas a partir del pre-compensado para cada técnica. De manera ordenada, se corresponden con SC, SP, SP ponderada, y resolución base. Así pues, las técnicas propuestas son capaces de representar de forma correcta frecuencias superiores cuando se compara con el sistema original.

5.5.2. Pre-compensado multimodal

Es posible evaluar la técnica propuesta de manera multi-modal, esto es, asumiendo que la imagen a representar puede ser descompuesta en un sub-conjunto de imágenes con diferentes PSFs asociadas. De manera similar al caso de una sola PSF, se estudia la técnica para el caso de SP y SC. Para no complicar excesivamente esta sección, los resultados obtenidos se han limitado al SC. Asimismo, las simulaciones realizadas conservan los mismos parámetros establecidos anteriormente.

En el pre-compensado multimodal, la elección de los círculos de desenfoque no resulta tan evidente como en el caso original. Previamente hemos evaluado el punto de operación óptimo para la técnica, encontrando este en los primeros tramos de solapamiento. Si bien esto es cierto, es posible evaluar en simulación, y de manera similar a lo presentado, diferentes casos de estudio. En la Figura 5.11 se muestra el resultado obtenido para los tres primeros rangos de solapamiento. La primera imagen se corresponde con el límite en resolución del sistema y sirve de referencia. A continuación aparece el resultado obtenido en el caso de un único modo, como se ha visto anteriormente. Debajo de estas dos últimas se indican el PSNR y SSIM con respecto a la resolución objetivo. Como se observa, la técnica propuesta mejora en ambas métricas con respecto a la resolución original del sistema.

La Figura 5.12 contiene la descomposición en instantes de tiempo del resultado mostrado en la Figura 5.11. Debajo de cada imagen se muestra el círculo de confusión asociado. En la descomposición multimodo, se ha observado que, del conjunto de imágenes, la asociada al círculo de confusión más pequeño gobierna en contraste e intensidad, por encima de las demás,

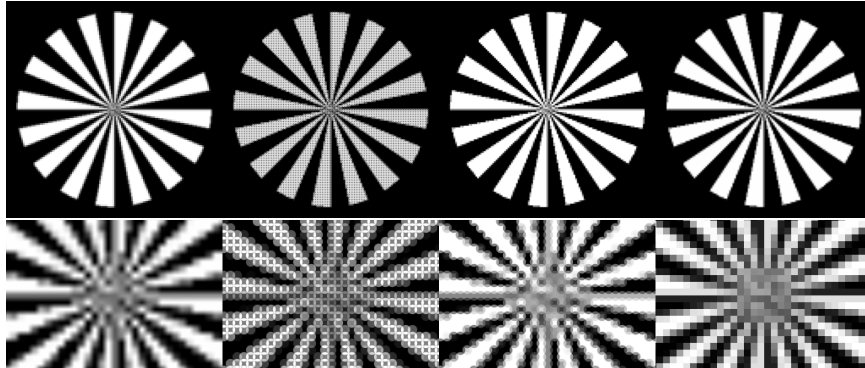
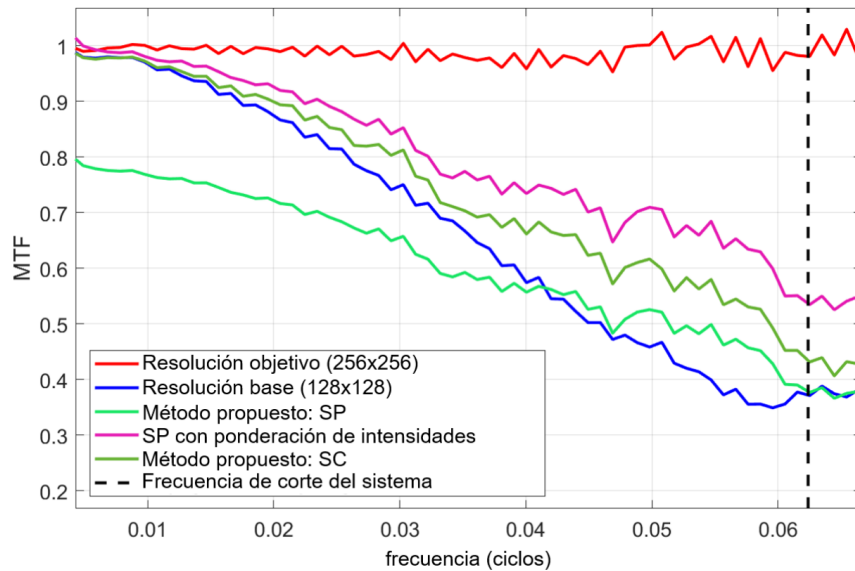


Figura 5.10 MTF del sistema para la técnica de pre-compensado. En la gráfica se muestra la MTF del sistema obtenida a partir de los patrones radiales propuestos para un sistema en (a) la resolución objetivo, (b) la resolución base, (c) el método SP, (d) el método SP con ponderación de intensidad y (e) el método SC. Debajo de la gráfica se muestran de izquierda a derecha las imágenes obtenidas de la carta radial propuesta para el caso SC, SP, SP con ponderación y sistema base.



Figura 5.11 Recuperación para el pre-compensado multimodo en el caso del SC haciendo uso de los tres primeros rangos en la Tabla 5.1. (5.11a) se corresponde con la resolución base del sistema, (5.11b) con la resolución objetivo, y (5.11c) el resultado de la operación. Debajo de (5.11a) y (5.11b) se muestran las métricas de PSNR y SSIM con respecto a (5.11c).

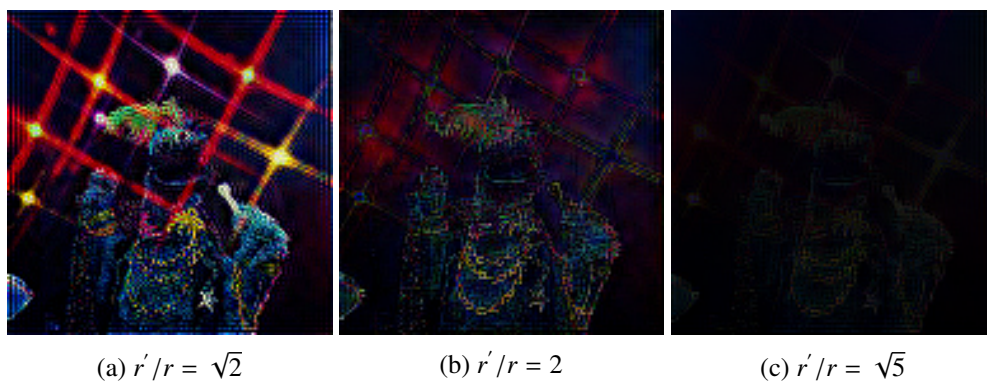


Figura 5.12 Descomposición en instantes de tiempo del método multimodo para el caso de del SC. Las imágenes están asociadas a los tres primeros rangos de la Tabla 5.1, de izquierda a derecha, respectivamente.

siendo estas atenuadas a razón del desenfoque. De manera intuitiva, se podría entender que los desenfoques mayores actúan como bajas frecuencias del sistema a diferencia de lo visto en el Capítulo 2, donde los instantes de tiempo obtienen un valor RMS similar entre ellos.

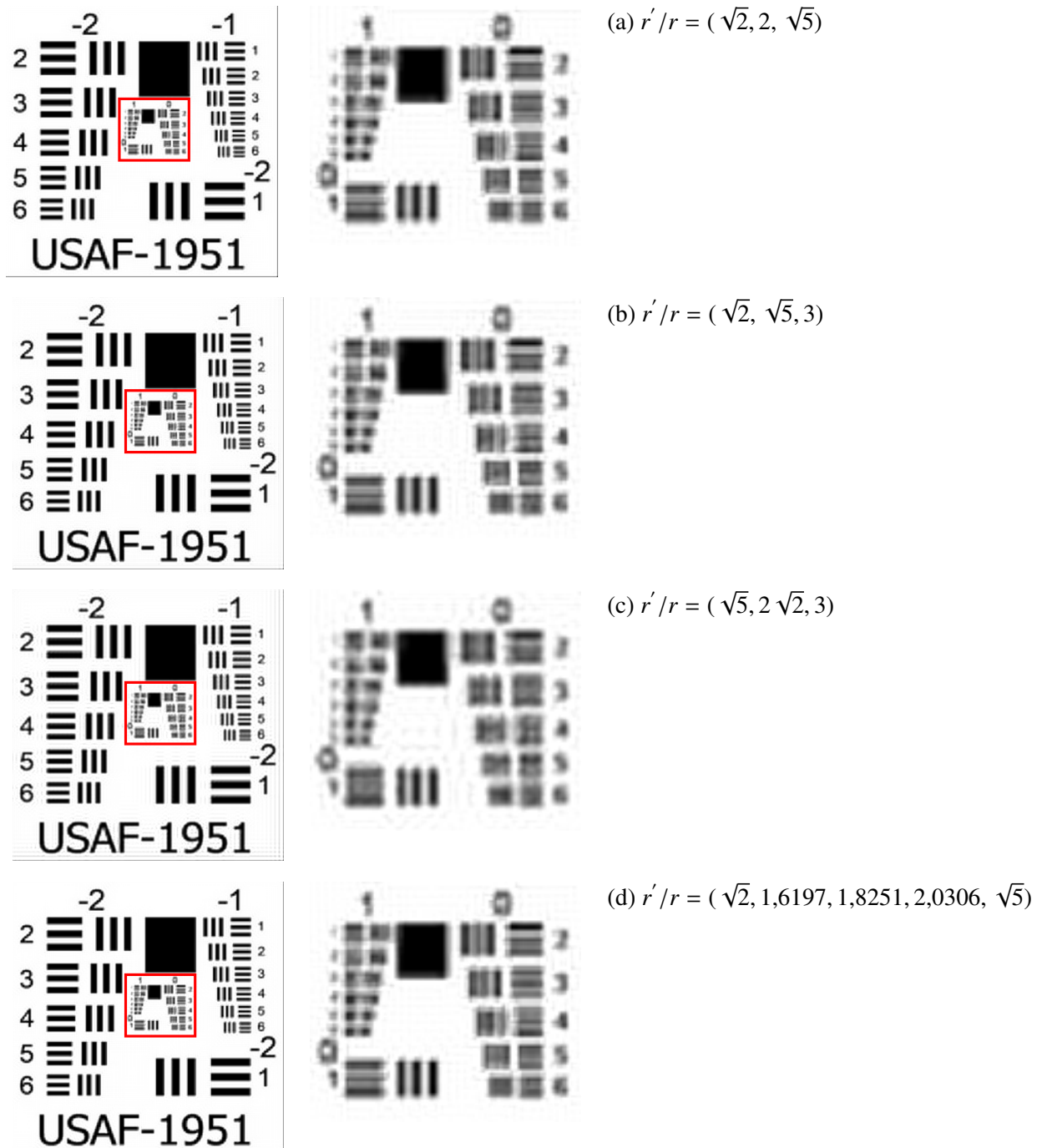


Figura 5.13 Comparativa entre diferentes rangos de desenfoque para la implementación multi-modo sobre una carta USAF-1951 en el caso SC. Se ha incluido un recorte del grupo central a la derecha de cada caso. Junto a este, se indica la relación de desenfoques empleada.

De manera más general, se ha realizado la simulación previa para diferentes combinaciones de rangos con el fin de estimar un posible punto de operación. Más concretamente, se ha aumentado el valor de la resolución objetivo a 512×512 para evitar un posible techo en términos de calidad. Las combinaciones elegidas se indican en la Tabla 5.3. Estas se han limitado a tres valores de desenfoque correspondientes a los límites de solapamiento. Asimismo, también se ha evaluado para valores arbitrarios espaciados linealmente entre $\sqrt{2}$ y $\sqrt{5}$. Los resultados obtenidos se pueden observar en la Figura 5.13. Se ha hecho uso de la carta USAF-1951, siendo ésta una imagen ampliamente empleada en el campo del procesamiento de imagen para la evaluación de la resolución de un sistema. En la columna de la izquierda se muestra la imagen final tras aplicar el pre-compensado para los rangos indicados. Junto a ésta, se ha incluido un aumento del grupo central. Como se puede observar, la mejora es similar en todos los casos, quedando algo por debajo el tercer conjunto de rangos. Este comportamiento pone de manifiesto que, siguiendo el razonamiento para el pre-compensado simple, para círculos de confusión de mayor radio, se reduce la ventana de mejora alcanzable. No obstante, es importante destacar que, en cualquiera de estos, la resolución obtenida se aproxima al valor objetivo (véase Figura 5.11), es decir, 512×512 , partiendo siempre de una resolución base de 128×128 .

5.6. Implementación sobre sistemas de visualización

En esta Sección se implementan los resultados anteriormente obtenidos en dos sistemas de visualización de uso cotidiano, como son el proyector y las gafas de realidad virtual.

5.6.1. Caso de estudio: Proyector

Una de las aplicaciones inmediatas en las que la técnica presentada puede ser validada es su implementación sobre un proyector. En esencia, un proyector ya contiene un elemento óptico regulable que permite controlar el grado de desenfoque que se introduce sobre una imagen. Esto permite, por lo tanto, elegir de forma manual el círculo de confusión que se desea introducir, y pre-compensar la imagen de acuerdo al mismo.

Para las pruebas experimentales, se ha hecho uso de un proyector EPSON EH-TW5900. Específicamente se trata de un proyector compuesto por tres pantallas LCD de 1920×1080 píxeles, sobre las que se hace pasar un haz de luz, y cuya combinación da lugar a una imagen. Como soporte para las proyecciones, se ha hecho uso de una pantalla blanca situada a una distancia de 37.5cm a partir de la lente de salida. La pantalla fue alineada con la normal del proyector con el fin de evitar distorsión de *keystone*, un tipo de deformación producida en la imagen al ser proyectada a un cierto ángulo sobre el soporte o pantalla. Con el fin de capturar el resultado, se dispuso de una cámara Sony $\alpha 6000$ con un objetivo Tamron 70-200 f2.8, posicionada a una distancia de 42.7cm de la pantalla y también alineada con ésta. El sistema de medida se muestra en la Figura 5.14. Siguiendo la estructura de la Sección anterior, se han



Figura 5.14 Montaje para la evaluación de la técnica sobre proyector EPSON EH-TW5900. Se hace uso de una cámara Sony α 6000 con un objetivo Tamron 70-200 f2.8 para la adquisición de imágenes

evaluado los casos SP y SC respectivamente. En ambos casos se ha introducido un desenfoque equivalente al punto óptimo de operación, esto es, $r'/r = \sqrt{2}$ y $r'/r = 2$ respectivamente.

Para producir las fuentes puntuales necesarias en el caso SP, gracias a la estructura del proyector, fue posible proyectar una rejilla de puntos sobre la superficie introduciendo tres píxeles de espaciado entre fuentes. En la Figura 5.15 se muestran los resultados obtenidos. La primera fila contiene una representación parcial de la rejilla de puntos sin desenfocar y desenfocada respectivamente. Se puede observar que, en el caso desenfocado, aparece una máscara debido a los solapamientos, como se indicó en la Sección 5.3. Debajo se muestra un fragmento de una imagen correspondiente a la carta de *Sloan* a la que se le ha aplicado la operación de pre-compensado. Como se puede observar, el caso en el que la imagen queda desenfocada se asemeja a las simulaciones.

De manera similar, se llevó a cabo la representación de imágenes para el caso del SC. Para simular esta estructura, fue necesario agrupar los píxeles en grupos de 4×4 . Esta aproximación fue realizada por dos motivos: simplificar de forma experimental la búsqueda del punto exacto de desenfoque y representar de la forma más cercana posible el caso estudiado en simulación. En la Figura 5.16 se muestran los resultados obtenidos para varias imágenes de muestra, correspondiéndose la de la izquierda con la resolución base, y la de la derecha con la pre-compensada. Como se observa en todos los casos, la imagen pre-compensada muestra una resolución superior a la del sistema, y se asemeja a los resultados mostrados en la Figura 5.9. Cabe destacar que en los tres casos hay una cierta reducción en el contraste.

Finalmente, y siguiendo lo presentado en la Sección 5.5.2, se ha realizado la evaluación práctica sobre la aproximación multimodal. Se ha limitado el caso de estudio al conjunto de desenfoques correspondientes a $r'/r = (\sqrt{2}, 2, \sqrt{5})$, siendo esta decisión motivada por

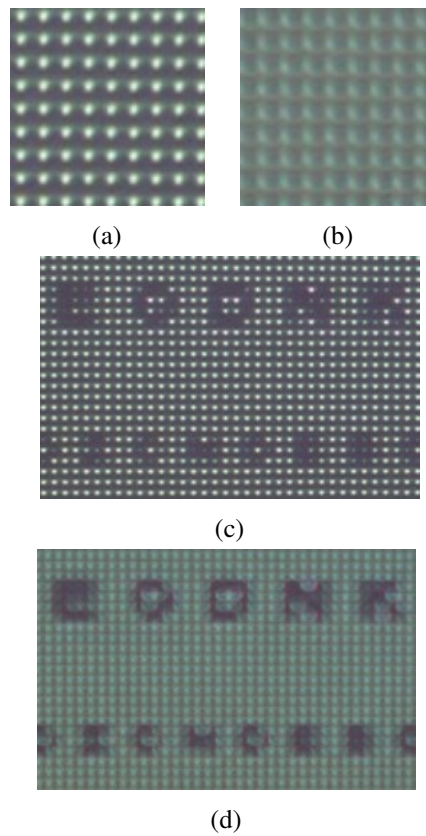


Figura 5.15 Imágenes capturadas para el caso del proyector con SP (5.15a) enfocadas, y (5.15b) desenfocadas con una relación r'/r de $\sqrt{2}$. Las imágenes percibidas por el sensor cuando se representa la imagen pre-compensada sin el solapamiento (5.15c), y con solapamiento (5.15d).

