

Proyecto Fin de Carrera

Reenfocado de imágenes de alto rango dinámico
con aperturas codificadas y aperturas múltiples

Autor

Luis García Girón

Directoras y ponente

Lara Presa Irazábal
Belen Masiá Corcoy
Dr. Diego Gutiérrez Pérez

Escuela de Ingeniería y Arquitectura
2012

Reenfocado de imágenes de alto rango dinámico con aperturas codificadas y aperturas múltiples

PROYECTO DE FIN DE CARRERA

Autor: Luis García Girón
Directora: Lara Presa Irazábal
Codirectora: Belén Masiá Corcoy
Ponente: Dr. Diego Gutiérrez Pérez

Ingeniería Superior de Telecomunicaciones
Curso 2011-2012

Departamento de Informática e Ingeniería de Sistemas
Escuela de Ingeniería y Arquitectura
Universidad de Zaragoza

Julio de 2012



Escuela de
Ingeniería y Arquitectura
Universidad Zaragoza



Departamento de
Informática e Ingeniería
de Sistemas
Universidad Zaragoza



Graphics and
Imaging Lab

Resumen

En los últimos años, la investigación en el campo de la Fotografía Computacional ha dado lugar a numerosos avances en las múltiples disciplinas que lo componen, rebasando los límites de la fotografía convencional. Uno de estos límites es la incapacidad de las cámaras fotográficas de representar toda la gama de luminancias presentes en una escena, y que se solucionó tras la aparición de la técnica del alto rango dinámico o HDR (High Dynamic Range). Esta técnica permite fusionar una serie de imágenes de una escena, capturadas con distintos parámetros de exposición, de modo que la imagen resultante recoge todo el rango dinámico conjunto de las imágenes empleadas para su obtención.

En este proyecto se ha trabajado con imágenes HDR, y en particular con el problema de reenfoque de las mismas. El desenfoque, provocado por la limitada profundidad de campo o por una incorrecta elección de los parámetros de la cámara, es un problema clásico en fotografía. Se han implementado y analizado dos técnicas diferentes de **reenfoque de imágenes HDR**. La *primera* de ellas consiste en el empleo de **aperturas codificadas**, máscaras que se colocan en la lente de la cámara y que permiten codificar el desenfoque, facilitando su corrección a posteriori mediante técnicas computacionales. El empleo de aperturas codificadas es habitual en el reenfoque de imágenes convencionales de bajo rango dinámico o LDR (Low Dynamic Range), pero hasta la fecha estas técnicas no se habían extendido a las imágenes HDR y se desconocía su viabilidad.

La *segunda* técnica que se estudia es la de generación y reenfoque de imágenes HDR con **aperturas múltiples**. Con esta técnica se emplean imágenes de una misma escena obtenidas variando el tamaño de la apertura, por lo que no sólo varía la exposición entre ellas, sino también el desenfoque. Gracias a ello, se permite obtener al mismo tiempo información sobre la luminancia y sobre la profundidad de la escena, permitiendo obtener fotografías HDR de la misma, que además pueden ser reenfoqueadas como se desee.

La primera parte del proyecto ha dado lugar a una publicación que ha sido enviada al **Congreso Español de Informática Gráfica** (CEIG 2012), y que a fecha de hoy está pendiente de aceptación.

Índice general

1. Introducción	8
2. Antecedentes	13
2.1. Fotografía Computacional	13
2.2. Fotografía de alto rango dinámico	13
2.2.1. Concepto del rango dinámico	13
2.2.2. Creación de imágenes de alto rango dinámico	15
2.2.3. Líneas de investigación en HDR	16
2.3. Aperturas codificadas para corrección de desenfoque	18
2.3.1. Desenfoque por profundidad de campo	18
2.3.2. Análisis frecuencial del desenfoque	18
2.3.3. Líneas de investigación con aperturas codificadas	21
3. Aperturas codificadas para corrección de desenfoques en imágenes HDR	22
3.1. Introducción	22
3.2. Modelos de procesado	22
3.3. Simulación de los modelos de procesado	25
3.3.1. Construcción de los priors	27
3.4. Resultados de simulación	27
3.4.1. Modelo de procesado 1	27
3.4.2. Modelo de procesado 2	27
3.4.3. Modelo de procesado 3	27
3.4.4. Comparativa de los modelos	28
3.4.5. Comparativa de los priors	28
3.5. Validación física de los modelos de procesado	28
3.5.1. Captura de las imágenes desenfocadas con apertura codificada	29
3.5.2. Obtención de las PSFs	30
3.5.3. Recuperación de las imágenes bien enfocadas	31
3.6. Obtención de patrones óptimos	33
3.6.1. Función objetivo	34
3.6.2. Algoritmo genético	35

3.6.3.	Evaluación de los patrones obtenidos	36
3.7.	Validación física de los patrones obtenidos	38
3.8.	Resultados de la validación física	39
4.	Aperturas múltiples	45
4.1.	Introducción	45
4.2.	Fotografía HDR con apertura variable	45
4.3.	Modelo de formación de imágenes	47
4.3.1.	Modelo de exposición	47
4.3.2.	Modelo de desenfoque	47
4.3.3.	Modelo de luminancia por capas	48
4.3.4.	Modelo completo de escena	49
4.4.	Optimización	50
4.4.1.	Función objetivo	50
4.4.2.	Método de optimización	50
4.4.3.	Inicialización	50
4.5.	Definición de los experimentos	51
4.5.1.	Recuperación de luminancia L en imágenes sintéticas .	51
4.5.2.	Recuperación de luminancia L y tamaño de blur en imágenes reales	52
4.6.	Resultados	54
4.6.1.	Imágenes sintéticas	54
4.6.2.	Imágenes reales	54
5.	Conclusiones	59
5.1.	Conclusiones del trabajo realizado	59
5.2.	Trabajo futuro	60
5.3.	Conclusiones personales	61
A.	Software y procedimientos de fusión de imágenes HDR	64
A.1.	Adobe Photoshop CS5	64
A.2.	Luminance HDR	65
A.3.	Photomatix Pro	65
B.	Inserción de aperturas codificadas en el objetivo <i>Canon EF</i> <i>50mm f/1.8 II</i>	66
C.	Artículo: Analysis of Coded Apertures for Defocus Deblu- rring of HDR Images	69

Índice de figuras

1.1. Diagrama de Gantt.	11
2.1. Imagen HDR creada a partir de tres exposiciones.	16
2.2. Modelo de lente delgada 2D.	18
2.3. Tamaño de la PSF en relación a la distancia del plano focal.	19
2.4. Objetivo con apertura codificada de Zhou y Nayar.	20
2.5. Respuesta frecuencial de la apertura de Zhou y Nayar y la circular.	20
3.1. Esquema del procesado 1.	24
3.2. Esquema del procesado 2.	24
3.3. Esquema del procesado 3.	24
3.4. Ejemplos de imágenes usadas en simulación.	26
3.5. Q media obtenida en simulación.	28
3.6. Ejemplo de imágenes HDR recuperadas con prior HDR y prior LDR.	29
3.7. Esquema de montaje para la captura de las exposiciones.	30
3.8. Detalle central de las imágenes empleadas para obtener las PSFs.	31
3.9. PSFs recuperadas para la deconvolución.	31
3.10. Q obtenida en el escenario real.	32
3.11. Detalle de las imágenes recuperadas empleando prior.	33
3.12. Resultados de los mejores procesados HDR y LDR.	34
3.13. Patrones de aperturas obtenidos.	36
3.14. Q media obtenida para cada apertura.	37
3.15. Respuesta frecuencial de las aperturas obtenidas con prior HDR.	38
3.16. Respuesta frecuencial de las aperturas obtenidas con prior LDR.	38
3.17. Aperturas impresas en papel de transparencia.	39
3.18. Resultados con la apertura circular.	40
3.19. Resultados de ambos procesados para el detalle de la derecha.	41
3.20. Resultados de ambos procesados para el detalle de la izquierda.	43

3.21. Métricas obtenidas para ambos detalles con cada una de las aperturas.	44
4.1. Variación de la profundidad de campo y de la exposición en función del tamaño de la apertura.	46
4.2. Simulación del desenfoque por profundidad con diferentes tamaños de blur.	48
4.3. Esquema del modelo de creación de la luminancia de la imagen por capas de profundidad.	49
4.4. Esquema de la optimización en el experimento con imágenes sintéticas.	51
4.5. Esquema de la optimización en el experimento con imágenes reales.	52
4.6. Imagen sintética con exposición artificial y división en capas de profundidad.	53
4.7. Simulación de captura con tres tamaños de apertura diferentes.	54
4.8. Resultado de la optimización para recuperar la luminancia de la imagen sintética.	55
4.9. Luminancia recuperada de la imagen real, imágenes de entrada y división en capas.	57
4.10. Luminancia HDR recuperada reenfocada en postcaptura a cada una de las capas.	58
4.11. Manipulación de la exposición en postcaptura con la escena totalmente enfocada.	58
4.12. Manipulación del tamaño de apertura en postcaptura.	58
B.1. Objetivo <i>Canon EF 50mm f/1.8 II</i> junto a los materiales empleados.	66
B.2. Objetivo montado, con la tapa trasera desenroscada.	67
B.3. Objetivo con la tapa lateral despegada.	67
B.4. Objetivo desmontado en sus tres piezas principales.	68
B.5. Pieza central con la apertura codificada.	68

Capítulo 1

Introducción

Uno de los objetivos de la Fotografía Computacional es obtener mejores representaciones de las escenas reales que las que se consiguen empleando técnicas de fotografía convencionales. En este sentido, uno de los problemas que presentan las cámaras fotográficas comerciales es la limitación del sensor a la hora de capturar imágenes de escenas con un rango de luminancias, o rango dinámico, elevado.

Así, cuando el fotógrafo se dispone a capturar una de estas escenas, debe escoger el rango de luminancias de interés y variar adecuadamente los parámetros de exposición de la cámara para realizar la captura según su necesidad. Sin embargo, se da una pérdida importante de información, ya que aquellas partes de la escena que presenten una luminancia fuera del rango seleccionado no podrán ser representadas de forma correcta. A aquellos píxeles que correspondan con las zonas de mayor luminancia, el sensor les asignará invariablemente el máximo valor, y al revés para aquellos que se correspondan con las zonas más oscuras, de modo que al tratar de representar escenas con un rango dinámico elevado la pérdida de detalle será muy alta.

En este contexto, surge el concepto de imágenes de alto rango dinámico o HDR (*High Dynamic Range*). Este tipo de imágenes permiten representar un rango de luminancias más amplio que las fotografías convencionales o de bajo rango dinámico, LDR (*Low Dynamic Range*), permitiendo una mejor representación de la extensa gama de niveles de intensidad presentes en las escenas reales.

La obtención de estas imágenes se consigue capturando una serie de fotografías de una misma escena con distintos tiempos de exposición y fusionándolas posteriormente, mediante técnicas que se encuentran implementadas en la mayoría del software fotográfico convencional [4].

Este proyecto gira en torno a este tipo de imágenes. En concreto se centra en la codificación de la señal en su captura, para poder realizar un reenfoque de las mismas a posteriori tras un procesado. En función de cómo se realizan esta codificación y este procesado se divide el trabajo en dos partes bien diferenciadas.

La *primera parte* consiste en la extensión a las imágenes HDR de las técnicas existentes de aperturas codificadas para corrección de desenfoque [13, 14, 19]. Es sabido que el efecto producido al capturar una fotografía desenfocada puede modelarse como una convolución entre la escena perfectamente enfocada y un kernel que viene determinado por la forma de la apertura. Así pues, conociendo esta forma, puede parecer sencillo recuperar la imagen bien enfocada mediante un simple proceso de deconvolución. Sin embargo, las aperturas circulares convencionales ofrecen una respuesta frecuencial muy desfavorable, ya que producen serias atenuaciones en las frecuencias más altas de la imagen, al tiempo que presentan varios cruces por cero en el dominio frecuencial, haciendo imposible recuperar correctamente la imagen.

Estos problemas motivan la aparición de las aperturas codificadas, como máscaras que se colocan delante de la lente de la cámara y que permiten cambiar la forma de la apertura y, por tanto, codificar el desenfoque, obteniendo patrones que eviten los problemas que presentan las aperturas circulares.

En la actualidad existe una gran variedad de aperturas que presentan un buen comportamiento para la corrección de desenfoques [13, 14, 19], y está sobradamente demostrada su viabilidad.

Sin embargo, hasta la fecha no se ha estudiado la utilidad de emplear estas aperturas codificadas en la captura de imágenes HDR. Además, las aperturas existentes han sido optimizadas para ser usadas con imágenes LDR, y dado que las HDR presentan una serie de diferencias en las estadísticas de imagen (e.g. espectro frecuencial o distribución del histograma) respecto de las primeras [16], cabe pensar que los patrones existentes quizá no sean óptimos para HDR. Así pues, se hace necesario un estudio a fondo del problema.

La segunda parte del proyecto consiste en el estudio de la novedosa técnica de generación y reenfoque de imágenes HDR mediante aperturas múltiples [9, 10], cuya principal peculiaridad radica en que la obtención de las distintas exposiciones de la escena se realiza dejando fijo el tiempo de exposición y modificando el tamaño de la apertura, teniendo en cuenta que

cuanto menor sea el diámetro de ésta, menor será la cantidad de luz que llega al sensor.

Manipulando el tamaño de la apertura, varía no sólo la exposición sino también el desenfoque, ya que una apertura más pequeña presentará una profundidad de campo mayor, permitiendo obtener una fotografía enfocada de la mayor parte de la escena, mientras que una apertura más grande lleva asociada una profundidad de campo menor, por lo que aquellas partes de la escena que se encuentren lejos de la focal aparecerán desenfocadas.

Esto puede parecer problemático a la hora de fusionar las exposiciones para obtener la imagen HDR. Sin embargo, puede obtenerse una ventaja de este problema, ya que capturando las exposiciones de esta manera se obtiene al mismo tiempo información sobre la luminancia de la escena y sobre el desenfoque, permitiendo, tras procesar las imágenes de forma adecuada, obtener una fotografía de la escena en HDR que además puede ser reenfocada como se desee.

Así pues, los objetivos principales de este proyecto son los siguientes:

- Estudio exhaustivo del **empleo de las aperturas codificadas en imágenes HDR**.
- Implementación y análisis de la técnica de **captura y reenfoque de imágenes HDR mediante aperturas múltiples** [9, 10].

Para la consecución del primero de los objetivos, se ha seguido el procedimiento que se detalla a continuación:

- Se ha comprobado la viabilidad del uso de aperturas codificadas en imágenes HDR, empleando algunas aperturas conocidas [19], simulando la captura de las imágenes desenfocadas y su posterior reenfoque.
- Se han obtenido patrones de aperturas calculados específicamente para imágenes HDR, empleando algoritmos genéticos.
- Se han obtenido las aperturas físicas y se han realizado una serie de experimentos reales tomando fotografías desenfocadas con ellas, evaluando su comportamiento.

Para la consecución del segundo objetivo, el procedimiento llevado a cabo es el siguiente:

- Se ha implementado un sistema simplificado de generación y reenfoque de imágenes HDR mediante aperturas múltiples a partir de los trabajos ya existentes [9, 10].

- Se ha comprobado el funcionamiento de dicho sistema con una serie de imágenes, tanto simuladas como capturadas en un escenario real.

Todo el trabajo se ha desarrollado dentro del Graphics and Imaging Lab, perteneciente al Grupo de Informática Gráfica Avanzada (GIGA) de la Universidad de Zaragoza. La parte de aperturas codificadas para imágenes HDR ha dado lugar a una publicación enviada al Congreso Español de Informática Gráfica (CEIG 2012), que se encuentra pendiente de aceptación.

La duración del proyecto ha sido de un año aproximadamente, durante el cual se han desarrollado una serie de actividades cuya distribución temporal se muestra en el diagrama de Gantt de la Figura 1.1.

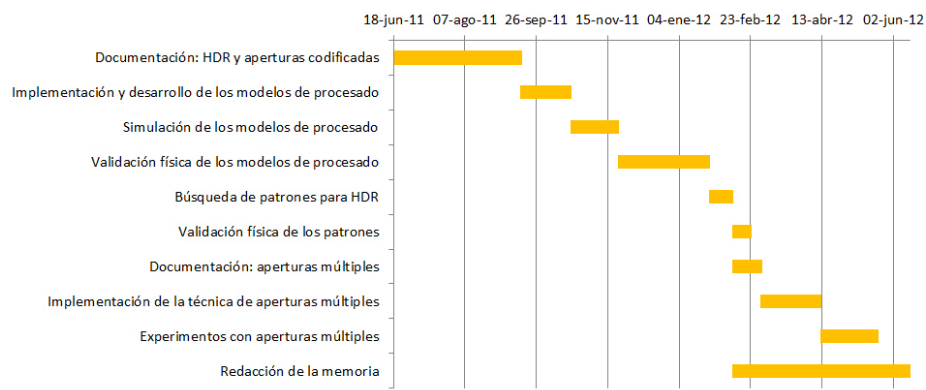


Figura 1.1: Diagrama de Gantt de las actividades realizadas

Esta memoria está estructurada en 5 capítulos, incluyendo el presente capítulo introductorio. En el Capítulo 2 se resumen los antecedentes del proyecto realizado, introduciendo los conceptos de las imágenes de alto rango dinámico y de las aperturas codificadas, explicando algunas de las técnicas en que se apoya este trabajo y resumiendo brevemente las líneas de investigación relacionadas con él.