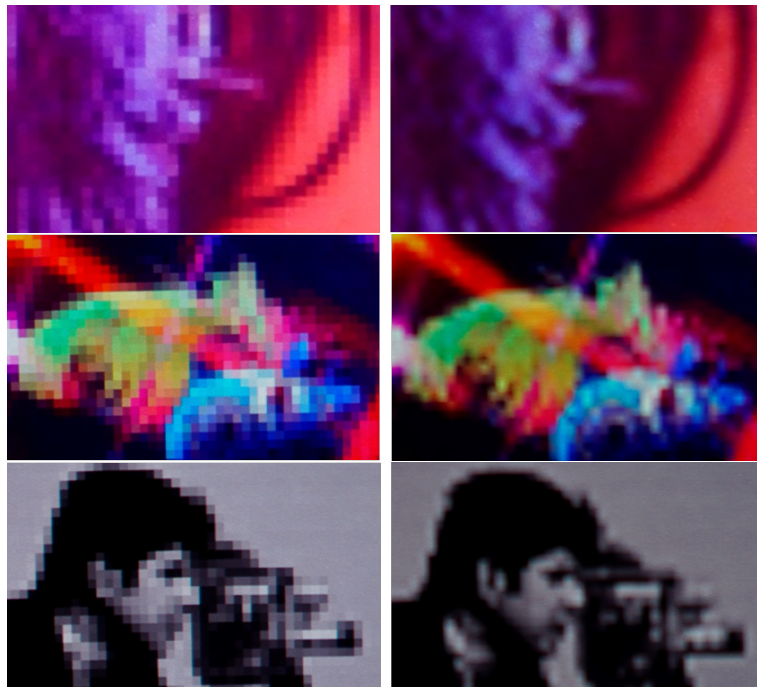


Figura 5.16 Comparación entre las imágenes capturadas para el caso SC en proyector en resolución base y pre-compensadas respectivamente.

la similitud entre los resultados obtenidos para otros conjuntos en simulación. No obstante, la metodología empleada se transfiere a los demás rangos. La adquisición de imagen se ha realizado sobre el mismo montaje ilustrado en la Figura 5.14, con la salvedad de que la imagen resultante se ha obtenido por medios computacionales, esto es, a partir de los diferentes instantes de tiempo muestreados en referencia a su desenfoque correspondiente.

Sobre la Figura 5.17 se muestra de izquierda a derecha una comparativa entre la resolución base del sistema y objetivo. El contraste de la imagen pre-compensada se reduce de manera similar al caso único. No obstante, se puede observar en las ampliaciones, cómo el resultado obtenido alcanza una representación de frecuencias muy superior a la imagen de referencia.

5.6.2. Caso de estudio: Sistema de realidad virtual

De forma alternativa, y por razones similares a las indicadas en el caso del proyector, se ha implementado la técnica en el caso de un sistema de realidad virtual. Un sistema de realidad virtual comercial incluye normalmente una lente magnificadora que permite, en esencia, reducir la distancia a la que se puede colocar el sistema de visualización del ojo del observador, facilitando así su miniaturización. La técnica propuesta se puede por tanto implementar sobre dicho sistema, asumiendo que la lente es regulable en distancia. Para las pruebas, se ha diseñado un montaje, representado en la Figura 5.18, en el que una cámara (UI-3370CP-M-GL, IDS Imaging Development Systems GmbH) se emplea para tomar imágenes de una pantalla LCD de resolución 1280x800 (HSD070PWW1-B00). En medio del camino óptico se ha situado



Figura 5.17 Comparación sobre proyector del pre-compensado multimodal para $r'/r = (\sqrt{5}, 2\sqrt{2}, 3)$. Se muestra de izquierda a derecha imagen de referencia y resultado, ambas representadas usando el mismo sistema.

un sistema de uso cotidiano de realidad virtual para terminales móviles. La regulación del desenfoque se ha realizado por medio de un objetivo de 16mm f1.4 unido a la cámara.

En la Figura 5.19 se muestran los resultados obtenidos en el caso de SP. Para representar correctamente esta topología fue necesario introducir un espaciado de un píxel entre cada fuente. En la primera imagen se muestra un texto enfocado en la resolución base del sistema. Por motivos comparativos se muestra debajo dicha imagen desenfocada y se compara con la misma pero tras ser pre-compensada con el círculo de confusión correspondiente. Como se puede observar, el texto es representado en la última imagen con mayor detalle cuando se compara a los dos casos previos. Cabe destacar que existe la presencia de una leve cromaticidad originada en el pre-compensado y el desplazamiento relativo entre canales de color.

Se ha analizado también el desempeño de la técnica en el caso SC sobre el sistema de realidad virtual. Los resultados obtenidos pueden observarse en la Figura 5.20. A la izquierda se muestra la imagen en resolución base. A pesar de que los píxeles siguen una estructura cuadrada, estos quedan ligeramente separados por un intersticio negro. Este efecto es conocido como *screen-door* en inglés por su similitud con una ventana y es debido a la representación de los píxeles como mosaicos, y no como un plano continuo. Se trata de un problema que ha sido evaluado múltiples veces en la literatura, y que afecta en mayor medida a los dispositivos de realidad virtual, ya que la pantalla es observada desde mucho más cerca debido a la lente magnificadora. Una forma de mitigar este efecto se encuentra precisamente en el empleo de un desenfoque controlado. Como se observa en la segunda columna de la Figura 5.20, en la misma imagen tras aplicar el desenfoque, se mitiga en gran medida el efecto anteriormente citado. Aplicar la técnica propuesta de pre-compensado permite no solo mejorar la resolución del sistema, sino ocultar por completo esta estructura. Esto se puede ver en la tercera columna.

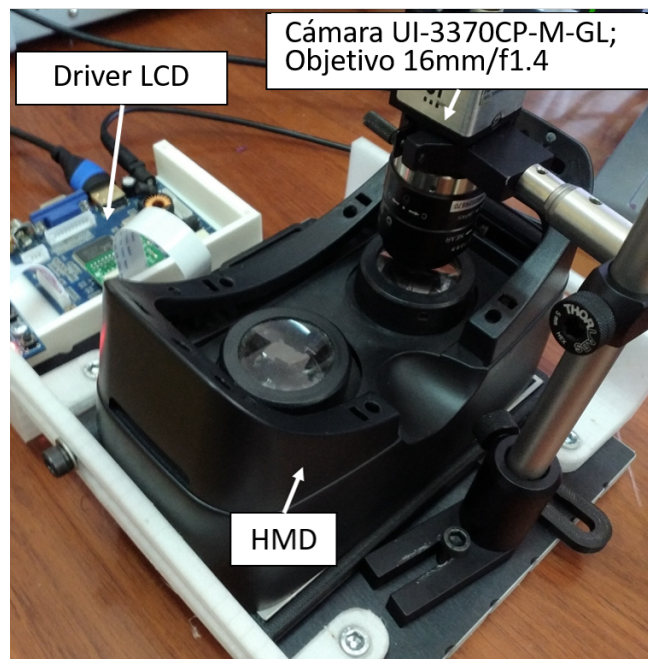


Figura 5.18 Montaje del sistema para realidad virtual. Una cámara IDS se encuentra montada en la posición del ojo derecho. La pantalla está adherida a la parte frontal del las gafas.

Una observación similar fue realizada por Zhang y Nayar (2006) en el caso de un proyector cuya rejilla era percibida en diferentes profundidades.

5.7. Sumario y Conclusiones

En este Capítulo se ha presentado una metodología basada en la deconvolución o pre-compensado de una imagen a razón de la introducción de un desenfoque artificial. Dicho análisis ha sido realizado en el contexto de los sistemas de visualización bidimensionales con el fin de mejorar la resolución lateral o espacial de estos. La introducción de un desenfoque ha permitido generar interacciones entre los píxeles del sistema de tal forma que posteriormente es posible hacer uso de estas con el fin de mejorar la resolución alcanzable. A este proceso se le denomina pre-compensado y ha sido anteriormente estudiado en la literatura (Damera-Venkata y Chang, 2007, 2009; Zhang y Nayar, 2006) con múltiples fines, entre ellos, la mejora de la resolución.

Se han analizado dos topologías de píxeles diferenciadas. En primer lugar, se ha presentado el análisis para sistemas cuya estructura de píxeles está formada por fuentes puntuales separadas una cierta distancia (SP). Alternativamente, hemos generalizado a una estructura más común, considerando los píxeles como fuentes cuadradas (SC). A diferencia de los trabajos previos, la metodología presentada se basa únicamente en la introducción de un desenfoque controlado, siendo este el encargado en introducir las interacciones entre píxeles. Se ha obtenido un punto de operación óptimo de desenfoque, y se ha evaluado en simulación la mejora en resolución obtenida para cada caso. Para ambas topologías se ha obtenido una mejora sustancial.

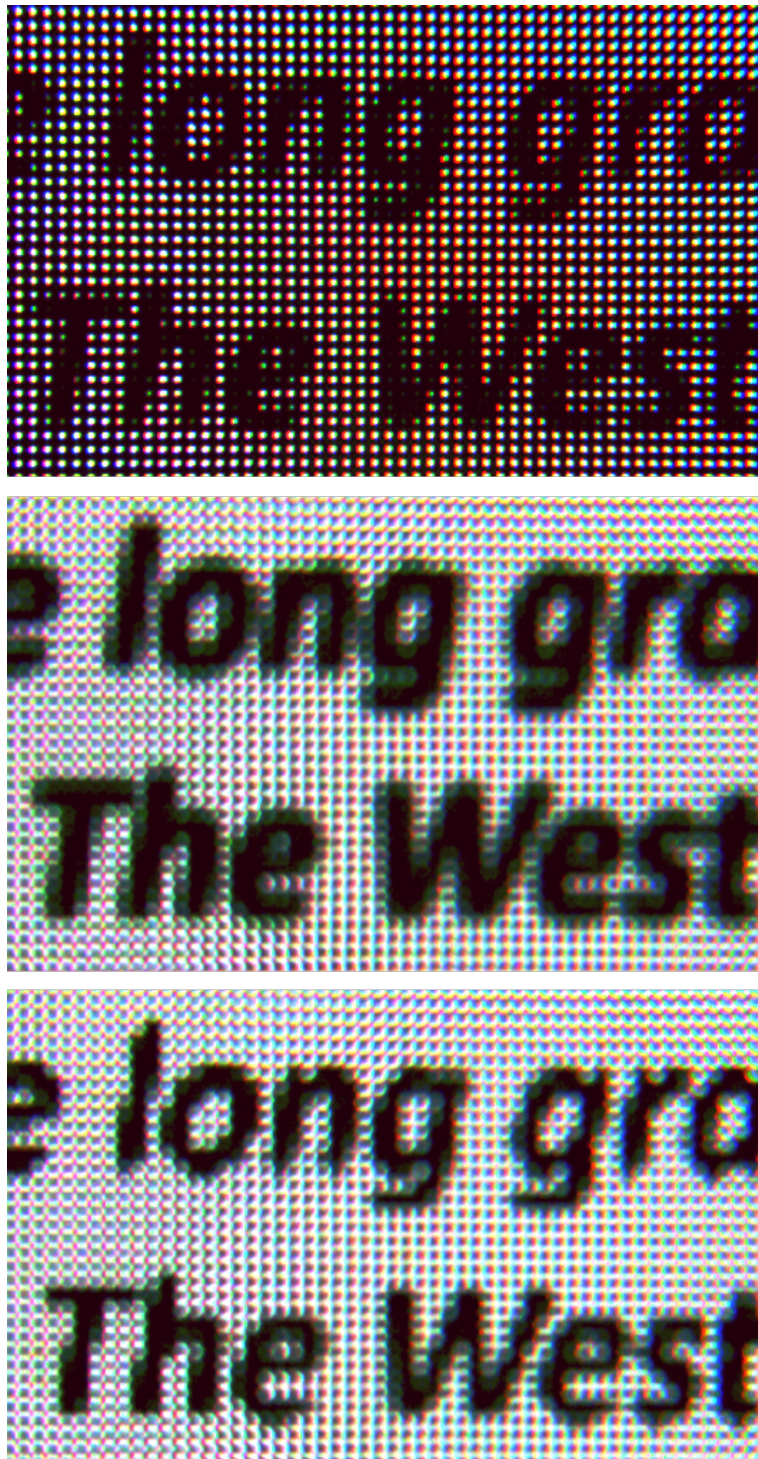


Figura 5.19 Resultados obtenidos para el caso de SP en el montaje de realidad virtual. De arriba a abajo, imagen en resolución base sin desenfoque, imagen sin pre-compensar desenfocada, e imagen pre-compensada desenfocada.

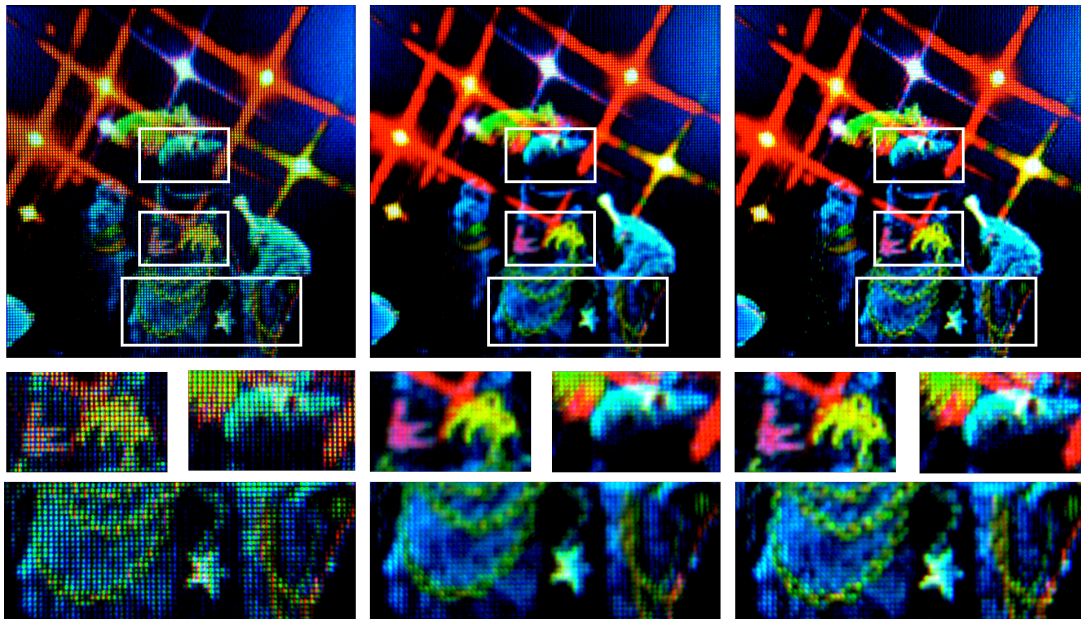


Figura 5.20 Resultados obtenidos para el caso de SC en el montaje de realidad virtual. De izquierda a derecha, imagen sin pre-compensar enfocada, imagen sin pre-compensar desenfocada, e imagen pre-compensada desenfocada. Debajo se muestra un conjunto de ampliaciones a modo comparativo.

A diferencia de trabajos anteriores que ya proponían el desenfoco como método para incrementar la resolución del sistema, en este trabajo se han analizado los límites de operación, así como se ha obtenido un punto óptimo para el método analizado. Específicamente, se ha partido del estudio de una topología de SP, cuyo análisis ha permitido una generalización sobre el caso de SC, encontrando en ambos una mejora en resolución.

Asimismo, partiendo de dicha metodología, se ha introducido la multiplexación temporal, de manera similar a otros trabajos Chang et al. (2018), con el fin de incrementar aún más la ganancia en resolución. Esta aproximación resulta muy prometedora, pues mejora considerablemente el resultado si lo comparamos con lo obtenido para la técnica sin multiplexación. Se ha observado que, a pesar de ello, existe un cierto techo en torno a los 3 modos o instantes de tiempo. A partir de esta cantidad, la mejora resulta poco significativa. Específicamente, se ha determinado que la combinación de diferentes círculos de desenfoco óptimo parte de los valores inferiores de la 5.1. No obstante, en un trabajo futuro resultaría interesante evaluar de forma analítica el alcance real de la técnica.

Para demostrar la viabilidad de la técnica, ésta se ha implementado sobre dos sistemas de uso relativamente extendido como son las gafas de realidad virtual o los proyectores. Se ha observado experimentalmente que los resultados obtenidos para ambas topologías, SP y SC, existe una similitud elevada con los valores obtenidos en simulación. Asimismo, y siguiendo las conclusiones obtenidas por Zhang y Nayar (2006), se ha observado que la técnica de pre-compensado en el caso SC permite mitigar el efecto de *screen-door*. No obstante, el análisis experimental se ha realizado desde un punto de vista meramente cualitativo, pues la

evaluación de este tipo de técnicas resulta compleja, especialmente en la elección y aplicación de métricas válidas (Ulichney et al., 2010).

El trabajo realizado sirve como una primera introducción, pues solo considera topologías concretas. Si bien su aplicabilidad en estos casos queda demostrada, en el siguiente Capítulo se optará por analizar estos sistemas conociendo *a priori* como se distribuyen sus píxeles.

Capítulo 6

Caracterización de respuesta al impulso de sistema integral

Tan exacto es que para el sabio no hay nada comparable al hecho descubierto por él que no se hallará acaso un investigador capaz de cambiar la paternidad de una conquista científica por todo el oro de la tierra.