En el Capítulo 3 se describe el trabajo realizado correspondiente a la primera parte del proyecto, sobre aperturas codificadas para corrección de desenfoques en imágenes HDR, explicando las técnicas que se han desarrollado y los procedimientos y experimentos que se han llevado a cabo y, finalmente, mostrando y evaluando los resultados.

En el Capítulo 4 se describe el trabajo referido a la segunda parte del proyecto, que trata del empleo de aperturas múltiples para obtención y reenfoque de imágenes HDR. Se explican las técnicas y el procedimiento empleados, se detallan los experimentos realizados y se muestran y discuten

los resultados obtenidos.

En el Capítulo 5 se exponen las conclusiones del trabajo realizado y se ofrecen una serie de sugerencias para investigaciones futuras, así como algunas apreciaciones de carácter personal.

Finalmente, se incluyen tres apéndices. El Apéndice A ofrece un breve resumen sobre el software comercial empleado para la fusión de imágenes HDR. El Apéndice B describe el procedimiento de inserción de una apertura codificada en el objetivo empleado durante los experimentos. Por último, el Apéndice C incluye el artículo *Analysis of Coded Apertures for Defocus Deblurring of HDR Images*, resultante de la investigación correspondiente a la primera parte del proyecto y sometido al Congreso Español de Informática Gráfica (CEIG 2012).

Capítulo 2

Antecedentes

2.1. Fotografía Computacional

La Fotografía Computacional es un campo multidisciplinar que combina elementos de óptica, procesamiento de imágenes, electrónica, visión por computador, informática gráfica o percepción. Este campo ha crecido enormemente en los últimos años, persiguiendo generalmente el objetivo de superar los límites de la fotografía convencional y obtener mejores representaciones y mayor cantidad de información de las escenas reales.

Tradicionalmente, las cámaras fotográficas se han diseñado con la idea de obtener representaciones bidimensionales y tricromáticas de las escenas, pero la Fotografía Computacional va más allá, introduciendo nuevas propiedades y conceptos como, entre otros muchos, espectros de color, variaciones espaciales y temporales, cambios en la dirección de la luz, extensión del rango dinámico o aperturas codificadas, siendo estos dos últimos objeto de estudio de este proyecto.

2.2. Fotografía de alto rango dinámico

2.2.1. Concepto del rango dinámico

En fotografía, se conoce como rango dinámico a la diferencia de luminosidad presente entre las partes más luminosas y las partes más oscuras de una escena. Se mide en diferencias de Valor de Exposición o *Exposure Value* (EV), también conocidas como *stops*. Estrictamente, el EV no es una medida de luminosidad propiamente dicha, sino una relación entre los distintos parámetros de la cámara que determinan la exposición de una fotografía, y viene dado por la ecuación:

$$EV = \log_2 \frac{N^2}{t}, \quad (2.1)$$

donde N es la apertura relativa, que determina el nivel de iluminación y t es el tiempo de exposición en segundos, que viene dado por la velocidad del obturador. No obstante, se emplea el EV como medida de luminosidad, relacionándolo con la luminancia para la cual una cámara emplearía ese EV para obtener la imagen correctamente expuesta. Por ejemplo, si en una escena determinada se necesita un Valor de Exposición de EV_1 stops para obtener una exposición correcta de las zonas más oscuras, y otro de EV_2 stops para obtener la exposición correcta de las zonas más luminosas, el rango dinámico de esa escena sería:

$$DR = EV_2 - EV_1 \quad (2.2)$$

El rango dinámico que las cámaras fotográficas convencionales son capaces de capturar correctamente es bastante inferior al que puede distinguir el ojo humano. Además de las limitaciones de los sensores, los conversores analógico-digitales restringen todavía más el rango dinámico de las fotografías. En el Cuadro 2.1 puede verse una relación entre los bits de precisión del conversor y el rango dinámico que permitirían obtener¹.

Bits de precisión	Ratio de contraste	Rango dinámico (stops)
8	256:1	8
10	1024:1	10
12	4096:1	12
14	16384:1	14
16	65536:1	16

Cuadro 2.1: Rango dinámico respecto al número de bits de precisión.

Así, un conversor A/D con 8 bits de precisión permitiría obtener imágenes con 8 bits de profundidad, es decir, con valores de píxel que irían entre 0 y 255, lo que se traduce en un rango dinámico de 8 stops. Sin embargo, todavía existe otra limitación debida a los niveles de ruido, que provoca una reducción aún mayor del rango dinámico. De este modo, aunque los conversores A/D que se emplean habitualmente en las cámaras digitales comerciales presentan entre 10 y 14 bits de precisión, el rango dinámico que son capaces de manejar se ve reducido hasta los 5-9 stops, muy por debajo de los cerca de 14 stops que el ojo humano es capaz de percibir. Estas limitaciones justifican la aparición de las imágenes de alto rango dinámico o HDR (*High Dynamic Range*), respondiendo a la necesidad de conseguir mejores aproximaciones del sistema visual humano a la hora de capturar fotografías.

¹Adaptado de <http://www.cambridgeincolour.com/tutorials/dynamic-range.htm>

2.2.2. Creación de imágenes de alto rango dinámico

En 1997, con la expansión de las cámaras fotográficas digitales en el mercado, y ante las limitaciones de rango dinámico comentadas anteriormente, surge la técnica de creación de imágenes de alto rango dinámico a partir de una serie de fotografías de la misma escena realizadas con distintos parámetros de exposición, desarrollada por Paul E. Debevec y Jitendra Malik [4].

Esta técnica se basa en la recuperación, a partir de esta serie de imágenes, de la curva de respuesta del proceso de formación de la imagen en la cámara (g). Conociendo esta función, se pueden procesar las distintas exposiciones y fusionarlas en una sola imagen de alto rango dinámico, cuyos píxeles presentan valores proporcionales a los valores reales de luminancia presentes en la escena, atendiendo a la siguiente fórmula:

$$\ln E_i = \frac{\sum_{j=1}^P w(Z_{ij})(g(Z_{ij}) - \ln \Delta t_j)}{\sum_{j=1}^P w(Z_{ij})} \quad (2.3)$$

Donde P es el número de fotografías empleadas, Z_{ij} es el valor del píxel i en la fotografía j , Δt_j es el tiempo de exposición de la fotografía j , y w es una función de pesos que se emplea para suavizar el resultado final. Habitualmente suele emplearse una función triangular o una gaussiana para cumplir este objetivo. Para información más detallada sobre el proceso de recuperación de la función g y otras especificaciones matemáticas de esta técnica se recomienda consultar la publicación original [4].

El rango dinámico de la imagen resultante comprende el rango dinámico total de las imágenes que la conforman, y los valores de los píxeles se almacenan en un vector de valores en coma flotante de 32 bits, en contraste con los 8 bits que se emplean habitualmente en las imágenes convencionales de bajo rango dinámico, lo que supone un aumento del contraste máximo de 256:1 hasta 4.2950e9:1.

En la Figura 2.1 se muestra un ejemplo de imagen de alto rango dinámico obtenida gracias a esta técnica, a partir de una serie de imágenes con distintos parámetros de exposición. Puede verse claramente cómo tanto las partes más luminosas como las más oscuras quedan perfectamente representadas en la imagen resultante, mientras que en cada una de las exposiciones individuales aparecen zonas subexpuestas, como la librería en 2.1(b) y sobreexpuestas, como la parte de la ventana en 2.1(d). A la imagen HDR se le ha aplicado una operación conocida como reproducción de tono o *tone-mapping*. Esta operación consiste en una reducción del contraste global de la imagen, generalmente a cambio de un aumento de los contrastes locales. En HDR se emplea para obtener un registro completo de todos los niveles

de brillo de la escena original, y se hace necesaria a la hora de mostrar una imagen HDR en displays LDR o en papel.