– Santiago Ramón y Cajal

En este Capítulo se estudiará, de forma similar a lo presentado en el Capítulo 5, cómo es posible utilizar la información conocida *a priori* de un sistema de imagen integral para mejorar la representación con respecto a un observador. De este estudio se derivan diferentes aplicaciones, como la corrección de vista defectuosa en un observador, mejora de la resolución, y sistemas multifoco. La técnica empleada no busca la coexistencia de la corrección visual y la representación de imagen tridimensional, sino que sustituye una por otra.

6.1. Consideraciones iniciales y trabajos previos

El uso de sistemas de InI no se ha limitado históricamente en la bibliografía a la representación de información tridimensional. Específicamente, y como se planteaba en el Capítulo anterior, existen implementaciones que buscan hacer uso del ancho de banda del sistema a favor de corregir defectos visuales. Tal es el caso de Pamplona et al. (2012), quienes plantean una técnica basada en un pre-compensado tomando como base un sistema con una rejilla de microlentes superpuesto. Son varios los estudios subsecuentes que parten de dicha base. Por ejemplo, Huang y Wetzstein (2014) proponen un híbrido entre el sistema multicapa y la metodología previa con el objetivo de mitigar la falta de contraste en el resultado final.

La mayoría de estos trabajos introducen una formulación matemática del sistema para, con posterioridad, hacer uso de ella en forma de un problema inverso. En el caso presente, se busca establecer un planteamiento alternativo, basado en la medición del sistema en una estructura que denominaremos *captura-representación*. Siguiendo lo expuesto en el Capítulo 5, se presenta en cierto sentido una generalización del método de pre-compensado, siendo en este caso particularizado para un sistema de imagen integral. Esta metodología es similar a lo propuesto por Trujillo-Sevilla et al. (2014) para la mejora de resolución de una cámara plenóptica por medio de la caracterización de su función de transferencia.

6.2. Respuesta impulsiva (IR) del sistema de imagen integral

En esta Sección se tratará de establecer las relaciones existentes entre el sistema de imagen integral y la adquisición de las respuestas en intensidad del mismo.

6.2.1. Respuesta en intensidad

Supongamos un sistema de imagen integral unidimensional compuesto por una pantalla situado en el eje x . Sea una rejilla de microlentes superpuesta a dicha pantalla de tal forma que el eje u en que se encuentra sea paralelo a x a razón de una focal f_m . Este sistema es registrado por un sistema de adquisición situado a una distancia d de dicho eje, y es teóricamente modelado por una lente con una cierta focal f_l . Justo detrás se ubica un sensor sobre un eje ξ . Aplicando óptica geométrica, es posible modelar de forma teórica la traslación de un píxel desde la pantalla del sistema hasta la retina del observador como:

$$\xi(x, u) = d \frac{s_i - f_l}{s_i} \left(u - \frac{x - u}{f_m} \right), \quad (6.1)$$

con s_o y s_i representando el píxel en el plano objeto e imagen respectivamente. Se define s_i como:

$$s_i = \left(\frac{1}{f_l} - \frac{u f_m}{x - u} + d \right)^{-1} \quad (6.2)$$

En la Figura 6.1 se muestra una representación esquemática del trazado de rayos del sistema planteado. Con esto se define la primera parte del método *captura-representación*, encargado de obtener del sistema de imagen integral un conjunto de respuestas al impulso (IR) para cada uno de los píxeles.

La traslación de píxeles desde la pantalla hasta el sensor indicada en la Ecuación 6.1 se puede entender como la conversión de un sistema de imagen integral a otro puramente bidimensional. Nótese que la evaluación de todas las respuestas lleva asociada una posición del observador fija, y por tanto, la representación de imagen queda limitada al espacio de la captura de similar a trabajos previos (Huang y Wetzstein, 2014; Huang et al., 2011; Pamplona et al., 2012).

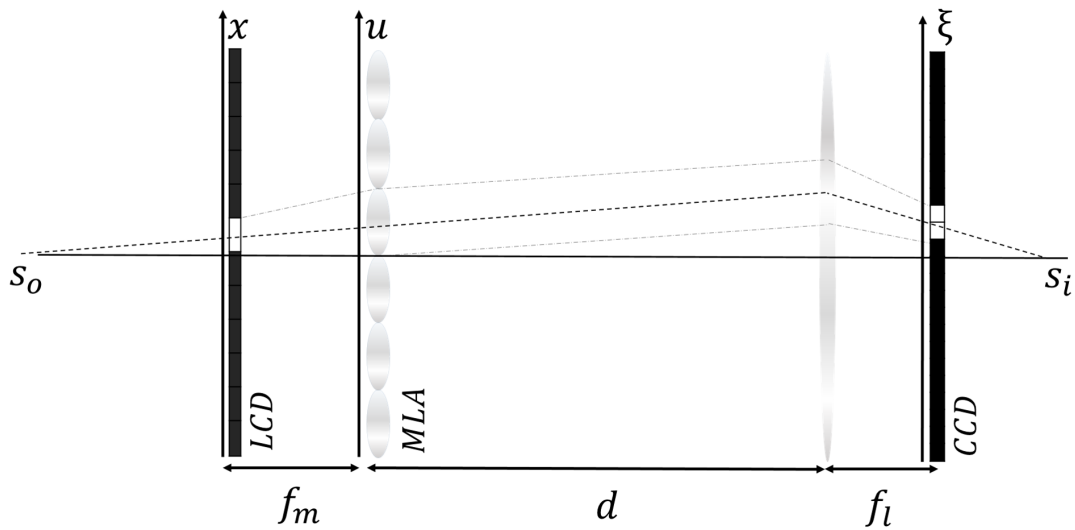


Figura 6.1 Esquema del método *captura-representación*. Una rejilla de microlentes se superpone a una pantalla, separados por su distancia focal f_l . La respuesta al impulso (IR) de cada píxel es obtenida en el sensor con una cámara a una distancia d del sistema.

6.2.2. Multiplicidad de la respuesta y distancia de muestreo

El planteamiento de la técnica de *captura-representación* tiene origen en diferentes observaciones realizadas sobre los sistemas de InI tradicionales. Este tipo de sistemas tienen un límite en el ancho de banda espacio-angular muy estricto (Chai et al., 2000). Asimismo, debido a esto, existe un criterio definido para evitar la presencia de submuestreo o *aliasing* (Zwicker et al., 2006), normalmente patente en forma de réplicas o *crosstalk* (Hoshino et al., 1998).

En el caso de estudio presente, se parte de la premisa de que al observar el dispositivo de imagen integral a una distancia dada, es posible obtener información que en otros casos no sería perceptible. El uso de dicha información puede ser considerado análogo a lo planteado en el Capítulo 5, donde las respuestas percibidas a razón de píxel pueden afectar a su vecindad, y por tanto producir información nueva. No obstante, este razonamiento lleva a establecer una distancia crítica de muestreo con base en diferentes consideraciones. Un ejemplo de este efecto puede ser observado en la Figura 6.2, donde para un mismo píxel contenido en el LCD, se generan diferentes contribuciones en el sensor.

En primer lugar, se debe garantizar la propagación del campo desde el sistema hasta el sensor, o más concretamente la retina del observador. Para simplificar, consideraremos que el plano ξ se encuentra situado en la superficie de la lente. Esto supondría que todos los rayos que atraviesan la pupila forman imagen en el sensor, lo cual no es estrictamente necesario. No obstante, posteriormente veremos que este criterio es válido en la práctica.

Siguiendo la aproximación de lente fina, y asumiendo que el u es conjugado de la focal de la lente del sistema de adquisición, se puede definir la siguiente expresión:

$$d \frac{x - u}{f} = \xi - u \quad (6.3)$$

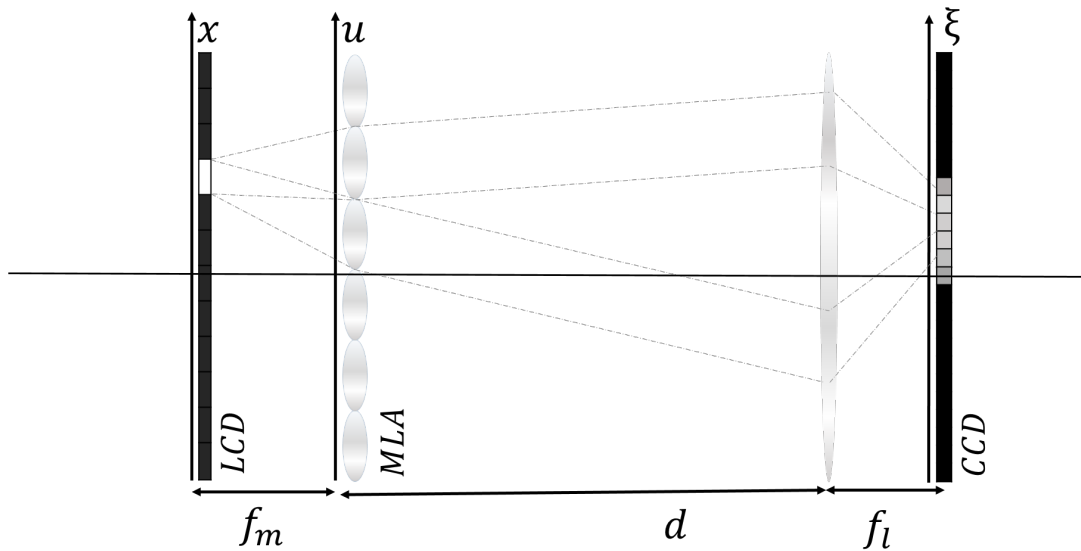


Figura 6.2 Representación esquematizada de la obtención de múltiples respuestas en el sensor a partir de un único píxel.

De esta manera, se definen todos los rayos que hacen imagen en la retina del observador. Si definimos la apertura del sistema como A , estando esta centrada en $\xi = 0$, se puede decir que los valores límite de la propagación geométrica se encuentran en $\xi_{min} = -A/2$ y $\xi_{max} = A/2$. Despejando ξ en la Ecuación 6.3:

$$\xi_{min} < d \frac{x - u}{f_m} - u < \xi_{max} \quad (6.4)$$

Siguiendo el razonamiento anterior, existen dos pares (x, u) tales que proyectados a una distancia d constituyen los extremos de la imagen formada en ξ . Más concretamente, estos vienen dados por u_{min} y u_{max} y, en coordenadas locales respecto a los centros de las lentes, por x_{max} y x_{min} . Es decir, para las lentes en los extremos, los píxeles subtendidos hacia el interior del sistema son los que producen, siguiendo la aproximación de óptica geométrica, los extremos de la imagen.

No obstante, intuitivamente se puede ver que existen multitud de soluciones válidas a la desigualdad en la Expresión 6.4. Dada una apertura cuyas dimensiones se correspondan con el ancho del plano de las MLA, es posible considerar $d \rightarrow \infty$. Esto significa, teóricamente, y siguiendo los principios de la InI, que solo se formará imagen de los píxeles situados en el eje óptico de cada microlente. Si bien este puede ser un caso de estudio, uno de los objetivos inherentes a la aplicación de esta metodología se basa en el aprovechamiento de información angular que normalmente se perdería en un sistema de visualización bidimensional.

Si se reduce el valor de d , el número de píxeles subtendidos por cada microlente aumentaría de forma teórica, provocando lo que en un sistema de InI se conoce como *crosstalk* o aparición de réplicas. En el caso de imagen tridimensional, este efecto se puede entender como la

percepción de vistas no pertenecientes al ángulo de observación, y viene determinado por la periodicidad de la función de *light-field* al ser representada en imágenes elementales en estos sistemas (Hoshino et al., 1998). Sin embargo, al muestrear esta información por el sistema de captura, y reducir el sistema a la representación bidimensional, esto deja de ser un inconveniente, y en última instancia define el número de píxeles que se muestrean por lente. Intuitivamente esto supone, como en el Capítulo 5, un conjunto de interdependencias que pueden ser empleadas para realizar diferentes mejoras con respecto al sistema base. Por ello, además de cumplir la condición impuesta por la Ecuación 6.4, es posible tomar en cuenta el valor de la distancia de muestreo de tal forma que obtengamos un número de respuestas variable.

6.3. Deconvolución basada en IR

En cierta forma, esta metodología guarda una similaridad con la técnica de solapamiento presentada en la Sección 5. Esto se debe a que, en ambas técnicas, se está introduciendo un elemento óptico que, en última instancia, origina una interacción entre los píxeles que componen una pantalla. Bien en el caso del previo, por medio de desenfoque, o bien en éste por la introducción de un elemento más complejo como es la rejilla de microlentes. La diferencia esencial entre ambos casos es que en la presente técnica, se tiene un conocimiento absoluto de cómo reacciona un sensor –tratado teóricamente como la retina de un observador– al impulso de cada uno de los píxeles que componen la pantalla, estando estos afectados por una función *a priori* no conocida. El muestreo de dicha función es lo que, nuevamente, permite operar de forma similar a lo presentado en el caso del pre-compensado para mejora de la resolución. Con dicha información es posible plantear un sistema de ecuaciones lineales que definan la relación píxel-observador, y por tanto, teóricamente su inversión presenta la posibilidad de representar una imagen de forma más precisa sobre el sistema de visualización.

Siguiendo la nomenclatura establecida en el Capítulo 5, sea $n=K \times M$ la resolución del área subtendida por la rejilla de microlentes, y por tanto el número total de IR, a la cual denominaremos nuevamente resolución base. De la misma manera, sea $s=P \times Q$ la resolución del sistema de adquisición, es decir, la de cada IR. Definimos una imagen objetivo **I** como la combinación lineal de las IR de tal forma que:

$$\mathbf{I}_p = \sum_s^S \mathbf{R}_{s,n} \cdot \rho_n, \quad (6.5)$$

siendo **R** una matriz donde cada columna contiene la respuesta vectorizada de cada píxel. El vector ρ equivale nuevamente a la combinación lineal de las respuestas tal que, una vez proyectadas, dan lugar a la imagen **I**. Comparativamente, el sistema presentado en la Ecuación 6.5 se corresponde al pre-compensado estudiado en el Capítulo anterior, con la salvedad de que, en este caso, las IRs no han sido convolucionadas con una PSF fija, sino que se obtienen condicionadas a las aberraciones propias del sistema.

La resolución de este problema se puede nuevamente plantear como un problema de mínimos cuadrados:

$$\min_{\rho} \|\mathbf{I} - \mathbf{R}\rho\|_2^2, \quad (6.6)$$

6.3.1. Adaptación de la IR como PSF del ojo humano

En la Ecuación 6.5 se plantea una función de transferencia que contiene todas las aberraciones implícitas del sistema de adquisición empleado durante la captura de las IRs. Esto puede entenderse, a diferencia del pre-compensado con una única PSF, como una deconvolución variante en el campo. A pesar de que la función de transferencia viene definida por la captura, sería posible, de forma computacional, introducir aberraciones de alto orden cuya introducción de forma experimental no sería fácilmente realizable:

$$\mathbf{R}'_n = (\mathbf{R}_n \otimes \text{PSF}_n) \odot A_n, \quad (6.7)$$

siendo PSF un conjunto de aberraciones definidas de manera teórica e introducidas computacionalmente sobre una matriz R ya definida. Se introduce asimismo la apertura A correspondiente a cada respuesta.

6.3.2. Sistema multifoco

Siendo el planteamiento inicial de la técnica la mejora de calidad en un sistema –bien en forma de un aumento de resolución o corrección de defectos–, es posible considerar dicha metodología de manera alternativa para la representación de pistas focales. En trabajos previos se ha evaluado la posibilidad de representar la información de foco en los sistemas de visualización (Maimone et al., 2013; Pamplona et al., 2012; Zhang y Nayar, 2006). Esto tiene origen principalmente en el conflicto vergencia-acomodación, presente mayoritariamente en los sistemas de imagen tridimensional Hoffman et al. (2015).

Supongamos que se cuenta con un conjunto de IRs para un rango correspondiente a diferentes potencias ópticas de desenfoque. Es posible entonces, a partir de un mapa de distancias DM, redefinir sobre una escena en particular la matriz \mathbf{R}' . Más concretamente:

$$\mathbf{R}'_n = \mathbf{IR} [DM_n] \quad (6.8)$$

De esta manera, resolviendo la Ecuación 6.6, se obtendría un ρ dependiente del mapa de distancias. Si se aplica el producto directo 6.5 para cada desenfoque del sistema, es posible recuperar una imagen enfocada a una profundidad concreta.