(a) Fusión HDR con reproducción de tono



(b) Imagen subexpuesta



(c) Exposición central



(d) Imagen sobreexpuesta

Figura 2.1: Imagen HDR (a) creada a partir de la fusión de tres exposiciones (b,c,d). Nótese cómo se consigue representar correctamente todo el rango de luminancias de la escena, mientras que en las tres exposiciones aparecen inevitablemente zonas sobreexpuestas y subexpuestas.

2.2.3. Líneas de investigación en HDR

En la actualidad, la técnica de creación de imágenes HDR desarrollada por Debevec y Malik sigue vigente y está implementada en gran parte del software fotográfico convencional (Photoshop, Luminance HDR, Photomatix Pro). Basándose en esta técnica, una de las líneas de investigación actuales trata de analizar la secuencia óptima de exposiciones que se deben tomar para emplearlas en la formación de la imagen de alto rango dinámico y obtener los mejores resultados, como hicieron Grosberg y Nayar en 2003 [6]. En la misma línea aparece en 2006 el trabajo de Akyüz y Reinhard [2], y el de Hasinoff *et al.* en 2010 [8], con el añadido de que ambos consideran la

ISO² de la cámara entre los parámetros variables para optimizar la captura de las exposiciones.

Otra línea de investigación se desarrolla en torno al análisis de las propiedades estadísticas de las imágenes de alto rango dinámico. En este campo destaca el trabajo de Pouli *et al.* [16], que establece una serie de regularidades estadísticas en imágenes LDR y HDR, observando claras diferencias entre ambas. La principal importancia de este estudio recae en que muchos procesos de optimización requieren del conocimiento de información a priori sobre las regularidades estadísticas de imágenes naturales para obtener resultados adecuados. A esta información se la conoce como prior, y entre otras muchas aplicaciones se emplea, como se verá más adelante, en la obtención de aperturas codificadas para corrección de desenfoques. El hecho de que existan estas claras diferencias entre las regularidades estadísticas de las imágenes de bajo rango dinámico y las de alto rango dinámico implicaría que los priors existentes para las primeras no serían óptimos para usarse en aplicaciones HDR.

Otra dirección a destacar es la que comenzaron Hasinoff y Kutulakos en 2007 [9, 10] que permite construir una imagen HDR a partir de fotografías tomadas variando el tamaño de la apertura en lugar del tiempo de exposición, y que es uno de los objetos de estudio de este proyecto.

En cuanto a hardware fotográfico también se han logrado varios avances importantes en los últimos años. A la hora de aumentar el rango dinámico que son capaces de capturar las cámaras en una sola fotografía, destaca el trabajo de Nayar y Branzoi [15], consistente en una adaptación individual de la exposición de cada píxel en el sensor, permitiendo capturar imágenes HDR con un solo disparo. También aparecen nuevas técnicas de hardware a la hora de facilitar la captura de las distintas exposiciones antes de fusionarlas en una imagen HDR. En este sentido, Aggarwal y Ahuja desarrollaron un dispositivo [1] para dividir el camino óptico en la cámara fotográfica empleando prismas, de modo que cada uno de los haces resultantes alcanza un sensor diferente, obteniendo distintas exposiciones de una misma imagen con un solo disparo y permitiendo incluso la captura de vídeo de alto rango dinámico.

²Sensibilidad de la cámara fotográfica. A mayor ISO, menor es la cantidad de luz que se requiere para realizar una fotografía, a costa de un aumento del nivel de ruido.

2.3. Aperturas codificadas para corrección de desenfoque

2.3.1. Desenfoque por profundidad de campo

Según el modelo de lente delgada, cuando un objeto está situado en el plano focal de la lente de la cámara, todos los rayos provenientes de un mismo punto de ese objeto convergen en un mismo punto del sensor. Sin embargo, al alejar el objeto del plano focal, los rayos ya no convergen en un mismo punto sino en varios, que se corresponden con el llamado círculo de confusión, dando lugar a una imagen desenfocada.

En la Figura 2.2 se muestra un esquema del modelo de lente delgada. Tal como se ha dicho, el objeto situado en el plano focal, a una distancia F de la lente, da lugar a una imagen nítida puesto que todos sus rayos convergen en un único punto del sensor. En cambio, los objetos situados a distancias $D1$ y $D2$ darán lugar a imágenes desenfocadas, puesto que los rayos que provienen de un mismo punto convergen en múltiples puntos del sensor de la cámara, dando lugar al círculo de confusión. Además, cuanto mayor sea la distancia del objeto al plano focal, mayor será el grado de desenfoque de la imagen, es decir, mayor será el diámetro del círculo de confusión.

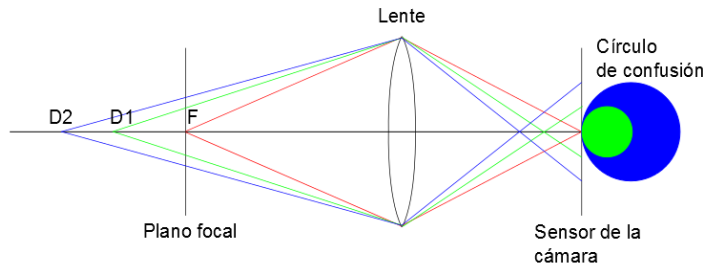


Figura 2.2: Modelo de lente delgada en 2D mostrando el efecto de desenfoque por profundidad de campo. Cuanto mayor es la distancia del objeto al plano focal, mayor es el círculo de confusión y, por tanto, el desenfoque.

2.3.2. Análisis frecuencial del desenfoque

Matemáticamente, el proceso de creación de la imagen f desenfocada en la cámara fotográfica se modela como una convolución entre la escena perfectamente enfocada f_0 y un kernel k que representa la PSF (*Point Spread Function*) o respuesta del sistema óptico a un impulso en la entrada, y que tiene la forma de la apertura con un tamaño variable en función de la profundidad, más un ruido η que habitualmente se asume gaussiano:

$$f = f_0 * k + \eta \quad (2.4)$$

Como puede verse en la Figura 2.3, cuanto más próxima esté la escena a la distancia focal, menor será el tamaño del kernel, y por tanto menor el desenfoque, de manera que para una fotografía bien enfocada el kernel se puede aproximar por un punto, de modo que el resultado de la convolución entre f_0 y k será la propia f_0 .

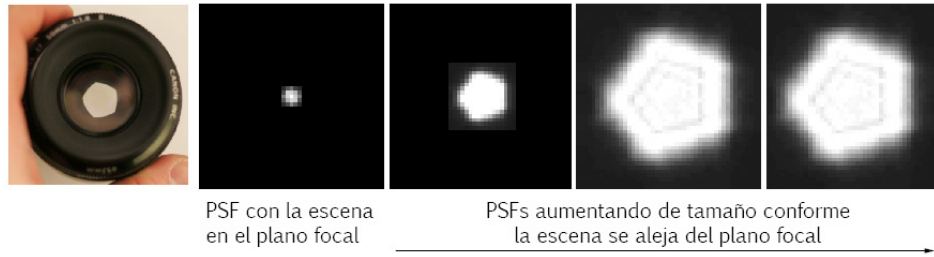


Figura 2.3: Tamaño de la PSF en relación a la distancia del plano focal. Cuanto mayor es esta distancia, mayor es el tamaño. Adaptada de Levin *et al.* [11].

En el dominio frecuencial, la ecuación se puede escribir como:

$$F = F_0 \cdot K + \zeta \quad (2.5)$$

Donde F , F_0 y ζ son las transformadas discretas de Fourier de f , f_0 y η , respectivamente.

Conociendo esto, puede parecer trivial recuperar una imagen enfocada a partir de otra desenfocada sin más que realizando la correspondiente deconvolución. Sin embargo el resultado no es todo lo bueno que cabría esperar, dado que las aperturas circulares presentan una respuesta frecuencial muy desfavorable, introduciendo una atenuación muy grande en las frecuencias más altas, así como varios cruces por cero que imposibilitan la recuperación de la información en esas componentes frecuenciales.

A partir de este problema surge la idea del empleo de aperturas codificadas para la corrección de desenfoques. Se trata de máscaras que, colocadas delante de la lente de la cámara (ver Figura 2.4), modifican la forma de la apertura, permitiendo codificar la luz que llega al sensor. Así se consigue una doble ventaja, ya que se obtiene un mayor control en la captura de la imagen codificando la forma del desenfoque, al tiempo que se resuelven los problemas de las aperturas circulares, creando patrones con respuestas frecuenciales más favorables para la corrección de desenfoque.