Siguiendo lo anteriormente expuesto, teóricamente es posible generar un sistema multifoco. Un requisito clave en estos sistemas es la continuidad de la representación en profundidad. Trabajos previos han analizado los requisitos necesarios para alcanzar esta representación, siendo uno de los más relevantes la densidad de enfoques o *focal stack*. Particularmente,

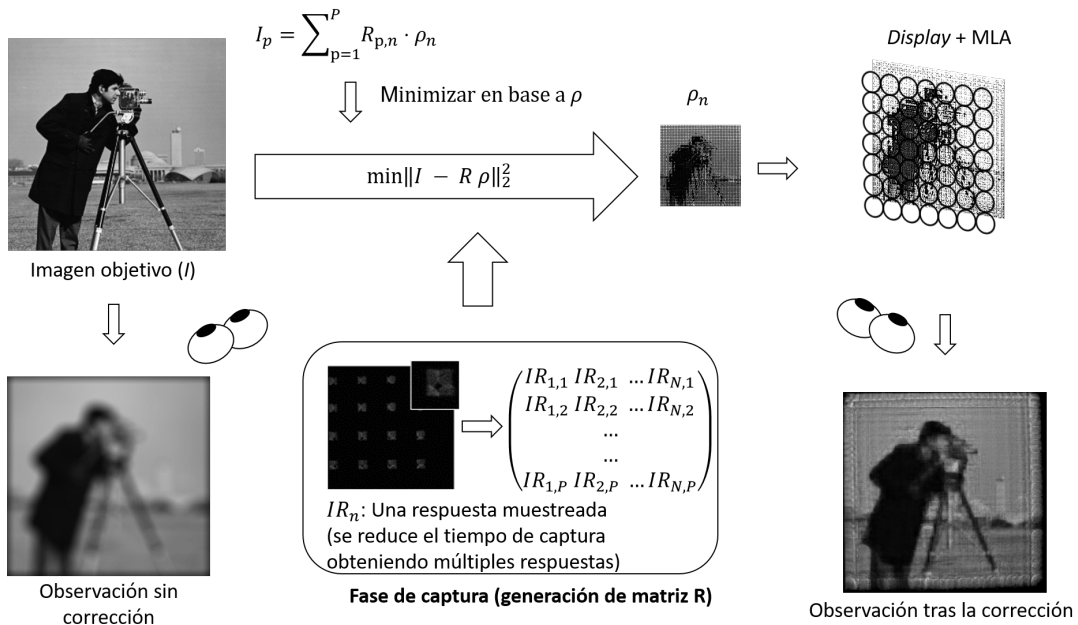


Figura 6.3 Esquema del método *captura-representación*

trabajos como los realizados por MacKenzie et al. (2011, 2010) y Chang et al. (2018) evalúan la importancia de la densidad de planos focales sobre un sistema multifoco, estableciendo un límite para alcanzar dicho requisito.

Comparativamente, en el sistema que planteamos, este rango de focales es virtualmente infinito. En el caso particular de Chang et al. (2018), este rango viene estrictamente ligado al refresco de pantalla del sistema. Veremos más adelante como con esta metodología es posible alcanzar los rangos de acomodación naturales de una persona, girando estos en torno a los $[25cm, \infty)$.

6.4. Implementación y resultados experimentales

La técnica propuesta es un método eminentemente experimental, así pues su validación requiere de un montaje adecuado. En esta Sección se presentan los resultados experimentales para el caso de la obtención de respuestas al impulso en intensidad. No ha sido posible la evaluación de la técnica en fase debido a la necesidad de un sensor de fase. La complejidad técnica de su implementación ha limitado su evaluación, y por tanto queda pendiente en forma de trabajo futuro. De forma ejemplar, se muestra en la Figura 6.3 un diagrama de evaluación de la técnica.

6.4.1. Generación de IR

En la Figura 6.4 se muestra el montaje básico de un sistema para la adquisición de las IRs. Siguiendo el esquema propuesto en la Figura 6.1 se ha situado una rejilla de microlentes, de focal $f_m = 3mm$ y apertura cuadrada de $1mm$ por lado, sobre una pantalla de resolución

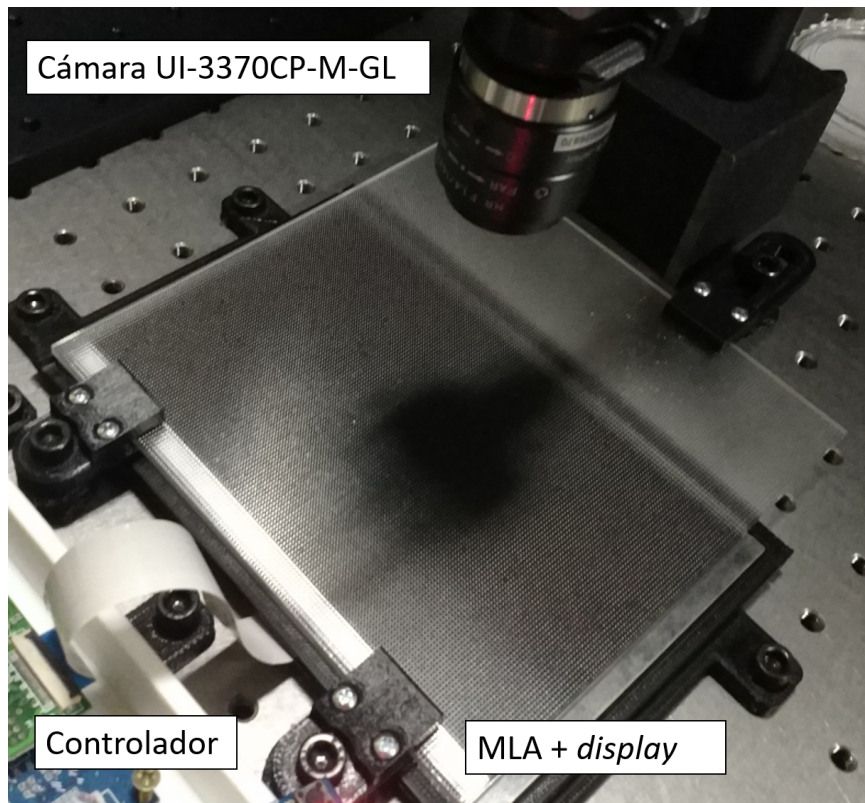


Figura 6.4 Montaje experimental del sistema *captura-representación*

1280x800 (HSD070PWW1-B00) a una distancia de 5mm . Como sistema de adquisición, se ha hecho uso de una cámara UI-3370CP-M-GL (IDS Imaging Development Systems GmbH.). La cámara fue dispuesta a una distancia de 15cm de las microlentes. Asimismo, se le superpuso un objetivo de 17mm de focal. La resolución base del sistema, esto es, el parche de pantalla a evaluar, de 256×256 píxeles.

Se realizó la captura de tres conjuntos de IRs, correspondientes a un observador con 0, 2 y 4 dioptrías respectivamente. El proceso de captura consiste en la obtención de la imagen producida por cada píxel tras su visualización a través de las microlentes. Cabe destacar que, para un sistema de 256×256 como el de estudio, sería necesario tomar 65.536 imágenes. Con el objetivo de reducir el tiempo de captura, esta se ha realizado en subconjuntos de 4×4 píxeles, garantizando que cada una de las respuestas observadas sea distinguible de sus vecindades y en un solo canal de color. Posteriormente, se llevó a cabo una segmentación para cada respuesta. Para reducir el espacio de almacenamiento, solo se almacenó la parte correspondiente de cada respuesta en una imagen, dejando el resto a 0, y recortando al área útil. Esto resultó en una resolución del sistema de adquisición de 977×992 . Asimismo, se obtuvo un *bias* para deducirlo de las imágenes adquiridas. En la Figura 6.5 se muestran la imagen percibida por la cámara del área de evaluación para cada uno de los casos. Como se puede observar, debido a que la rejilla de microlentes no se encuentra en un plano conjugado de la pantalla, un mismo píxel da lugar por respuesta a varios puntos en el sensor.

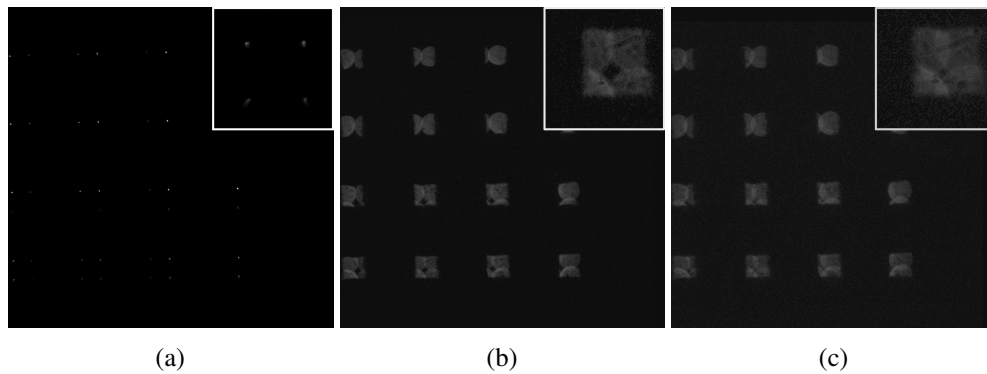


Figura 6.5 Respuesta al impulso obtenida para una rejilla de 4x4 píxeles para diferentes desenfocos. De izquierda a derecha, 0 (6.5a), 2 (6.5b), y 4 (6.5c) dioptrías respectivamente. En la esquina superior derecha se muestra una respuesta. Obsérvese que en cada uno de los casos está presente la apertura de las microlentes.

Un problema inherente a la técnica radica en el ajuste de las exposiciones para cada captura en relación a su PSF. Por ejemplo, la exposición de captura para el caso emélope no puede ser la misma que en el caso de 2D; si se mantuviera un tiempo de exposición suficientemente elevado como para percibir correctamente el caso desenfocado, podría darse la circunstancia de que las IR saturasen en intensidad al ser muestreadas por el sensor cuando no hay desenfoco. De la misma manera, las respuestas con desenfoco podrían no ser correctamente capturadas si el tiempo de exposición fuera demasiado corto.

6.4.2. Corrección de defectos visuales

Partiendo de las respuestas almacenadas, se hizo uso de una solución implementada sobre *MatlabR2016* con el fin de minimizar la Ecuación 6.6. Para ello, de manera similar al Capítulo 5, se empleó un algoritmo de minimización *trust-region reflective*, como parte de la función LSQNLIN. El resultado se limitó al rango dinámico de la pantalla. Dicha implementación se llevó a cabo en una computadora con un Intel i7-7700 3.60GHz CPU y 64GB de memoria RAM.

En la Figura 6.6 se muestran los resultados obtenidos para tres desenfocos diferentes. La primera columna de la Figura muestra las imágenes de referencia empleadas. A su derecha se muestra el vector ρ tras ser reordenado para tres casos distintos (0, 2 y 4 dioptrías). Las dos últimas columnas se corresponden con los resultados obtenidos en simulación y en el montaje experimental. En la línea de lo anteriormente explicado, hay una discrepancia entre los contrastes obtenidos en la imagen final con el caso experimental. Salvando este inconveniente, se puede observar que comparativamente, los resultados en simulación y montaje son equivalentes. En relación a la columna de impulsos, cabe destacar que, para el caso sin desenfoco, la imagen obtenida se asemeja a las imágenes elementales de un sistema integral. Un detalle importante es la presencia de pequeños intersticios en la imagen. Estos tienen correspondencia con aquellos presentes en la rejilla de microlentes. Al aumentar el desenfoco,

sin embargo, ésta comienza a dispersar los puntos de intensidad, posiblemente compensando el tamaño de las respuestas.

De manera comparativa, se muestra en la Figura 6.7 el caso de 4D. De izquierda a derecha, la imagen de referencia, el sistema capturado sin corrección con MLA, imagen de referencia desenfocada con una PSF equivalente a 4D, y la imagen resultante de la corrección en simulación. Tras comparar (6.7c) y (6.7d), se puede determinar que el sistema propuesto ofrece una corrección parcial de la imagen presentada.

En la Figura 6.8 se muestra una comparativa similar de los resultados obtenidos tanto en simulación como experimentalmente al realizar la corrección sobre una carta de USAF-1951. Junto a cada caso se ha añadido una ampliación del grupo -1. En todos los casos se demuestra que la resolución alcanzable por el sistema permite definir correctamente dicho grupo.

6.4.3. Mejora de la resolución

El problema de la deconvolución sobre la imagen ya se demostró en el Capítulo 5 como una metodología potencial a la hora de superar los límites impuestos por el ancho de banda del sistema de visualización. Dado que en este caso la propuesta *captura-representación* parte de una base similar, se ha evaluado la capacidad subyacente para la mejora de la resolución como alternativa dentro del método. No obstante, debido a la complejidad del sistema –MLA y desenfoco externo– se estudiará el caso más simple, esto es, cuando no se introduce un desenfoco al sistema.

En la Figura 6.9 se muestra nuevamente una carta USAF-1951 con la finalidad de definir el incremento en resolución del sistema. Se presenta además dicha imagen rotada con un ángulo de 45grados. De izquierda a derecha se muestran la imagen de referencia a la resolución del sistema de adquisición, a resolución base, y capturada en el sistema experimental. En ambos ejemplos, el grupo central, específicamente las columnas de 0 y 1, es distinguible únicamente en la imagen tratada, mientras que no lo es en aquella a la resolución base. Cabe destacar que en el caso con rotación presente, esta mejora es algo más discreta.

Este comportamiento se repite en trabajos anteriores (Sajadi et al., 2013), donde ciertas configuraciones ópticas y las interdependencias entre las respuestas pueden favorecer un conjunto particular de frecuencias. Precisamente se observaba en el Capítulo 5 como para un texto dado, existía una discrepancia entre la calidad obtenida en función de posición y ángulos a representar. No obstante, en cualquiera de los casos se percibe un incremento en las frecuencias representadas.

6.4.4. Sistema multifocal

Se ha evaluado el sistema en la representación de información multifoco. En la Figura 6.10 una matriz de 3×3 imágenes. Cada imagen está compuesta por tres franjas correspondientes a una carta de Robson-Pelli Campbell y Robson (1968) dispuestas horizontalmente una sobre otra. Esto se ha hecho con el fin de descartar posibles discrepancias a lo largo del campo. La

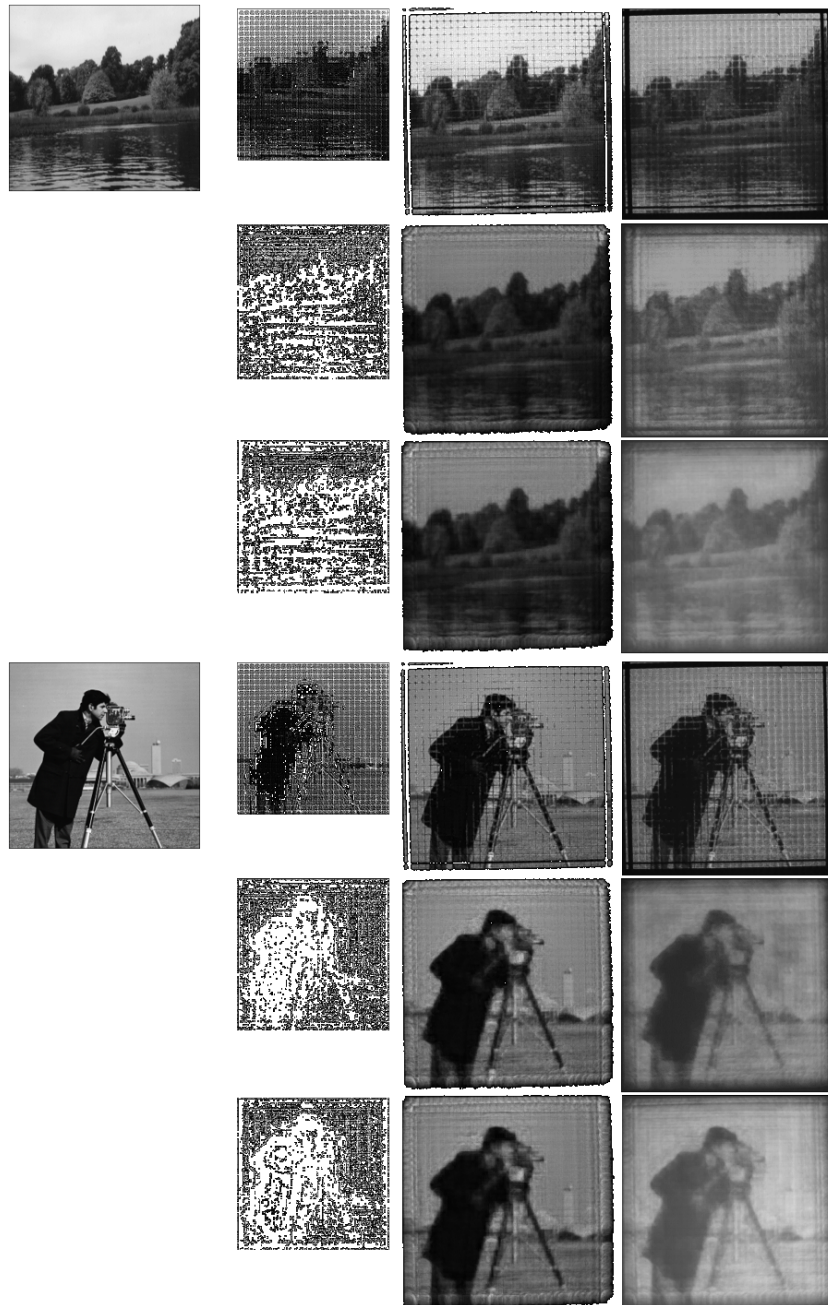


Figura 6.6 Imágenes obtenidas por el sistema de adquisición tras la minimización. Se han usado (*Cameraman* y *Autumn*), pertenecientes al repositorio de Matlab. En cada grupo, se representa el resultado para 0, 2 y 4 dioptrías. La primera columna muestra la imagen de referencia. La segunda columna contiene el resultado de la combinación de IRs obtenida a partir de la Ecuación 6.6. Las dos últimas columnas contienen los resultados en simulación y experimentales respectivamente. Debido al muestreo con diferentes valores de exposición, ha sido necesario hacer un leve ajuste de contraste para facilitar la visualización.