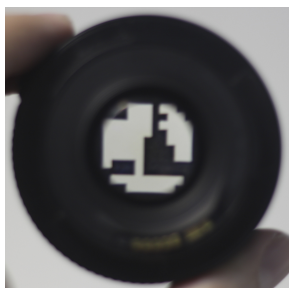


Figura 2.4: Objetivo con una de las aperturas codificadas desarrolladas por Zhou y Nayar [19]. Colocando una máscara en la lente del objetivo se consigue codificar el desenfoque, haciendo más sencilla su corrección.

En la Figura 2.5 puede verse la respuesta frecuencial de una de las aperturas de Zhou y Nayar [19], la misma que aparece en la Figura 2.4, frente a la respuesta frecuencial de la circular.

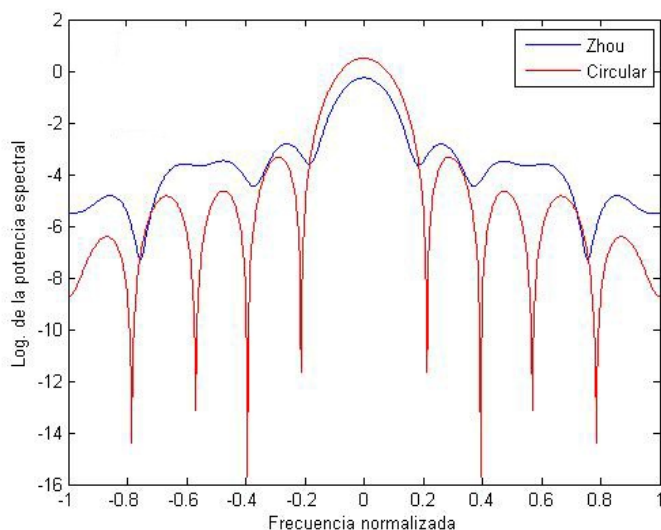


Figura 2.5: Logaritmo de la potencia espectral de una de las aperturas codificadas de Zhou [19] frente a la de la apertura circular. La apertura codificada ofrece una respuesta más favorable, al eliminar los cruces por cero y reducir la atenuación en las frecuencias altas.

Puede comprobarse cómo esta apertura tiene una respuesta frecuencial mucho más favorable que la circular, ya que la atenuación en las frecuencias altas es mucho menor y evita los cruces por cero.

2.3.3. Líneas de investigación con aperturas codificadas

El origen de las aperturas codificadas se remonta a los años 60. Habitualmente se empleaban en astronomía para resolver los problemas de ruido en la formación de imágenes sin lentes a partir de rayos X y rayos γ [3]. En esta línea aparecieron muchos trabajos y se desarrollaron numerosos patrones de aperturas, de entre los que destacan los patrones MURA (*Modified Uniformly Redundant Array*) [5].

Con el auge de la fotografía computacional, aparece el trabajo de Veeraghavan *et al.*, que emplea aperturas codificadas para obtener light fields 4D mediante una cámara convencional modificada [17].

En lo referente al problema del desenfoque y su relación con la profundidad (*depth from defocus*), Levin *et al.* consiguen una recuperación simultánea de la imagen enfocada y el mapa de profundidad, diseñando una apertura óptima basada en un criterio de discriminación de profundidad y explotando la caracterización estadística de las imágenes [11]. Zhou *et al.* consiguen el mismo objetivo empleando pares de aperturas codificadas, obtenidas con algoritmos genéticos [18]. Zhou *et al.* también desarrollan una métrica para evaluar la calidad de las aperturas codificadas para recuperación de enfoque, basada en la calidad de las imágenes reenfocadas y teniendo en cuenta las estadísticas de imágenes naturales, obteniendo aperturas óptimas [19]. Masiá *et al.* introducen las métricas perceptuales en el proceso de obtención de aperturas óptimas para corrección de desenfoques [14] y exploran el uso de valores no binarios en las aperturas codificadas [13].

Capítulo 3

Aperturas codificadas para corrección de desenfoques en imágenes HDR

3.1. Introducción

En esta parte del proyecto se estudia la aplicación de las técnicas de aperturas codificadas para corrección de desenfoques en imágenes HDR.

En primer lugar se proponen tres modelos de procesamiento diferentes para adaptar estas técnicas a las imágenes HDR, y se estudia su viabilidad mediante una serie de simulaciones, contrastadas con la captura y procesamiento de una imagen HDR real con apertura codificada.

A continuación se estudia la obtención de aperturas codificadas para HDR basándose en algoritmos genéticos y con priors estadísticos HDR y LDR.

Finalmente se realiza una validación por soporte físico de las aperturas obtenidas, insertando el patrón de la apertura codificada en el objetivo de la cámara y capturando las imágenes necesarias para su posterior procesamiento.

3.2. Modelos de procesamiento

Para estudiar la viabilidad del empleo de aperturas codificadas en imágenes HDR se simula el proceso de captura de las mismas y se trata de recuperar la imagen bien enfocada a partir de la imagen simulada. Como se ha comentado anteriormente, el proceso de creación de una imagen f en la cámara fotográfica, con un kernel k a partir de una escena f_0 viene dado

por la Ecuación 2.4:

$$f = f_0 * k + \eta$$

Así, para simular la captura de una imagen LDR desenfocada con una apertura codificada basta con realizar una convolución entre una imagen bien enfocada y el patrón de la apertura y añadirle un ruido gaussiano $\eta \sim N(0, \sigma^2)$. Sin embargo, esta simulación que es válida para LDR sólo sería válida para HDR en el caso de que pudiéramos capturar la imagen de alto rango dinámico a partir de una sola fotografía. Aunque existen cámaras capaces de capturar imágenes con rangos dinámicos más altos que las cámaras convencionales, como se ha comentado anteriormente en la mayoría de los casos la obtención de una imagen HDR se realiza capturando una serie de N exposiciones de la escena y fusionándolas posteriormente. Así, siendo $f_{0n}^{LDR}, (n = [1...N])$ las distintas exposiciones de una escena f_0^{HDR} correctamente enfocada, podemos simular la captura de una fotografía HDR desenfocada de esa misma escena f^{HDR} simulando las capturas por separado de cada una de las exposiciones f_n^{LDR} como:

$$f_n^{LDR} = f_{0n}^{LDR} * k + \eta \quad (3.1)$$

Y fusionándolas posteriormente:

$$f^{HDR} = g(f_1^{LDR}, f_2^{LDR}, ..., f_N^{LDR}) \quad (3.2)$$

Una vez obtenida f^{HDR} se recupera la imagen HDR enfocada \hat{f}_0^{HDR} realizando una deconvolución. Sin embargo, teniendo las imágenes LDR desenfocadas de cada una de las exposiciones, también será posible recuperar las exposiciones enfocadas $\hat{f}_{0n}^{LDR}, (n = 1, ...N)$ por separado, realizando N deconvoluciones, y fusionarlas posteriormente para obtener la imagen HDR recuperada:

$$\hat{f}_0^{HDR} = g(\hat{f}_{01}^{LDR}, \hat{f}_{02}^{LDR}, ..., \hat{f}_{0N}^{LDR}) \quad (3.3)$$

A partir de lo anterior definimos tres tipos de procesamiento diferentes que se resumen de la siguiente manera:

1. **Procesado de la imagen HDR obtenida a partir de una sola captura:** Éste procedimiento sólo es válido con cámaras fotográficas que permitan capturar imágenes con un rango dinámico superior al de las cámaras convencionales, por lo que sólo podrá ser validado en simulación al no disponer del material necesario para realizar experimentos reales. Su esquema se muestra en la Figura 3.1.

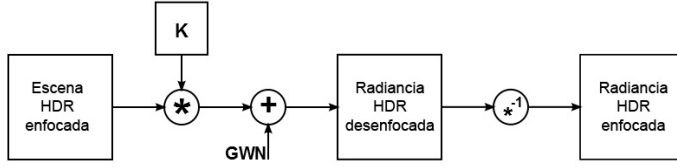


Figura 3.1: Esquema del procesamiento 1. GWN representa el ruido gaussiano, K el kernel del desenfoque y $*$ el operador de convolución.