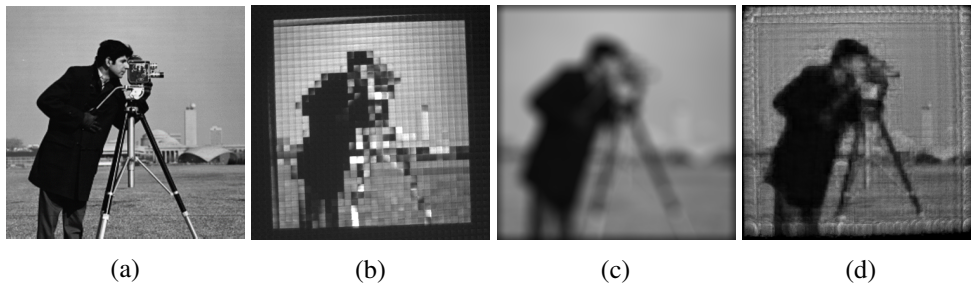


Figura 6.7 Comparación entre (6.7a) imagen de referencia, (6.7b) resolución máxima con el MLA, (6.7c) convolucionada con el círculo de desenfoque correspondiente a 4D, y (6.7d) corregida.

matriz está compuesta por todas las combinaciones de casos estudiados anteriormente. El eje vertical se corresponde con la imagen del impulso representada en la pantalla, mientras que el eje horizontal hace referencia a un observador teórico para cada desenfoque.

En consonancia con los resultados obtenidos previamente, la diagonal se corresponde con la imagen de estudio enfocada. Por encima de dicha diagonal se encuentran los casos en los que el observador está enfocando fuera del plano en el que se sitúa la imagen, correspondientemente ocasionando un cierto desenfoque sobre la imagen. Sin embargo, las imágenes correspondientes a un observador cuya potencia es inferior a la representada en la pantalla dan lugar a una representación incorrecta. Esto se debe a que, como se había concluido anteriormente, la solución del problema se dispersa de acuerdo con la potencia del sistema de adquisición (véase Figura 6.7), subsecuentemente ocasionando que el desenfoque aplicado a dicha imagen no se ajuste a lo esperado.

No obstante, esta situación no entraña un problema a nivel práctico. Teóricamente, el valor de desenfoque, salvando la magnificación, da lugar a un círculo de confusión igual tanto si el observador enfoca una cierta distancia por delante o detrás del sistema. Por ello, es posible hacer uso del impulso obtenido por ejemplo para el caso de 2D independientemente de si esta potencia es positiva o negativa, ya que su percepción debería ser idéntica. A pesar de ello, esto introduce un nuevo inconveniente, y es que el observador podría percibir el contenido de manera bifocal, esto es, con dos profundidades diferentes enfocadas al mismo tiempo.

Una manera de sortear este inconveniente es forzando el enfoque más cercano como el valor mínimo de distancia al que un objeto puede hacer imagen correctamente en la retina de dicho observador. Esto puede conseguirse trasladando el plano del sistema hacia dicha posición, bien óptica o mecánicamente. Sin embargo, adaptar esto para un observador particular implica repetir el proceso de captura de la respuesta al impulso, con lo que, en un caso práctico, lo ideal sería establecer esta posición de manera general, para un límite de potencia positiva; esto es, un ojo con miopía.

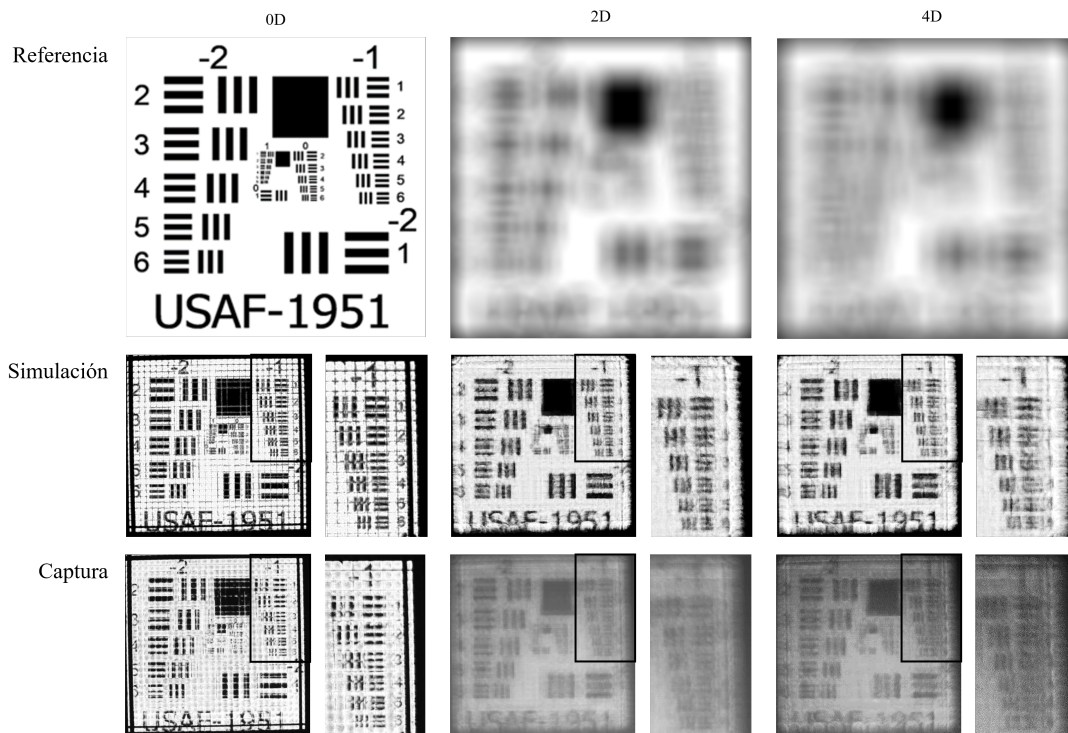


Figura 6.8 Carta USAF-1951 normalizada indicando el límite en resolución del sistema para cada caso. Los resultados incluyen la simulación y la captura del prototipo experimental en la primera y segunda columna respectivamente. Junto a cada carta se ha añadido una ampliación del grupo -1

6.5. Sumario y conclusiones

El modo de uso de un sistema de visualización tradicional parte de la idealización de la función de transferencia del píxel. Si bien los diseñadores de estos dispositivos compensan las limitaciones –especialmente optoelectrónicas– derivadas de los procesos de manufactura, no hay una especial atención a la cuantificación y medida de la información característica de la estructura que los componen.

Por ello, en este Capítulo, que sirve como cierre del segundo bloque de la memoria, se ha estudiado cómo puede afectar la medida de dispersión del píxel en un sistema de imagen integral, siendo ésta empleada posteriormente para la correcta representación de una imagen bidimensional. Con este fin, se ha desarrollado una metodología *captura-representación* que engloba la adquisición de la función de transferencia de cada píxel y su posterior uso a efectos de pre-compensar el contenido representado.

En líneas similares al Capítulo 5, se ha observado como la multiplicidad de la respuesta evaluada por cada píxel y la interacción con sus vecinos da lugar a un incremento efectivo en el ancho de banda alcanzable por el sistema. No obstante, debido a la variabilidad de los artefactos introducidos en la imagen por el MLA, no ha sido posible evaluar completamente este incremento.

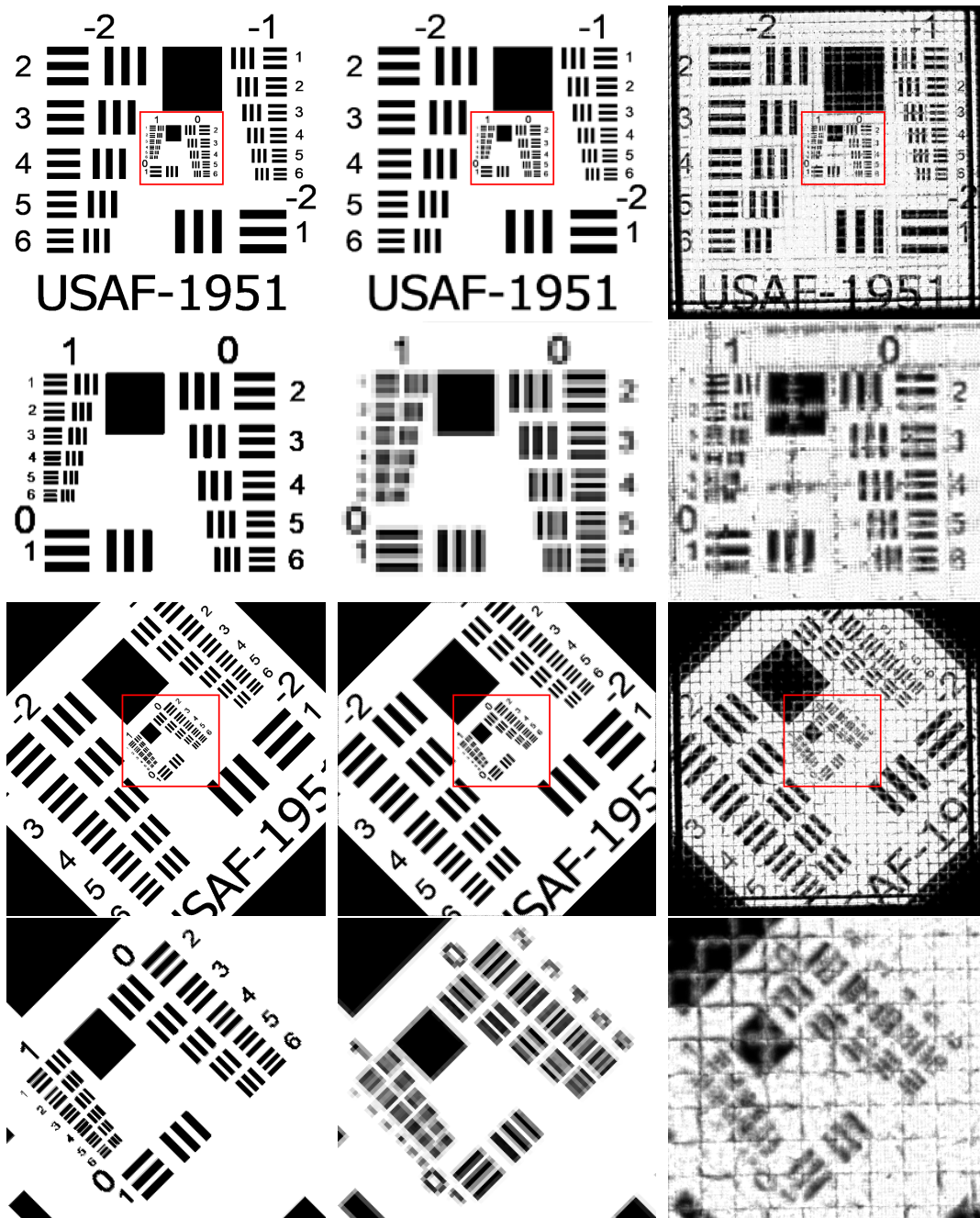


Figura 6.9 Resultados obtenidos en simulación para la resolución base. Dos imágenes de muestra fueron empleadas, *Cookie Monster* (primera fila) y *Lenna* (segunda fila). La primera columna se corresponde con la imagen en resolución base. La segunda columna es el resultado obtenido después de mostrar la imagen precompensada. En la última columna se muestra una ampliación de algunas zonas clave de las imágenes.

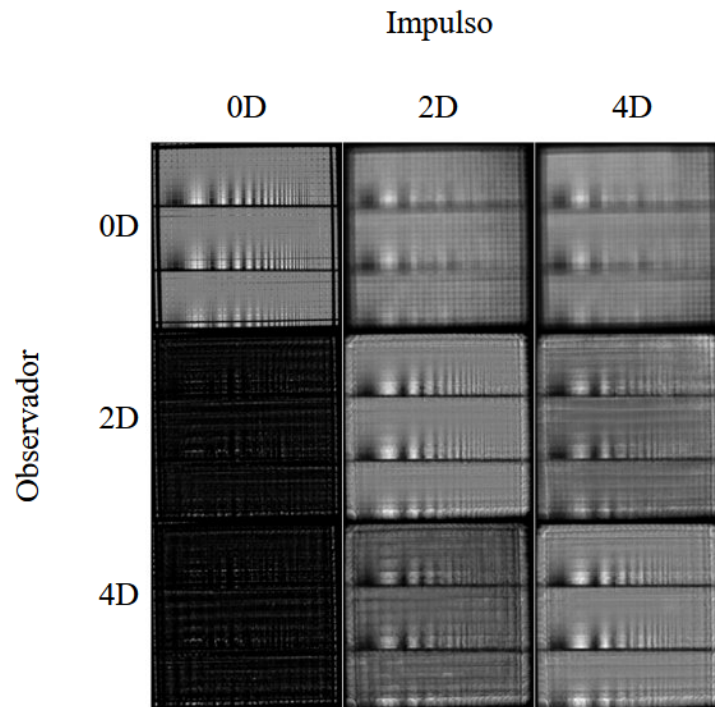


Figura 6.10 Matriz de correlación para diferentes valores de desenfoque aplicados sobre una carta de *Robson-Pelli*. Cada imagen se corresponde con la evaluación del sistema propuesto en la Ecuación 6.6 sustituyendo R y ρ por las diferentes combinaciones posibles de potencias.

Una aplicación directa de esta técnica radica en la corrección de defectos visuales por medio de la introducción de estos en la fase de adquisición. Más concretamente, se ha evaluado cualitativamente el grado de corrección obtenido para acomodaciones posteriores al sistema. Esto ha resultado en una mejora considerable en términos de frecuencias espaciales observadas por un sujeto de hasta 4D dioptrías.

No obstante, la implementación práctica de esta técnica manifiesta una considerable reducción en el contraste, similar a lo presentado por Huang et al. (2011). Se teoriza que este efecto, que también puede ser observado en el pre-compensado realizado en el Capítulo 5, tiene relación con el número de interacciones entre las respuestas observables de cada píxel, o de forma general, en el grado de interdependencia generado. Alternativamente, en trabajos futuros podría añadirse un conjunto de reglas adicionales sobre el operador de pre-compensado que tomara este problema en consideración.

De forma paralela, y repitiendo la experiencia del Capítulo previo, se ha analizado el sistema desde un punto de vista de la mejora de resolución espacial. De este análisis se extraen múltiples conclusiones. Por un lado, se demuestra que es posible recuperar una imagen con una resolución espacial superior a la del área de estudio del sistema base, esto es la pantalla sin MLA. Para los casos con defecto visual, se observa también una recuperación casi total. Esto resulta interesante, ya que hasta ahora los sistemas de imagen integral habían quedado relegados a la representación de información tridimensional en detrimento de su resolución

espacial. Esta técnica permite alternar de forma efectiva entre un teórico modo 2D y 3D. Comparativamente a otros trabajos (Hong et al., 2012; Kim et al., 2007; Zhang et al., 2020), la técnica propuesta está únicamente limitada a un proceso computacional una vez se ha realizado la adquisición de las IR. Es importante destacar que los resultados obtenidos por Pamplona et al. (2012) son similares a los presentados en este Capítulo, si bien no se evalúa la técnica con ese fin. En un trabajo futuro puede ser interesante evaluar ambos en términos de ganancia de resolución.

Por otro lado, se ha demostrado que es posible simular diferentes desenfoques para una misma escena. Este análisis se ha limitado a una simulación y se plantea como una aplicación alternativa para los sistemas de imagen integral. Otros trabajos ya han evaluado antes metodologías que potencialmente pueden resolver el conflicto vergencia-acomodación (Akşit et al., 2015; Chang et al., 2018; Hua y Javidi, 2014; Lanman y Luebke, 2013; Maimone et al., 2013). Sin embargo, la propuesta de este sistema radica en la generación de pistas focales a partir de un solo plano de representación, es decir, el sistema de imagen integral. Con ello, se teoriza que podría ser posible su implementación en dispositivos de realidad virtual con el fin de computacionalmente solucionar este inconveniente.

Capítulo 7

Conclusiones y trabajo futuro

7.1. Sumario

Partiendo de la técnica de representación de imagen tridimensional por sistema multicapa se han realizado los siguientes avances:

- Se ha analizado la influencia de la solución inicial sobre el algoritmo WNMF para el sistema multicapa en la representación de imagen tridimensional. Se ha concluido que es posible mejorar la calidad en términos de señal ruido si se pondera el valor inicial frente a la profundidad de la escena para mitigar artefactos. Asimismo se ha presentado una implementación alternativa de las reglas multiplicativas del algoritmo WNMF consistente en la reordenación de sus términos para hacer un uso eficiente de la memoria en tramos intermedios del algoritmo. Por otro lado, se ha corroborado que dicho reordenamiento no afecta en forma a los resultados obtenidos. Finalmente se ha diseñado un soporte experimental multicapa compuesto por tres pantallas LCD con el objetivo de mitigar los defectos de alineamiento. Además, se ha presentado una metodología básica para, por medio de los elementos de calibración introducidos, corregir dichos defectos.
- Se ha planteado el sistema multicapa de forma alternativa para solventar la cuestión de privacidad en sistemas de uso cotidiano. Más concretamente, se ha introducido la implementación de WNMF con el fin de evitar de manera dinámica que un observador no autorizado tenga acceso al contenido de dicho sistema, mientras su usuario puede seguir su actividad con normalidad. Esto se ha evaluado de forma objetiva para un caso teórico.
- Se ha implementado la técnica WNTF de en el proceso de la tomografía atmosférica. Se ha demostrado como dicha implementación mejora en torno a un 50 %, de acuerdo a las métricas propuestas, los resultados obtenidos por los algoritmos de uso más extendido en la actualidad para la reconstrucción de un perfil atmosférico basado en medidas reales. No obstante, los resultados tienen un valor cualitativo, pues es necesaria la introducción de otros agentes no evaluados y que pueden encontrarse en el proceso de la AO.