2. **Procesado de la imagen HDR obtenida a partir de una serie de exposiciones:** En primer lugar se simula la captura de una serie de exposiciones desenfocadas de la imagen (Ecuación 3.1) para fusionarlas en una sola imagen HDR (Ecuación 3.2) de la que se obtendrá la imagen recuperada a partir de una deconvolución. Su esquema se muestra en la Figura 3.2

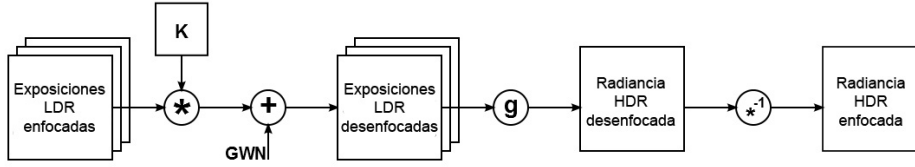


Figura 3.2: Esquema del procesamiento 2. GWN representa el ruido gaussiano, K el kernel del desenfoque y $*$ el operador de convolución.

3. **Procesado de las distintas exposiciones antes de obtener la imagen HDR:** Se simula la captura de una serie de exposiciones desenfocadas (Ecuación 3.1) a partir de las cuales se obtendrán las exposiciones recuperadas mediante una serie de deconvoluciones. A continuación se obtiene la imagen HDR recuperada fusionando las exposiciones recuperadas (Ecuación 3.3). Su esquema se muestra en la Figura 3.3

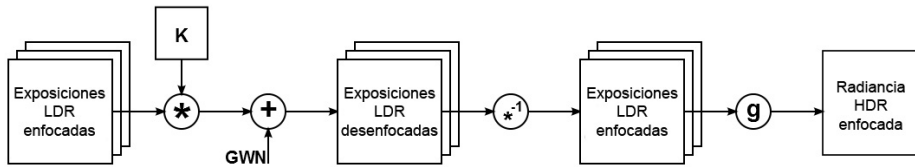


Figura 3.3: Esquema del procesamiento 3. GWN representa el ruido gaussiano, K el kernel del desenfoque y $*$ el operador de convolución.

Las operaciones necesarias para llevar a cabo cada una de las simulaciones se resumen en el Cuadro 3.1. Como puede verse, el procesamiento 1 es el que

menos coste computacional presenta, al requerir tan sólo una convolución y una deconvolución, si bien, como ya se ha comentado, sólo podrá comprobarse su viabilidad en simulación. El procesado 3 es el más costoso al requerir una deconvolución por cada exposición. Sin embargo este último procesado es el único que ofrece garantías de funcionamiento a priori, ya que está demostrada la viabilidad del empleo de aperturas codificadas en imágenes LDR [19], por lo que es de suponer que la recuperación de las distintas exposiciones ofrecerá buenos resultados, que se extenderán a la imagen HDR recuperada tras la fusión.

	Procesado 1	Procesado 2	Procesado 3
Convolución	1	N	N
Fusión HDR	0	1	1
Deconvolución	1	1	N

Cuadro 3.1: Operaciones necesarias para cada procesado empleando N exposiciones.

3.3. Simulación de los modelos de procesado

Para realizar las simulaciones de los distintos modelos de procesado, empleamos una de las aperturas de Zhou, la misma que se muestra en el Apartado 2.3.2, cuya validez para la corrección de desenfoques en imágenes LDR está demostrada [19].

Estudiamos los distintos modelos realizando simulaciones con cuatro niveles de ruido gaussiano diferentes (desviaciones típicas: $\sigma = 0.0005, 0.001, 0.005, 0.05$) y para tres modelos de deconvolución a partir de la deconvolución de Wiener, que viene dada por la siguiente ecuación:

$$\hat{F}_0 = \frac{F \cdot \bar{K}}{|K|^2 + |C|^2} \quad (3.4)$$

Donde \hat{F}_0 es la imagen recuperada, \bar{K} es el complejo conjugado de K , $|K|^2 = K \cdot \bar{K}$ y $|C|^2 = |\sigma/F_0|^2$ es la matriz de relación ruido a señal NSR (*Noise to Signal Ratio*) de la imagen original.

Así pues, a partir de esta fórmula estudiaremos estas tres variantes:

- **Deconvolución de Wiener sin prior, con NSR constante.** Sustituyendo $|C|^2$ en la Ecuación 3.4 por una constante. Se probaron varios valores, encontrando que al aumentarlo se obtienen resultados menos

ruidosos a costa de un aumento del *ringing*¹. Finalmente se decide emplear un valor intermedio $|C|^2 = 0,005$, alcanzando un compromiso entre ambos efectos.

- **Deconvolución de Wiener empleando un prior de imágenes naturales HDR.** Sustituyendo F_0 en $|C|^2 = |\sigma/F_0|^2$ por un prior estadístico A^{HDR} , una matriz que promedia la potencia espectral de una serie de 198 imágenes HDR naturales, en la Ecuación 3.4.
- **Deconvolución de Wiener empleando un prior de imágenes naturales LDR.** Sustituyendo F_0 en $|C|^2 = |\sigma/F_0|^2$ como en el anterior, pero con un prior A^{LDR} de 198 imágenes LDR.

También se estudió la posibilidad de emplear el modelo de deconvolución de Richardson-Lucy, pero al tratarse de una deconvolución iterativa era necesario realizar un número de repeticiones muy elevado para alcanzar soluciones aceptables y, en todo caso, peores que las obtenidas con la deconvolución de Wiener, incrementando considerablemente el tiempo de cálculo, por lo que se descartó esta idea.

Para estudiar la validez de los modelos se escoge un set de siete fotografías HDR obtenidas cada una de ellas a partir de tres exposiciones, con diversos rangos dinámicos y distintos parámetros de captura. Algunas de estas imágenes se muestran en la Figura 3.4. Se trata de recuperar la imagen original y se emplea la métrica *HDR-VDP2*²[12] para obtener una medida de calidad de los resultados.

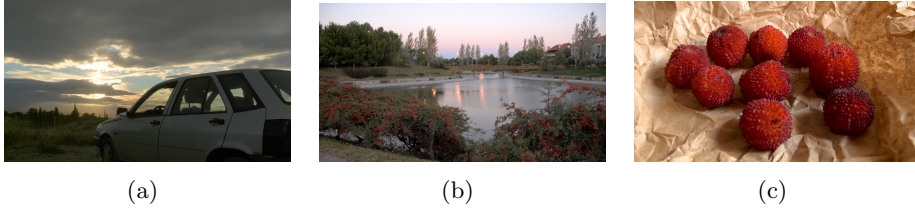


Figura 3.4: Algunas de las imágenes HDR empleadas en la simulación de los procesos. Todas se muestran tras aplicar reproducción de tono.

¹Artefacto indeseado que aparece habitualmente en las transiciones bruscas de una señal. En el caso de las imágenes, se muestran como oscilaciones alrededor de los bordes, y aparecen frecuentemente tras una deconvolución al perder información en las altas frecuencias.

²Se trata de una métrica perceptual que compara una imagen de referencia con otra distorsionada, dando medidas de *calidad* y de *visibilidad* (probabilidad de que el usuario detecte las diferencias entre ambas imágenes).

3.3.1. Construcción de los priors

Para construir los priors escogemos una muestra suficientemente grande de imágenes HDR y LDR, de modo que la información estadística obtenida sea lo bastante fiable y generalizada. Empleamos 198 imágenes de cada tipo, obtenidas de la base de datos de Pouli³[16]. De cada imagen se calcula la potencia espectral como:

$$S = F \cdot \bar{F} \quad (3.5)$$

Siendo F la transformada de Fourier de una imagen y \bar{F} su complejo conjugado. Se obtiene el promedio de las potencias espectrales de todas las imágenes HDR para obtener el prior A^{HDR} y lo mismo con las LDR para A^{LDR} .

3.4. Resultados de simulación

3.4.1. Modelo de procesado 1

La calidad Q media de las siete imágenes, obtenida con la métrica *HDR-VDP2*[12], para cada tipo de deconvolución y nivel de ruido se muestra en la Figura 3.5(a). Puede comprobarse cómo, para niveles altos de ruido, la deconvolución de Wiener se comporta mejor empleando priors. Igualmente, se observa que el empleo de un prior HDR tiene ventajas claras sobre el empleo de un prior LDR cuando el ruido es muy alto. Para niveles de ruido bajos, los tres tipos de deconvolución de Wiener presentan resultados similares.

3.4.2. Modelo de procesado 2

La calidad Q media de las siete imágenes, obtenida con la métrica *HDR-VDP2*[12], para cada tipo de deconvolución y nivel de ruido se muestra en la Figura 3.5(b). Puede observarse cómo el empleo de prior LDR en la deconvolución ofrece peores resultados que el empleo de prior HDR. En general, la deconvolución con NSR constante parece ofrecer bastante similares, aunque levemente superiores, a la deconvolución con prior HDR.

3.4.3. Modelo de procesado 3

Dado que las exposiciones que se procesan con este modelo son de bajo rango dinámico, no tiene sentido emplear la deconvolución de Wiener con un prior HDR para recuperar las imágenes, de modo que se omite este tipo de deconvolución para este modelo de procesado.

³<http://taniapouli.co.uk/research/statistics/>

La calidad Q media de las siete imágenes, obtenida con la métrica $HDR-VDP2$ [12], para cada tipo de deconvolución y nivel de ruido se muestra en la Figura 3.5(c). Puede observarse cómo empleando NSR constante se obtienen resultados ligeramente mejores, aunque muy similares.

3.4.4. Comparativa de los modelos

Observando la figura 3.5 puede verse cómo el procesado 1 ofrece globalmente mejores resultados que los otros dos. En cuanto al procesado 2, aunque para nivel de ruido alto parece ofrecer resultados de calidad similar a los otros dos, conforme disminuye el ruido se comporta mucho peor.

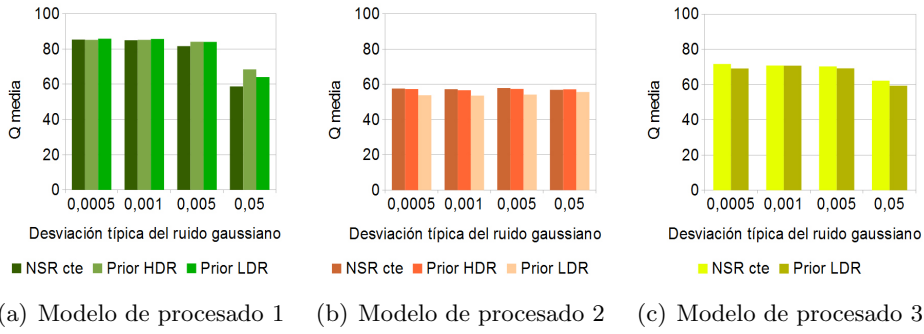


Figura 3.5: Q media obtenida con cada tipo de deconvolución (con NSR constante, con prior HDR y con prior LDR) para los distintos niveles de ruido, empleando los tres modelos de procesado diferentes.

3.4.5. Comparativa de los priors

A la vista de los resultados vistos en la Figura 3.5 se entiende que el prior HDR ofrece ventajas sobre el prior LDR a la hora de realizar la deconvolución de una imagen HDR (procesados 1 y 2). En la Figura 3.6 se observan los resultados de una de las imágenes recuperada con los dos priors distintos con el ruido de desviación típica $\sigma = 0,0005$. Se puede comprobar cómo la recuperada con prior HDR ofrece un resultado más favorable que la recuperada con deconvolución con prior LDR, al reducir sensiblemente el *ringing*. Nótese que se trata de una imagen particularmente desfavorable y la distorsión presente en ambas imágenes es alta, pero en todo caso inferior al emplear prior HDR.

3.5. Validación física de los modelos de procesado

Como se ha comentado anteriormente, al no disponer de una cámara que capture imágenes de alto rango dinámico con un solo disparo, no se

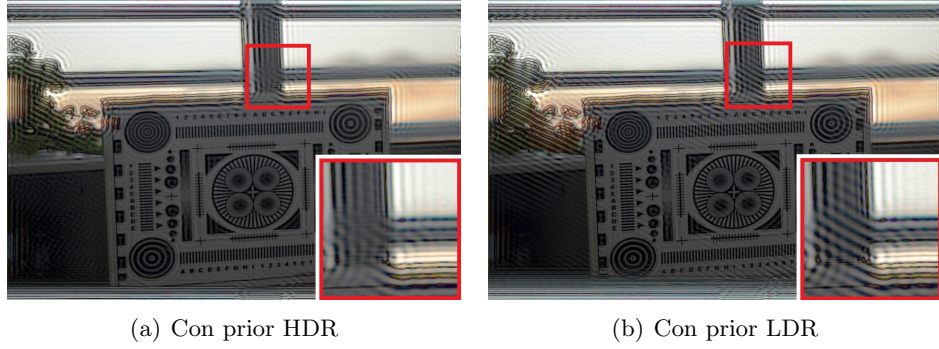


Figura 3.6: Ejemplo de imágenes HDR poco ruidosas, recuperadas los dos priors diferentes empleando el procesado 2. Nótese la reducción de *ringing* que se consigue al emplear un prior HDR en lugar de un prior LDR. A ambas imágenes se les ha aplicado reproducción de tono.

puede validar físicamente el modelo de procesado 1, por lo que, de ahora en adelante, el proyecto se centrará en los modelos 2 y 3, que en lo sucesivo y por comodidad se denominarán Procesado HDR y Procesado LDR respectivamente.

3.5.1. Captura de las imágenes desenfocadas con apertura codificada

Para evaluar estos métodos se capturan tres exposiciones distintas de una misma escena⁴. La escena está colocada a 60 cm del plano de enfoque, y éste a 120 cm de la cámara, como se muestra en la Figura 3.7. Para obtener las distintas exposiciones se emplea la opción de *multibracketing* de la cámara con EV relativos de -2, 0 y 2 stops. La misma apertura descrita en simulación [19] (Ver Figura 2.5) se imprime en papel de transparencia y se coloca en la lente del objetivo *Canon EF 50mm f/1.8 II* que se va a emplear. El procedimiento para insertar una apertura codificada en la lente está descrito en el Apéndice B. Para las capturas se emplea una cámara reflex *Canon EOS 500D*.

Además de estas exposiciones, se capturan las tres exposiciones correspondientes con la escena bien enfocada, a fin de tener una imagen de referencia de la escena.

⁴Parámetros de captura: $ISO = 100$; $F = 2,0$ stops; $texp_1 = 1/5$ s; $texp_2 = 1/20$ s; $texp_3 = 1/80$ s



Figura 3.7: Esquema de montaje para la captura de las exposiciones.

3.5.2. Obtención de las PSFs

Para poder recuperar las imágenes bien enfocadas a partir de las desenfocadas se debe conocer la PSF del sistema de captura (respuesta a un impulso), para poder emplearlo como kernel en la deconvolución. Para conseguir este impulso, empleamos un sencillo montaje con un LED colocado tras una cartulina negra a la que se ha aplicado un pequeño orificio, de modo que el efecto de la luz del LED al atravesar el agujero sea lo más puntual posible. Se coloca este montaje en el lugar de la escena, manteniendo fijo el plano de enfoque. Cabe pensar que existirán diferencias entre las PSF de las tres distintas exposiciones que se han capturado de la escena, de modo que se deben obtener una PSF para cada una de ellas. Para ello, en primer lugar se capturan tres imágenes del montaje con los mismos parámetros de exposición que se han empleado anteriormente.

Aunque existe una distorsión radial de la lente, que provoca una ligera variación de la PSF conforme se aleja del centro de la imagen, ésta se puede considerar despreciable, por lo que no se tiene en cuenta y se toman las capturas con la PSF centrada en la imagen. En la Figura 3.8 puede observarse el detalle central de las tres imágenes obtenidas con este procedimiento. Se puede comprobar cómo la forma de la respuesta al impulso es similar a la forma de la apertura, tal como se indica en el Apartado 2.3.2.

Una vez obtenidas estas imágenes, se recortan las zonas donde aparecen los patrones, se les aplica un umbral a cada uno de los píxeles que los conforman⁵, de modo que a aquellos que estén por debajo de éste se les da un valor cero, y se obtienen las PSF normalizando los resultados por la suma de todos sus píxeles.

También se obtiene una PSF para emplear en la deconvolución de la imagen HDR, fusionando las tres imágenes para obtener una fotografía del patrón de alto rango dinámico. Ésta se procesa de la misma manera que

⁵Umbral de 0.39, 0.5 y 0.8 para las imágenes subexpuesta, intermedia y sobreexpuesta, respectivamente. Nótese que el umbral a emplear aumenta al aumentar la exposición debido a la mayor luminosidad de la imagen.