En el ámbito de sistemas bidimensionales se han realizado los siguientes avances:

- Estudio y análisis del proceso de pre-compensado de imagen para la mejora de la resolución lateral. Se ha analizado la influencia de dicho pre-compensado en dos topologías propuestas, obteniendo los puntos de operación óptimos. Alternativamente, se ha propuesto una implementación multimodal que toma como ventaja la multiplexación temporal, mejorando la propuesta previa. Se han implementado ambas técnicas sobre soportes de visualización actuales demostrando la viabilidad de su uso.
- Se ha propuesto un paradigma de representación de imagen basado en un proceso de *captura-representación* para la corrección de defectos visuales sobre el observador de un sistema. Se ha demostrado la viabilidad de dicho sistema no solo para la corrección de defectos visuales, sino como sistema multifocal, siendo esto una potencial solución al conflicto vergencia-acomodación en los sistemas de imagen tridimensional. La técnica de *captura-representación* ha demostrado mejorar la resolución base de un sistema de manera similar a la propuesta de pre-compensado sobre pantalla bidimensional, siendo este caso un sistema de imagen integral. Finalmente se ha demostrado que dicha técnica permite intercambiar de forma efectiva un sistema de imagen integral entre un teórico modo bidimensional y tridimensional, resolviendo la limitación en resolución lateral propia de estos sistemas.

7.2. Publicaciones

7.2.1. Publicaciones relevantes

1. David Carmona-Ballester, Juan M. Trujillo-Sevilla, Sergio Bonaque-González, Óscar Gómez-Cárdenas, y José M. Rodríguez-Ramos. Weighted nonnegative tensor factorization for atmospheric tomography reconstruction. *Astronomy and Astrophysics*, 614 (A41), 2018a. doi: <https://doi.org/10.1051/0004-6361/201832597>
2. David Carmona-Ballester, Juan M Trujillo-Sevilla, Lara Díaz-García, Daniel Walo, Ángela Hernández-Delgado, Juan J Fernández-Valdivia, Óscar Casanova-González, Óscar Gómez-Cárdenas, y José M Rodríguez-Ramos. Privacy-enabled displays. volume 10666, page 1066609, 2018b. URL <https://doi.org/10.1117/12.2306029>
3. David Carmona-Ballester, Viana L Guadalupe-Suárez, Juan M Trujillo-Sevilla, Sergio Bonaque-González, Ricardo Oliva-García, Jose G Marichal-Hernández, Miguel J Sicilia-Cabrera, Lara Díaz-García, y Jose M Rodríguez-Ramos. Calibration method applied to a tunable tensor display system. In *Proc.SPIE*, volume 10997, may 2019a. URL <https://doi.org/10.1117/12.2518865>
4. David Carmona-Ballester, Juan M. TrujilloSevilla, Sabato Ceruso, y Oscar Casanova-Gonzalez. Shoulder-surfing prevention on multilayered light-field displays. In *ESODM-2019*, 2019c

7.2.2. Patentes

1. David Carmona-Ballester, Juan M. Trujillo-Sevilla, Sergio Bonaque-González, y José M. Rodríguez-Ramos. Method and optical system for characterizing and improve displays operation, 2019b

7.2.3. Otras publicaciones

1. Óscar; Gómez-Cárdenas, José Gil; Marichal-Hernández, Juan M.; Trujillo-Sevilla, David; Carmona-Ballester, y José M.; Rodríguez-Ramos. Focus measurement in 3D focal stack using direct and inverse discrete radon transform. In *Proceedings Volume 10219 , Three-Dimensional Imaging, Visualization, and Display 2017;*, Anaheim, California, United States, 2017. doi: 10.1117/12.2261921
2. Sergio Bonaque-González, Óscar Casanova-González, Juan M. Trujillo-Sevilla, Juan J. Fernández-Valdivia, Ángela Hernández-Delgado, y David Carmona-Ballester. Optical and physics differential characteristics of optical plastics. *Optica Pura y Aplicada*, 51 (2), 2018. ISSN 21718814. doi: 10.7149/OPA.51.2.50010
3. Ricardo Oliva-García, Óscar Gómez-Cárdenas, David Carmona-Ballester, José G Marichal-Hernández, y José M Rodríguez-Ramos. Simultaneous computation of discrete Radon transform quadrants for its efficient implementation on real time systems. In *Proc.SPIE*, volume 10996, 2019. URL <https://doi.org/10.1117/12.2518601>
4. Sabato Ceruso, Sergio Bonaque-González, Alicia Pareja-Ríos, David Carmona-Ballester, y Juan Trujillo-Sevilla. Reconstructing wavefront phase from measurements of its slope, an adaptive neural network based approach. *Optics and Lasers in Engineering*, 126: 105906, 2020. ISSN 0143-8166. doi: <https://doi.org/10.1016/j.optlaseng.2019.105906>. URL <http://www.sciencedirect.com/science/article/pii/S0143816619313697>
5. Sergio Bonaque-González, José M. Rodríguez-Ramos, Juan M. Trujillo-Sevilla, Oscar Casanova-González, David Carmona-Ballester, y Miguel J. Sicilia-Cabrera. Extremely high resolution ocular aberrometry up to 2.4 million points. In *ARVO Annual Meeting - Investigative Ophthalmology & Visual Science*, Vancouver, 2019

Bibliografía

- Edward H. Adelson y John Y. A. Wang. Single Lens Stereo with a Plenoptic Camera. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):99–106, 1992. ISSN 01628828. doi: 10.1109/34.121783.
- Kaan Akşit, Jan Kautz, y David Luebke. Slim near-eye display using pinhole aperture arrays. *Applied Optics*, 54(11):3422, 2015. ISSN 0003-6935. doi: 10.1364/AO.54.003422. URL <http://www.opticsinfobase.org/viewmedia.cfm?uri=ao-54-11-3422>. URL <http://www.opticsinfobase.org/viewmedia.cfm?uri=ao-54-11-3422>.
- K R Alexander, W Xie, y D J Derlacki. Spatial-frequency characteristics of letter identification. *Journal of the Optical Society of America. A, Optics, image science, and vision*, 11(9): 2375–2382, 1994. ISSN 1084-7529. doi: 10.1364/JOSAA.11.002375. URL <http://www.ncbi.nlm.nih.gov/pubmed/7931762>.
- Mohammed Eunus Ali, Anika Anwar, Ishrat Ahmed, Tanzima Hashem, Lars Kulik, y Egemen Tanin. Protecting mobile users from visual privacy attacks. In *UbiComp 2014 - Adjunct Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 1–4, 2014. ISBN 9781450330473. doi: 10.1145/2638728.2638788.
- Daniel G. Aliaga, Yu Hong Yeung, Alvin Law, Behzad Sajadi, y Aditi Majumder. Fast high-resolution appearance editing using superimposed projections. *ACM Transactions on Graphics*, 31(2):1–13, 2012. ISSN 07300301. doi: 10.1145/2159516.2159518.
- Jukka Antikainen, Jiří Havel, Radovan Jošth, Adam Herout, Pavel Zemčík, y Markku Hauta-Kasari. Nonnegative tensor factorization accelerated using GPGPU. *IEEE Transactions on Parallel and Distributed Systems*, 2011. ISSN 10459219. doi: 10.1109/TPDS.2010.194.
- Günter Auzinger. On choosing layer profiles in atmospheric tomography. In *Journal of Physics: Conference Series*, volume 595, 2015. doi: 10.1088/1742-6596/595/1/012001.
- Günter Auzinger, Miska Le Louarn, Andreas Obereder, y Daniela Saxenhuber. Effects of reconstruction layer profiles on atmospheric tomography in E-ELT AO systems. In *Adaptive Optics for Extremely Large Telescopes 4 – Conference Proceedings*, 2015. doi: 10.20353/K3T4CP1131383.
- A. Bakushinsky y A. Goncharsky. *Ill-Posed Problems: Theory and Applications*. 1994. doi: 10.1007/978-94-011-1026-6.
- E. C. Beckmann. CT scanning the early days. In *British Journal of Radiology*, volume 79, pages 5–8, 2006. ISBN 0007-1285 (Print)\r0007-1285 (Linking). doi: 10.1259/bjr/29444122.
- Gareth Bell, Robert Craig, y Tony Simmiss. Moiré interference in multilayered displays. *Journal of the Society for Information Display*, 15(11):883, 2007. ISSN 10710922. doi: 10.1889/1.2812988.
- Floraine Berthouzoz y Raanan Fattal. Resolution enhancement by vibrating displays. *ACM Trans. Graph.*, 31:15:1–15:14, 2012.

- Blender Foundation. *Blender - a 3D modelling and rendering package*. Blender Foundation, Stichting Blender Foundation, Amsterdam, 2018. URL <http://www.blender.org>.
- Vincent Blondel, Ngoc-Diep Ho, y Paul Van Dooren. Weighted Nonnegative Matrix Factorization and Face Feature Extraction. In *Image and Vision Computing - IVC*, 2007.
- Sergio Bonaque-González, Óscar Casanova-González, Juan M. Trujillo-Sevilla, Juan J. Fernández-Valdivia, Ángela Hernández-Delgado, y David Carmona-Ballester. Optical and physics differential characteristics of optical plastics. *Optica Pura y Aplicada*, 51(2), 2018. ISSN 21718814. doi: 10.7149/OPA.51.2.50010.
- Sergio Bonaque-González, José M. Rodríguez-Ramos, Juan M. Trujillo-Sevilla, Oscar Casanova-González, David Carmona-Ballester, y Miguel J. Sicilia-Cabrera. Extremely high resolution ocular aberrometry up to 2.4 million points. In *ARVO Annual Meeting - Investigative Ophthalmology & Visual Science*, Vancouver, 2019.
- Hans Brettel, Lei Shi, y Hans Strasburger. Temporal image fusion in human vision. *Vision Research*, 2006. ISSN 00426989. doi: 10.1016/j.visres.2005.07.005.
- F. W. Campbell y J. G. Robson. Application of fourier analysis to the visibility of gratings. *The Journal of Physiology*, 1968. ISSN 00223751. doi: 10.1113/jphysiol.1968.sp008574.
- D. Carmona-Ballester, J. M. Trujillo-Sevilla, Sergio Bonaque-González, A. Hernández-Delgado, y J. M. Rodríguez-Ramos. Analyzing the impact of the initialization for the nonnegative tensor factorization in tensor displays (in press). *Optical Engineering*, 57(6), 2017. doi: 10.1117/1.OE.57.6.061603.
- David Carmona-Ballester, Juan M. Trujillo-Sevilla, Sergio Bonaque-González, Óscar Gómez-Cárdenas, y José M. Rodríguez-Ramos. Weighted nonnegative tensor factorization for atmospheric tomography reconstruction. *Astronomy and Astrophysics*, 614(A41), 2018a. doi: <https://doi.org/10.1051/0004-6361/201832597>.
- David Carmona-Ballester, Juan M Trujillo-Sevilla, Lara Díaz-García, Daniel Walo, Ángela Hernández-Delgado, Juan J Fernández-Valdivia, Óscar Casanova-González, Óscar Gómez-Cárdenas, y José M Rodríguez-Ramos. Privacy-enabled displays. volume 10666, page 1066609, 2018b. URL <https://doi.org/10.1117/12.2306029>.
- David Carmona-Ballester, Viana L Guadalupe-Suárez, Juan M Trujillo-Sevilla, Sergio Bonaque-González, Ricardo Oliva-García, Jose G Marichal-Hernández, Miguel J Sicilia-Cabrera, Lara Díaz-García, y Jose M Rodríguez-Ramos. Calibration method applied to a tunable tensor display system. In *Proc.SPIE*, volume 10997, may 2019a. URL <https://doi.org/10.1117/12.2518865>.
- David Carmona-Ballester, Juan M. Trujillo-Sevilla, Sergio Bonaque-González, y José M. Rodríguez-Ramos. Method and optical system for characterizing and improve displays operation, 2019b.
- David Carmona-Ballester, Juan M. TrujilloSevilla, Sabato Ceruso, y Oscar Casanova-Gonzalez. Shoulder-surfing prevention on multilayered light-field displays. In *ESODM-2019*, 2019c.
- Luca Cerina, Pierandrea Cancian, Giuseppe Franco, y Marco Domenico Santambrogio. A hardware acceleration for surface EMG non-negative matrix factorization. In *Proceedings - 2017 IEEE 31st International Parallel and Distributed Processing Symposium Workshops, IPDPSW 2017*, 2017. ISBN 9781538634080. doi: 10.1109/IPDPSW.2017.66.
- Sabato Ceruso, Sergio Bonaque-González, Alicia Pareja-Ríos, David Carmona-Ballester, y Juan Trujillo-Sevilla. Reconstructing wavefront phase from measurements of its slope, an

- adaptive neural network based approach. *Optics and Lasers in Engineering*, 126:105906, 2020. ISSN 0143-8166. doi: <https://doi.org/10.1016/j.optlaseng.2019.105906>. URL <http://www.sciencedirect.com/science/article/pii/S0143816619313697>.
- Jin-Xiang Chai, Xin Tong, Shing-Chow Chan, y Heung-Yeung Shum. Plenoptic sampling. *Proceedings of the 27th annual conference on Computer graphics and interactive techniques - SIGGRAPH '00*, pages 307–318, 2000. ISSN 0097-8930. doi: 10.1145/344779.344932. URL <http://portal.acm.org/citation.cfm?doid=344779.344932>.
- Jen Hao Rick Chang, B. V.K. Vijaya Kumar, y Aswin C. Sankaranarayanan. Towards multifocal displays with dense focal stacks. In *SIGGRAPH Asia 2018 Technical Papers, SIGGRAPH Asia 2018*, 2018. ISBN 9781450360081. doi: 10.1145/3272127.3275015.
- Renjie Chen, Andrew Maimone, Henry Fuchs, Ramesh Raskar, y Gordon Wetzstein. Wide field of view compressive light field display using a multilayer architecture and tracked viewers. *Journal of the Society for Information Display*, 22(10):525–534, 2014. ISSN 19383657. doi: 10.1002/jsid.285.
- Jen Tzung Chien y Yi Ting Bao. Tensor-factorized neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, 2018. ISSN 21622388. doi: 10.1109/TNNLS.2017.2690379.
- Stephen D Christman, Frederick L Kitterle, y Christopher L Niebauer. Hemispheric asymmetries in the identification of band-pass filtered letters. *Psychonomic Bulletin and Review*, 4(2):277–284, 1997. ISSN 10699384. doi: 10.3758/BF03209406.
- Andrzej Cichocki, Rafal Zdunek, Anh Huy Phan, y Shun Ichi Amari. *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-Way Data Analysis and Blind Source Separation*. John Wiley & Sons, Hoboken, NJ, USA, first edition, 2009. ISBN 9780470746660. doi: 10.1002/9780470747278.
- T Coleman y Y Li. A Reflective Newton Method for Minimizing a Quadratic Function Subject to Bounds on Some of the Variables. *SIAM Journal on Optimization*, 6(4):1040–1058, nov 1996. ISSN 1052-6234. doi: 10.1137/S1052623494240456. URL <https://doi.org/10.1137/S1052623494240456>.
- A. M. Cormack. Representation of a function by its line integrals, with some radiological applications. II. *Journal of Applied Physics*, 35(10):2908–2913, 1964. ISSN 00218979. doi: 10.1063/1.1713127.
- A Cortes, B Neichel, A Guesalaga, J Osborn, F Rigaut, y D Guzman. New results on a Cn2 profiler for GeMS. *Proceedings of SPIE - The International Society for Optical Engineering*, 8447(June):1–12, 2012. ISSN 0277786X. doi: 10.1117/12.925242.
- Niranjan Damera-Venkata y Nelson L. Chang. Realizing super-resolution with superimposed projection. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, (June), 2007. ISSN 10636919. doi: 10.1109/CVPR.2007.383463.
- Niranjan Damera-Venkata y Nelson L. Chang. Display supersampling. *ACM Transactions on Graphics*, 28(1):1–19, 2009. ISSN 07300301. doi: 10.1145/1477926.1477935.
- Gregory C. Deangelis. Seeing in three dimensions: The neurophysiology of stereopsis, 2000. ISSN 13646613.
- James J. DiCarlo, Davide Zoccolan, y Nicole C. Rust. How does the brain solve visual object recognition?, 2012. ISSN 08966273.

- Piotr Didyk, Elmar Eisemann, Tobias Ritschel, Karol Myszkowski, y Hans-Peter Seidel. Apparent Display Resolution Enhancement for Moving Images. *ACM Transactions on Graphics (Proceedings SIGGRAPH 2010, Los Angeles)*, 29(4), 2010.
- A. Dorado, S. Hong, G. Saavedra, J. C. Barreiro, B. Javidi, y M. Martinez-Corral. Toward 3D integral-imaging broadcast with increased viewing angle and parallax. *Optics and Lasers in Engineering*, 2018. ISSN 01438166. doi: 10.1016/j.optlaseng.2018.03.017.
- Francisco J. Duarte. Multiple-return-pass beam divergence and the linewidth equation. *Applied Optics*, 2001. ISSN 0003-6935. doi: 10.1364/ao.40.003038.
- Malin Eiband, Mohamed Khamis, Emanuel von Zezschwitz, Heinrich Hussmann, y Florian Alt. Understanding Shoulder Surfing in the Wild. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*, pages 4254–4265, 2017. ISBN 9781450346559. doi: 10.1145/3025453.3025636. URL <http://dl.acm.org/citation.cfm?doid=3025453.3025636>.
- Brent L Ellerbroek. Efficient computation of minimum-variance wave-front reconstructors with sparse matrix techniques. *Journal of the Optical Society of America. A, Optics, image science, and vision*, 19(9):1803–16, 2002. ISSN 1084-7529. doi: Doi10.1364/Josaa.19.001803. URL <http://www.ncbi.nlm.nih.gov/pubmed/12216874>.
- Shaojing Fan, Tian Tsong Ng, Jonathan S. Herberg, Bryan L. Koenig, Cheston Y.C. Tan, y Rangding Wang. An automated estimator of image visual realism based on human cognition. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2014. ISBN 9781479951178. doi: 10.1109/CVPR.2014.535.
- A Glassner. Inside Moire Patterns. *IEEE Comput. Graph. Appl.*, 17(6):97–101, 1997. ISSN 0272-1716. doi: 10.1109/38.626975. URL <https://doi.org/10.1109/38.626975>.
- Óscar; Gómez-Cárdenas, José Gil; Marichal-Hernández, Juan M.; Trujillo-Sevilla, David; Carmona-Ballester, y José M.; Rodríguez-Ramos. Focus measurement in 3D focal stack using direct and inverse discrete radon transform. In *Proceedings Volume 10219, Three-Dimensional Imaging, Visualization, and Display 2017*; , Anaheim, California, United States, 2017. doi: 10.1117/12.2261921.
- Joseph W. Goodman. *Introduction to Fourier Optics, Third Edition*. 2005. ISBN 0974707724. doi: 10.1117/1.601121.
- Richard Gordon, Robert Bender, y Gabor T Herman. Algebraic Reconstruction Techniques (ART) for three-dimensional electron microscopy and X-ray photography. *Journal of Theoretical Biology*, 29(3):471–476, 1970. ISSN 10958541. doi: 10.1016/0022-5193(70)90109-8.
- Steven J. Gortler, Radek Grzeszczuk, Richard Szeliski, y Michael F. Cohen. The lumigraph. In *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 1996*, 1996. ISBN 0897917464. doi: 10.1145/237170.237200.
- Hironobu Gotoda. A multilayer liquid crystal display for autostereoscopic 3D viewing. *Proc. SPIE*, 7524(Stereoscopic Displays and Applications XXI):75240P 1—8, 2010. ISSN 0277786X. doi: 10.1117/12.840286. URL <http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=775969>.
- L. Grippo y M. Sciandrone. On the convergence of the block nonlinear Gauss-Seidel method under convex constraints. *Operations Research Letters*, 2000. ISSN 01676377. doi: 10.1016/S0167-6377(99)00074-7.

- Gabor T. Herman y Lorraine B. Meyer. Algebraic Reconstruction Techniques Can Be Made Computationally Efficient. *IEEE Transactions on Medical Imaging*, 12(3):600–609, 1993. ISSN 1558254X. doi: 10.1109/42.241889.
- Matthew Hirsch, Douglas Lanman, Gordon Wetzstein, y Ramesh Raskar. Construction and Calibration of Optically Efficient Lcd-based Multi-layer Light Field Displays.
- David M Hoffman, Ahna R Girshick, y Martin S Banks. Vergence – accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of Vision*, 8(2008):1–30, 2015. doi: 10.1167/8.3.33.Introduction. URL <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2879326/>.
- Jisoo Hong, Jonghyun Kim, y Byoungcho Lee. Two-dimensional/three-dimensional convertible integral imaging using dual depth configuration. *Applied Physics Express*, 5(1):10–12, 2012. ISSN 18820778. doi: 10.1143/APEX.5.012501.
- H. Hoshino, F. Okano, H. Isono, e I. Yuyama. Analysis of resolution limitation of integral photography. *Journal of the Optical Society of America A*, 1998. ISSN 1084-7529. doi: 10.1364/josaa.15.002059.
- G. N. Hounsfield. Computerized transverse axial scanning (tomography): I. Description of system. *British Journal of Radiology*, 46(552):1016–1022, 1973. ISSN 00071285. doi: 10.1259/0007-1285-46-552-1016.
- Hong Hua y Bahram Javidi. A 3D integral imaging optical see-through head-mounted display. 22(11):198–203, 2014. doi: 10.1364/OE.22.013484.
- Fc Huang y G Wetzstein. Eyeglasses-free display: towards correcting visual aberrations with computational light field displays. *ACM Transactions on Graphics (SIGGRAPH)*, pages 1–12, 2014. ISSN 15577333. doi: 10.1145/2601097.2601122. URL <http://dl.acm.org/citation.cfm?id=2601122>.
- Jian Huang, Armando Barreto, Miguel Alonso, y Malek Adjouadi. Vision correction for computer users based on image pre-compensation with changing pupil size. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, pages 4868–4871, 2011. ISSN 1557170X. doi: 10.1109/IEMBS.2011.6091206.
- Frederic E. Ives. Parallax Stereogram and Process of making same, 1903. URL <http://www.google.com/patents?hl=en&lr={&}vid=USPAT725567&id=ouBYAAAAEBAJ&oi=fnd&dq=Parallax+Stereogram+and+Process+of+making+same&printsec=abstract>.
- Herbert E. Ives. Parallax Panoramagrams Made with a Large Diameter Lens. *Journal of the Optical Society of America*, 1930. ISSN 0030-3941. doi: 10.1364/josa.20.000332.
- Prateek Jain y Sewoong Oh. Provable tensor factorization with missing data. In *Advances in Neural Information Processing Systems*, 2014.
- Yeong Min Ji y Jae Hyeung Park. Dual layered display that presents auto-stereoscopic 3D images to multiple viewers in arbitrary positions. *Journal of the Society for Information Display*, 2016. ISSN 19383657. doi: 10.1002/jsid.509.
- Jae-Hyun Jung, Soon-gi Park, Youngmin Kim, y Byoungcho Lee. Integral imaging using a color filter pinhole array on a display panel. *Optics express*, 20(17):18744–56, 2012. ISSN 1094-4087. doi: 10.1364/OE.20.018744. URL <http://www.ncbi.nlm.nih.gov/pubmed/23038515>.

- S. Kaczmarz. Angenäherte Auflösung von Systemen linearer Gleichungen. *Bulletin International de l'Académie Polonaise des Sciences et des Lettres*, 35:355–357, 1937.
- Avinash C. Kak, Author, Malcolm Slaney, Ge Wang, y Reviewer. Principles of Computerized Tomographic Imaging. *Medical Physics*, 29:107–107, 2002. ISSN 00942405. doi: 10.1118/1.1455742. URL <http://scitation.aip.org/getabs/servlet/GetabsServlet?prog=normal{id=MPHYA6000029000001000107000003{idtype=cvips{&}gifs=yes{&}5Cnhttp://link.aip.org/link/?MPH/29/107>.
- Mohamed Khamis, Malin Eiband, Martin Zürn, y Heinrich Hussmann. EyeSpot: Leveraging Gaze to Protect Private Text Content on Mobile Devices from Shoulder Surfing, 2018.
- Yunhee Kim, Joohwan Kim, Jin-Mo Kang, Jae-Hyun Jung, Heejin Choi, y Byoungcho Lee. Point light source integral imaging with improved resolution and viewing angle by the use of electrically movable pinhole array. *Optics express*, 15(26):18253–18267, 2007. ISSN 1094-4087. doi: 10.1364/OE.15.018253.
- A.N. Kolmogorov. The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers. *Doklady Akademii Nauk Sssr*, 30(1890):301–305, 1941. ISSN 13645021. doi: 10.1098/rspa.1991.0075.
- R. J. Krauzlis y S. G. Lisberger. Temporal properties of visual motion signals for the initiation of smooth pursuit eye movements in monkeys. *Journal of Neurophysiology*, 72(1):150–162, 1994. ISSN 00223077. doi: 10.1152/jn.1994.72.1.150.
- T. Kreis. Applications of digital holography: From microscopy to 3D-television. *Journal of the European Optical Society*, 2012. ISSN 19902573. doi: 10.2971/jeos.2012.12006.
- Mi Young Kwon y Gordon E. Legge. Spatial-frequency requirements for reading revisited. *Vision Research*, 2012. ISSN 00426989. doi: 10.1016/j.visres.2012.03.025.
- MiYoung Kwon y Gordon E. Legge. Spatial-frequency cutoff requirements for pattern recognition in central and peripheral vision. *Vision Research*, 2011. ISSN 00426989. doi: 10.1016/j.visres.2011.06.020.
- MiYoung Kwon y Gordon E Legge. Higher-contrast requirements for recognizing low-pass-filtered letters. *Journal of vision*, 13(1):13, 2013. ISSN 1534-7362. doi: 10.1167/13.1.13. URL <http://www.ncbi.nlm.nih.gov/pubmed/23302216>.
- R. G. Lane, A. Glindemann, y J. C. Dainty. Simulation of a Kolmogorov phase screen. *Waves in Random Media*, 2(3):209–224, 1992. ISSN 09597174. doi: 10.1088/0959-7174/2/3/003.
- Amy Langville y Carl Meyer. ALS Algorithms for The Nonnegative Matrix Factorization in Text Mining, 2005.
- Amy N Langville, Carl D Meyer, y Russell Albright. Initializations for the Nonnegative Matrix Factorization. 2006. doi: 10.1.1.131.4302.
- Amy N. Langville, Carl D. Meyer, Russell Albright, James Cox, y David Duling. Algorithms, Initializations, and Convergence for the Nonnegative Matrix Factorization. In *Proc. 12th ACM SIGKDD Int. Conf. Knowl. Disc. & Data Mining*, number 919, 2014. URL <http://arxiv.org/abs/1407.7299>.
- D Lanman, G Wetzstein, M Hirsch, W Heidrich, y R Raskar. Polarization Fields: Dynamic Light Field Display using Multi-Layer LCDs. *ACM Trans. Graph.*, 30(6), 2011.
- Douglas Lanman y David Luebke. Near-eye light field displays. *ACM SIGGRAPH 2013 Emerging Technologies on - SIGGRAPH '13*, page 1, 2013. ISSN 07300301. doi: 10.1145/2503368.2503379. URL <http://dl.acm.org/citation.cfm?doid=2503368.2503379>.

- Douglas Lanman, Matthew Hirsch, Yunhee Kim, y Ramesh Raskar. Content-adaptive parallax barriers: optimizing dual-layer 3D displays using low-rank light field factorization. *ACM Transactions on Graphics*, 29(6):1, 2010. ISSN 07300301. doi: 10.1145/1866158.1866164. URL <http://portal.acm.org/citation.cfm?doid=1866158.1866164>{% }5Cn<http://dl.acm.org/citation.cfm?id=1866164>.
- Douglas Lanman, Felix Heide, Dikpal Reddy, Jan Kautz, Kari Pulli, y David Luebke. Cascaded displays: spatiotemporal superresolution using offset pixel layers. *ACM SIGGRAPH 2014 Emerging Technologies*, page 5, 2014. ISSN 07300301. doi: 10.1145/2601097.2601120.
- M Le Louarn, N Hubin, M Sarazin, y A Tokovinin. New challenges for adaptive optics: extremely large telescopes. *Monthly Notices of the Royal Astronomical Society*, 317(3): 535–544, 2000. doi: doi:10.1046/j.1365-8711.2000.03607.x. URL +<http://dx.doi.org/10.1046/j.1365-8711.2000.03607.x>.
- D D Lee y H S Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755):788–91, 1999. ISSN 0028-0836. doi: 10.1038/44565. URL <http://www.ncbi.nlm.nih.gov/pubmed/10548103>.
- Dd Lee y Hs Seung. Algorithms for non-negative matrix factorization. *Advances in neural information processing systems*, (1):556–562, 2001. ISSN 10987576. doi: 10.1109/IJCNN.2008.4634046. URL <http://papers.nips.cc/paper/1861-algorithms-for-non-negative-matrix-factorization>.
- Hyekyoung Lee, Yong Deok Kim, Andrzej Cichocki, y Seungjin Choi. Nonnegative tensor factorization for continuous EEG classification. In *International Journal of Neural Systems*, 2007. doi: 10.1142/S0129065707001159.
- Marc Levoy y Pat Hanrahan. Light field rendering. *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques - SIGGRAPH '96*, pages 31–42, 1996. ISSN 00978930. doi: 10.1145/237170.237199. URL <http://portal.acm.org/citation.cfm?doid=237170.237199>.
- G. Lippmann. Épreuves Réversibles Donnant la Sensation du Relief. *Journal de Physique Théorique et Appliquée*, 7(1):821–825, 1908. ISSN 0368-3893. doi: 10.1051/jphystap:019080070082100.
- J P Lüke, F Rosa, J G Marichal-Hernández, J C Sanluís, C Domínguez Conde, y J M Rodríguez-Ramos. Depth From Light Fields Analyzing 4D Local Structure. *Journal of Display Technology*, 11(11):900–907, 2015. ISSN 1551319X. doi: 10.1109/JDT.2014.2360992.
- K. J. MacKenzie, R. A. Dickson, y S. J. Watt. Vergence and accommodation to multiple-image-plane stereoscopic displays: 'Real world' responses with practical image-plane separations? In *Stereoscopic Displays and Applications XXII*, 2011. ISBN 9780819484000. doi: 10.1117/12.872503.
- Kevin J. MacKenzie, David M. Hoffman, y Simon J. Watt. Accommodation to multiple-focal-plane displays: Implications for improving stereoscopic displays and for accommodation control. *Journal of Vision*, 2010. ISSN 15347362. doi: 10.1167/10.8.22.
- Andrew Maimone, Gordon Wetzstein, Matthew Hirsch, Douglas Lanman, Ramesh Raskar, y Henry Fuchs. Focus 3D. *ACM Transactions on Graphics*, 32(5):1–13, 2013. ISSN 07300301. doi: 10.1145/2503144. URL <http://dl.acm.org/citation.cfm?doid=2516971.2503144>.
- André Maréchal. Étude des effets combinés de la diffraction et des aberrations géométriques sur l'image d'un point lumineux. *Éditions de la Revue d'optique théorique et instrumentale*, 1948.

- MATLAB. *version 9.1 (R2016b)*. The MathWorks Inc., Natick, Massachusetts, 2016.
- Alonso Miguel y Armando Barreto. An Image Processing Approach to Pre-compensation for Higher-Order Aberrations in the Eye F. *Systemics, Cybernetics and Informatics*, 2(3):1–4, 2003.
- Alonso Miguel, Armando Barreto, y Julie A Jacko. Inverse Point Spread Function Processing. (Lcd):421–427, 2005.
- Lee; Mun-Kyu, Jin Bok; Kim, y Matthew K.; Franklin. Enhancing the Security of Personal Identification Numbers with Three-Dimensional Displays. *Mobile Information Systems*, 2016(Article ID 8019830):9 pages, 2016. doi: <https://doi.org/10.1155/2016/8019830>.
- F Natterer. *The Mathematics of Computerized Tomography*. Society for Industrial and Applied Mathematics, 2001. doi: 10.1137/1.9780898719284. URL <http://epubs.siam.org/doi/abs/10.1137/1.9780898719284>.
- Ren Ng. Fourier slice photography. *ACM Transactions on Graphics*, 24(3):735, 2005. ISSN 07300301. doi: 10.1145/1073204.1073256.
- Maximilian Nickel, Kevin Murphy, Volker Tresp, y Evgeniy Gabrilovich. A review of relational machine learning for knowledge graphs, 2016. ISSN 15582256.
- Ricardo Oliva-García, Óscar Gómez-Cárdenes, David Carmona-Ballester, José G Marichal-Hernández, y José M Rodríguez-Ramos. Simultaneous computation of discrete Radon transform quadrants for its efficient implementation on real time systems. In *Proc.SPIE*, volume 10996, 2019. URL <https://doi.org/10.1117/12.2518601>.
- Pentti Paatero y Unto Tapper. Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics*, 5(2):111–126, 1994. ISSN 1099095X. doi: 10.1002/env.3170050203.
- Vitor F Pamplona, Manuel M Oliveira, Daniel G Aliaga, y Ramesh Raskar. Tailored displays to compensate for visual aberrations. *ACM Transactions on Graphics*, 31(4):1–12, 2012. ISSN 07300301. doi: 10.1145/2185520.2335432. URL <papers3://publication/uuid/C4623603-9203-4ECB-BC3A-025BBEE56C7A>.
- Charles A. Primmerman, Daniel V. Murphy, Daniel A. Page, Byron G. Zollars, y Herbert T. Barclay. Compensation of atmospheric optical distortion using a synthetic beacon. *Nature*, 1991. ISSN 00280836. doi: 10.1038/353141a0.
- R. Ramlau y M. Rosensteiner. An efficient solution to the atmospheric turbulence tomography problem using Kaczmarz iteration. *Inverse Problems*, 28(9), 2012. ISSN 02665611. doi: 10.1088/0266-5611/28/9/095004.
- R. Ramlau, D. Saxenhuber, y M. Yudytskiy. Iterative reconstruction methods in atmospheric tomography: FEWHA, Kaczmarz and Gradient-based algorithm. In *Proc. SPIE*, pages 91480Q–91480Q–15, 2014. ISBN 9780819496164. doi: 10.1117/12.2057379.
- Maximilian Riesenhuber y Tomaso Poggio. Hierarchical models of object recognition in cortex. *Nature Neuroscience*, 1999. ISSN 10976256. doi: 10.1038/14819.
- François Roddier. *Adaptive Optics in Astronomy*. 1999. ISBN 9780511525179. doi: 10.1017/CBO9780511525179.
- Behzad Sajadi, Meenakshisundaram Gopi, y Aditi Majumder. Edge-guided resolution enhancement in projectors via optical pixel sharing. *ACM Trans. Graph.*, 31:79:1–79:122, 2012.

- Behzad Sajadi, Duy-Quoc Lai, Alexander T Ihler, Meenakshisundaram Gopi, y Aditi Majumder. Image enhancement in projectors via optical pixel shift and overlay. *IEEE International Conference on Computational Photography (ICCP)*, pages 1–10, 2013.
- Helin T. Saxenhuber D, Auzinger G, Louarn ML. Comparison of methods for the reduction of reconstructed layers in atmospheric tomography. *Appl Opt.*, 56(10):2621–2629, 2017. doi: 10.1364/AO.56.002621.
- Semple J. and Kneebone G. *Algebraic Projective Geometry*. Oxford University Press, 1979.
- Thomas Serre, Lior Wolf, y Tomaso Poggio. Object recognition with features inspired by visual cortex. In *Proceedings - 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005*, 2005. ISBN 0769523722. doi: 10.1109/CVPR.2005.254.
- Thomas Serre, Lior Wolf, Stanley Bileschi, Maximilian Riesenhuber, y Tomaso Poggio. Robust object recognition with cortex-like mechanisms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007. ISSN 01628828. doi: 10.1109/TPAMI.2007.56.
- C Shen, H Liu, M Yang, Y Hung, y S Pei. Viewing-Distance Aware Super-Resolution for High-Definition Display. *IEEE Transactions on Image Processing*, 24(1):403–418, 2015. ISSN 1057-7149. doi: 10.1109/TIP.2014.2375639.
- Xin Shen, Manuel Martinez-Corral, y Bahram Javidi. Extended view angle integral imaging 3D display using smart pseudoscopic-to-orthoscopic conversion with head tracking. In *Optics InfoBase Conference Papers*, 2016. ISBN 9781943580156. doi: 10.1364/3D.2016.TW1A.3.
- Xin Shen, Manuel Martinez-Corral, y Bahram Javidi. An overview of head tracking integral imaging three-dimensional display using smart pseudoscopic-to-orthoscopic conversion. In *Three-Dimensional Imaging, Visualization, and Display 2017*, 2017. ISBN 9781510609396. doi: 10.1117/12.2262376.
- Allan Shickman. "Turning Pictures in Shakespeare's England. *The Art Bulletin*, 1977. ISSN 00043079. doi: 10.2307/3049597.
- Louise L. Sloan. New test charts for the measurement of visual acuity at far and near distances. *American Journal of Ophthalmology*, 1959. ISSN 00029394. doi: 10.1016/0002-9394(59)90626-9.
- K. Strehl. Aplanatische und fehlerhafte Abbildung im Fernrohr. *Zeitschrift für Instrumentenkunde*, (15 (Oct.)):362–370, 1895.
- Thomas Strohmer y Roman Vershynin. A Randomized Kaczmarz Algorithm with Exponential Convergence. *Journal of Fourier Analysis and Applications*, 15(2):262, 2009. ISSN 1531-5851. doi: 10.1007/s00041-008-9030-4. URL <https://doi.org/10.1007/s00041-008-9030-4>.
- Keita Takahashi, Yuto Kobayashi, y Toshiaki Fujii. From Focal Stack to Tensor Light-Field Display. *IEEE Transactions on Image Processing*, 2018. ISSN 10577149. doi: 10.1109/TIP.2018.2839263.
- M Tallon y R Foy. Adaptive Telescope with Laser Probe : Isoplanatism and Cone Effect. *Astronomy and Astrophysics*, 235(1-2):549–557, 1990.
- A. N. Tikhonov. Solution of incorrectly formulated problems and the regularization method. *Soviet Mathematics*, 1963. ISSN 10634584. doi: 10.1557/PROC-1201-H02-06.
- J M Trujillo-Sevilla, L. F. Rodríguez-Ramos, I Montilla, y J. M. Rodríguez-Ramos. High resolution imaging and wavefront aberration correction in plenoptic systems. *Optics Letters*, 39(17):5030, 2014. ISSN 0146-9592. doi: 10.1364/OL.39.005030.

- URL <http://ol.osa.org/abstract.cfm?URI=ol-39-17-5030>{%}5Cn<https://www.osapublishing.org/abstract.cfm?URI=ol-39-17-5030>.
- J. M. Trujillo-Sevilla, Vladimir Katkovnik, Bahram Javidi, y J. M. Rodríguez-Ramos. Restoring Integral Images from Focal Stacks Using Compressed Sensing Techniques. *Journal of Display Technology*, 2016. ISSN 1551319X. doi: 10.1109/JDT.2016.2522922.
- JM. Trujillo-Sevilla. *Nuevas técnicas de adquisición de light field y fase de frente de onda*. PhD thesis, Universidad de La Laguna, 2017.
- Robert Ulichney, Ali Ghajarnia, y Niranjan Damera-Venkata. Quantifying performance of overlapped displays. In *Image Quality and System Performance VII*, 2010. ISBN 9780819479228. doi: 10.1117/12.838555.
- Stephen A. Vavasis. On the Complexity of Nonnegative Matrix Factorization. *SIAM Journal on Optimization*, 20(3):1364–1377, 2010. doi: 10.1137/070709967. URL <https://doi.org/10.1137/070709967>.
- A. Frieze Vempala, R. Kannan, y S. Fast Monte-Carlo algorithms for finding low-rank approximations. In *Proceedings 39th Annual Symposium on Foundations of Computer Science (Cat. No.98CB36280)*, pages 370–378, 1998. doi: 10.1109/SFCS.1998.743487.
- Daniel Vogel y Ravin Balakrishnan. Interactive public ambient displays: Transitioning from implicit to explicit, public to personal, interaction with multiple users. In *UIST: Proceedings of the Annual ACM Symposium on User Interface Software and Technology*, 2004.
- Shizheng Wang, Zhenfeng Zhuang, Phil Surman, Junsong Yuan, Yuanjin Zheng, y Xiao Wei Sun. Two-layer optimized light field display using depth initialization. *2015 Visual Communications and Image Processing, VCIP 2015*, pages 1–4, 2015. doi: 10.1109/VCIP.2015.7457837.
- Yu-Xiong Wang y Yu-Jin Zhang. Nonnegative Matrix Factorization: A Comprehensive Review. *IEEE Transactions on Knowledge and Data Engineering*, 25(6):1336–1353, 2013. ISSN 1041-4347. doi: 10.1109/TKDE.2012.51. URL <http://dx.doi.org/10.1109/TKDE.2012.51>{%}5Cn<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6165290>.
- Max Welling y Markus Weber. Positive tensor factorization. *Pattern Recognition Letters*, 22(12):1255–1261, 2001. ISSN 01678655. doi: 10.1016/S0167-8655(01)00070-8.
- Stephen Welstead. *Fractal and Wavelet Image Compression Techniques*. 2009. doi: 10.1117/3.353798.
- G Wetzstein, D Lanman, M Hirsch, y R Raskar. Tensor Displays: Compressive Light Field Synthesis using Multilayer Displays with Directional Backlighting. *ACM Trans. Graph. (Proc. SIGGRAPH)*, 31(4):1–11, 2012a.
- Gordon Wetzstein, Douglas Lanman, Wolfgang Heidrich, y Ramesh Raskar. Layered 3D. *ACM Transactions on Graphics*, 30(4):1, 2011. ISSN 07300301. doi: 10.1145/2010324.1964990.
- Gordon Wetzstein, Douglas R. Lanman, Matthew Waggener Hirsch, y Ramesh Raskar. Tensor displays: compressive light field synthesis using multilayer displays with directional backlighting. *ACM Transactions on Graphics*, 31(4):1–11, 2012b. ISSN 07300301. doi: 10.1145/2185520.2185576.
- Gordon Wetzstein, Douglas R. Lanman, Matthew Waggener Hirsch, y Ramesh Raskar. Tensor displays: compressive light field synthesis using multilayer displays with directional backlighting. *ACM Transactions on Graphics*, 31(4):1–11, 2012c. ISSN 07300301. doi: 10.1145/2185520.2185576.

- Christian Winkler, Jan Gugenheimer, Alexander De Luca, Gabriel Haas, Philipp Speidel, David Dobbstein, y Enrico Rukzio. Glass unlock: Enhancing security of smartphone unlocking through leveraging a private near-eye display. In *Conference on Human Factors in Computing Systems - Proceedings*, 2015. ISBN 9781450331456. doi: 10.1145/2702123.2702316.
- Xiao Xiao, Bahram Javidi, Manuel Martinez-Corral, y Adrian Stern. Advances in three-dimensional integral imaging: sensing, display, and applications [Invited]. *Applied Optics*, 52(4):546, 2013. ISSN 0003-6935. doi: 10.1364/AO.52.000546. URL <https://www.osapublishing.org/abstract.cfm?URI=ao-52-4-546>.
- Jinwei Ye, Yu Ji, Mingyuan Zhou, Sing Bing Kang, y Jingyi Yu. Content Aware Image Pre-Compensation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8828(c):1–14, 2018. ISSN 01628828. doi: 10.1109/TPAMI.2018.2839115.
- Liu Yiqun. Separable and Three-Dimensional Optical Reversible Data Hiding with Integral Imaging Cryptosystem. In Leonard Barolli, Mingwu Zhang, y Xu An Wang, editors, *Advances in Internetworking, Data & Web Technologies*, pages 13–21, Cham, 2018. Springer International Publishing. ISBN 978-3-319-59463-7.
- David Young. Homography Matrix estimation, 2008. URL <https://es.mathworks.com/matlabcentral/profile/authors/415894-david-young>.
- Han Le Zhang, Huan Deng, Hui Ren, Min Yang He, Da Hai Li, y Qiong Hua Wang. See-through 2D/3D compatible integral imaging display system using lens-array holographic optical element and polymer dispersed liquid crystal. *Optics Communications*, 456(May 2019):124615, 2020. ISSN 00304018. doi: 10.1016/j.optcom.2019.124615. URL <https://doi.org/10.1016/j.optcom.2019.124615>.
- Li Zhang y Shree Nayar. Projection defocus analysis for scene capture and image display. *ACM SIGGRAPH 2006 Papers on - SIGGRAPH '06*, page 907, 2006. doi: 10.1145/1179352.1141974. URL <http://portal.acm.org/citation.cfm?doid=1179352.1141974>.
- Qiang Zhang, Michael W Berry, Brian T Lamb, y Tabitha Samuel. A Parallel Nonnegative Tensor Factorization Algorithm for Mining Global Climate Data BT - Computational Science – ICCS 2009. *Computational Science – ICCS 2009*, 2009.
- Huiyuan Zhou, Khalid Tearo, Aniruddha Waje, Elham Alghamdi, Thamara Alves, Vinicius Ferreira, Kirstie Hawkey, y Derek Reilly. Enhancing mobile content privacy with proxemics aware notifications and protection. In *Conference on Human Factors in Computing Systems - Proceedings*, 2016. ISBN 9781450333627. doi: 10.1145/2858036.2858232.
- Matthias Zwicker, Wojciech Matusik, Frédo Durand, Hanspeter Pfister, y Clifton Forlines. Antialiasing for automultiscopic 3D displays. *ACM SIGGRAPH 2006 Sketches on - SIGGRAPH '06*, page 107, 2006. doi: 10.1145/1179849.1179983. URL <http://portal.acm.org/citation.cfm?doid=1179849.1179983>.

Apéndice A

Pseudocódigo para las reglas multiplicativas de WNMF

A.1. Generación de matriz de proyección del *light-field*

```

1  function [projectionMatrix , offVector] =
    generateProjectionMatrix(lf , screen , sel)
2
3      nloop          = prod(lf.POV);
4      nelem          = prod(lf.sizeIm);
5      cchannels      = sel.cchannels;
6      height         = lf.sizeIm(1);
7
8      limit          = length(screen.poslayer);
9      limit_u        = size(lf.lightField , 4);
10     limit_v         = size(lf.lightField , 3);
11
12     projectionMatrix = nan(nloop , nelem , cchannels , limit);
13
14     offVector       = zeros(nloop , limit);
15
16     for k = 1:limit
17
18         count = 1;
19         layer = screen.poslayer(k);
20
21         for POVu = 1:limit_u
22
23             slope_u = lf.slope.u(POVu) * layer;
24             chix    = 1.0 + round(slope_u);
25             chix_h  = (chix - 1.0) * height;
26
27             for POVv = 1:limit_v
28
29                 slope_v = lf.slope.v(POVv) * layer;
30                 chiy    = 1.0 + round(slope_v);
31                 chi     = chiy + chix_h;
32
33                 trimm   = chi - 1.0;
34                 offVector(count ,k) = -trimm;
35
36                 abs_trimm = abs(trimm);
37

```

```
38         for c = 1:cchannels
39
40             imgVec = lf.lightField(:, :, POVv, POVu,
41                                     c);
42             imgVec = imgVec(:);
43
44             if trimm > 0
45                 projectionMatrix(count, (abs_trimm +
46                                         1):end, c, k) = imgVec(1:(nelem -
47                                             abs_trimm));
48
49             elseif trimm < 0
50                 projectionMatrix(count, 1:(end - (
51                                         abs_trimm)), c, k) = imgVec((
52                                             abs_trimm + 1):end);
53
54             else
55                 projectionMatrix(count, :, c, k) =
56                     imgVec;
57
58             end
59
60         end
61
62         count = count + 1;
63
64     end
```

A.2. Producto capa-matriz de *light-field*

```
1 function num = layerProjectionMatrixProduct(projectionMatrix ,  
      offVec , vectorlayer , nview , numel)  
2  
3     vector = 1:numel;  
4     num     = zeros(1, numel);  
5  
6     for i = 1:nview  
7  
8         ind = vector + 2*offVec(i);  
9  
10        offset = (ind >= 1) & (ind <= numel);  
11  
12        values = projectionMatrix(i, :, 1);  
13  
14        layer = circshift(vectorlayer', -2*offVec(i))';  
15        res = values(offset) .* layer(offset);  
16        num(offset) = num(offset) + res;  
17  
18    end  
19  
20    num(isnan(num))=0;  
21  
22 end
```

A.3. Producto entre capas

```
1 function projectionMatrixAux = prodVec(f, g, offVec, nviews,
    numel)
2
3     projectionMatrixAux = nan(nviews, numel);
4
5     for i = 1:nviews
6
7         abs_recorte = abs(offVec(i));
8         view = f .* circshift(g', -2*offVec(i))';
9
10        if offVec(i) < 0
11            projectionMatrixAux(i, (abs_recorte + 1):end) =
                view((abs_recorte + 1):end);
12        elseif offVec(i) > 0
13            projectionMatrixAux(i, 1:(end - abs_recorte)) =
                view(1:(numel - abs_recorte));
14        else
15            projectionMatrixAux(i, :) = view;
16        end
17    end
18 end
19
20 end
```

A.4. Reglas multiplicativas para dos capas

```

1  function [F, G] = layerSynthesis(lf)
2
3  %dos pantallas
4  nbucle = prod(lf.POV);
5  nelem  = prod(lf.sizeIm);
6  F      = ones(1, nelem);
7  G      = ones(1, nelem);
8
9  %numerador
10 projectionMatrix = lf.projectionMatrix;
11 offVector = lf.offVector;
12
13 numiter = 50;
14
15 for n = 1:numiter
16
17     %F
18     num = layerProjectionMatrixProduct(projectionMatrix
19         (:,:,1,1), squeeze(offVector(:,1)), G, nbucle, nelem);
20     projectionMatrixAux = prodVec(F, G, squeeze(offVector
21         (:,1)), nbucle, nelem);
22     den = layerProjectionMatrixProduct(projectionMatrixAux,
23         squeeze(offVector(:,1)), G, nbucle, nelem);
24
25     F = F .* num./(den+0.01);
26     F(F > 1) = 1;
27
28     %G
29     num = layerProjectionMatrixProduct(projectionMatrix
30         (:,:,1,2), squeeze(offVector(:,2)), F, nbucle, nelem);
31     projectionMatrixAux = prodVec(G, F, squeeze(offVector
32         (:,2)), nbucle, nelem);
33     den = layerProjectionMatrixProduct(projectionMatrixAux,
34         squeeze(offVector(:,2)), F, nbucle, nelem);
35
36     G = G .* num./(den+0.01);
37     G(G > 1) = 1;

```

33 end

34

35 end

A.5. Síntesis

```
1 function wnmfLayerSynthesis_Optimization(lf, sel, screen)
2
3     [projectionMatrix, offVector] = generateProyectionMatrix(lf
4         , screen, sel)
5
6     lf.matrizProy = matrizProy;
7     lf.offVector = offVector;
8
9     [F, G] = layerSynthesis(lf)
10 end
```

Apéndice B

Correlación cruzada para tomografía atmosférica entre *Kaczmarz iterativo* y WNTF

		Dos capas			Tres capas			
		TS	TSO	SS	SSO	TS	TSO	SS
Dos capas	TSO	Desde 23.00arcsec						
	SS	Toda la curva	0-28.86arcsec					
	SSO	Desde 26.86arcsec	Desde 36.07arcsec	Toda la curva				
Tres capas	TS	Toda la curva	Desde 11.51arcsec	Toda la curva	Desde 5.37arcsec			
	TSO	Toda la curva	Toda la curva	Toda la curva	Toda la curva	0-16.88arcsec		
	SS	Iguales	Desde 14.58arcsec	Toda la curva	Desde 11.51arcsec	Desde 35.30arcsec	0-33.00arcsec	
	SSO	Toda la curva	Toda la curva	Toda la curva	Toda la curva	0-10.74arcsec	Iguales	

Figura B.1 Relación estadística entre los métodos propuestos. El color de cada celda indica qué método es mejor en términos estadísticos tras realizar la corrección atmosférica, siendo WNTF negro y *Kaczmarz* blanco. Las celdas grises indican que no es posible discriminar entre las técnicas. Se han añadido los rangos de la curva en los que el análisis tiene validez.