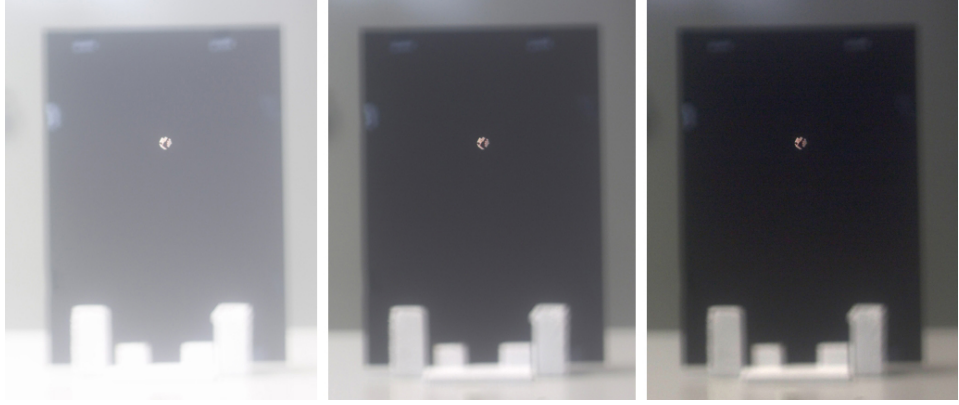


Figura 3.8: Detalle central de las imágenes empleadas para obtener las PSFs. De izquierda a derecha: imagen sobreexpuesta, exposición central e imagen subexpuesta

se indica en el párrafo anterior⁶. Tanto las exposiciones como las imágenes empleadas para recuperar las PSFs están tomadas en formato RAW con un tamaño de 4752x3168, pero por motivos de coste computacional se les aplica un *resize* de factor 0.2, hasta un tamaño de 951x634. En la Figura 3.9 pueden verse las cuatro PSFs recuperadas tras el *resize*. Su tamaño aproximado es de unos 14x14 píxeles.

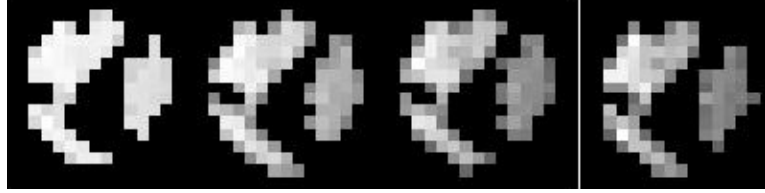


Figura 3.9: PSFs recuperadas tras el *resize* para la deconvolución. De izquierda a derecha: PSF recuperadas de la imagen sobreexpuesta, de la exposición central y de la imagen subexpuesta, y PSF recuperada de la imagen HDR.

3.5.3. Recuperación de las imágenes bien enfocadas

Una vez obtenidas las PSFs se recuperan las imágenes bien enfocadas empleando los dos tipos de procesado. Para el procesado HDR se fusionan las tres exposiciones desenfocadas y se trata de recuperar la imagen bien enfocada a partir de ella con una sola deconvolución (como se muestra en el esquema de la Figura 3.2), empleando la PSF obtenida especialmente para HDR. Para ello se emplea la deconvolución de Wiener sin prior de imágenes, con prior HDR y con prior LDR.

⁶El umbral empleado en este caso es de 0.2.

Para el procesado LDR se trata de recuperar las exposiciones bien enfocadas mediante una deconvolución para cada una de ellas (como se muestra en el esquema de la Figura 3.3), empleando en cada caso la PSF correspondiente a su exposición. Se prueban la deconvolución de Wiener sin prior y con prior LDR. Finalmente se fusionan las exposiciones recuperadas para obtener la imagen HDR resultante.

Se utiliza la métrica *HDR-VDP2*[12] para obtener la medida de calidad Q en comparación con la imagen de referencia. Esta medida puede observarse en la Figura 3.10.

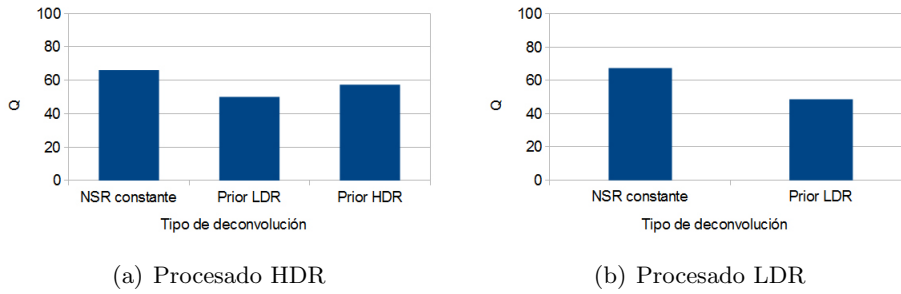
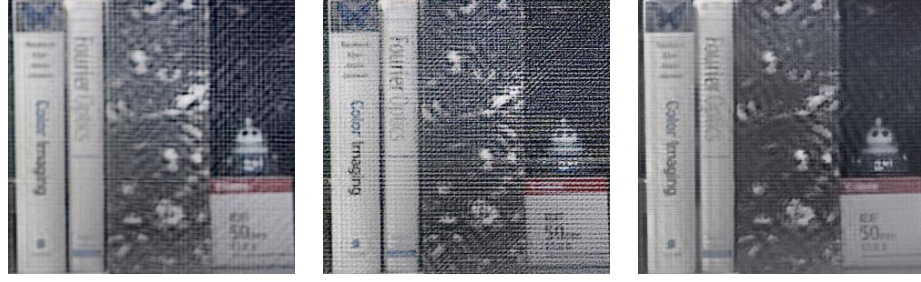


Figura 3.10: Q obtenida en el escenario real con la métrica *HDR-VDP2* para cada tipo de procesado y deconvolución. Aunque las simulaciones indican que el procesado LDR se comporta mucho mejor que el HDR, en la realidad ofrecen resultados similares.

Lo más llamativo es que, aunque las simulaciones indican que el procesado LDR se comporta mucho mejor que el procesado HDR, en realidad ofrecen resultados de calidad muy semejante. Además, mientras que las simulaciones indican que los resultados son bastante similares tanto empleando priors como sin ellos, en realidad ambos procesados se comportan mucho mejor cuando no se usa ningún prior en la deconvolución.

Al observar detenidamente las imágenes obtenidas con las deconvoluciones con prior se aprecia una clara distorsión en forma de rejilla en todas ellas, que no aparece en la simulación, como se muestra en la Figura 3.11.

Esta distorsión reduce sensiblemente la calidad visual de las imágenes obtenidas tras deconvolución con prior. No obstante puede verse cómo, para el procesado HDR, la imagen recuperada con el prior HDR 3.11(a) ofrece mejor resultado que la recuperada con prior LDR 3.11(b), por lo que se confirma que el empleo de un prior de imágenes HDR en la deconvolución de una imagen HDR ofrece resultados de una calidad superior que empleando un prior de imágenes LDR.



(a) Procesado HDR con prior HDR (b) Procesado HDR con prior LDR (c) Procesado LDR con prior LDR

Figura 3.11: Detalle de las imágenes recuperadas empleando prior en la deconvolución. Como puede verse, aparece una distorsión en forma de rejilla que reduce la calidad de las imágenes.

En la Figura 3.12 puede observarse el resultado de los dos mejores procesados reales en términos de Q , correspondientes a ambos modelos de procesado (HDR y LDR) empleando deconvolución sin prior. Junto a ellos se presenta la imagen original HDR desenfocada, obtenida mediante la fusión de las tres exposiciones desenfocadas, y la imagen HDR de referencia perfectamente enfocada. Estas imágenes son la muestra de que el empleo de aperturas codificadas para la captura de imágenes HDR es viable, como se quería demostrar. Observando las partes aumentadas se ve claramente cómo los detalles que en la original son imposibles de distinguir, en las recuperadas aparecen mucho más nítidos.

3.6. Obtención de patrones óptimos

Tras comprobar que el empleo de aperturas codificadas es viable en imágenes HDR, debido a que la apertura que se ha empleado hasta ahora [19] está diseñada para ser usada con imágenes LDR se va a tratar de obtener patrones de aperturas óptimos para imágenes de alto rango dinámico. Para ello se sigue el método empleado por Zhou y Nayar [19], basado en un algoritmo genético y en el empleo de priors de imágenes naturales. Se obtendrán una serie de aperturas empleando el prior de imágenes HDR descrito en el Apartado 3.3.1. Además, para poder realizar una comparativa rigurosa, también se obtendrán aperturas empleando el prior de imágenes LDR (ver 3.3.1).

Una vez obtenidos los patrones, éstos se emplean para realizar una serie de simulaciones como las detalladas en el Apartado 3.2, siguiendo los esquemas de procesado HDR (Figura 3.2) y LDR (Figura 3.3). A los resultados obtenidos se les aplica nuevamente la métrica $HDR-VDP2$ [12] para evaluar

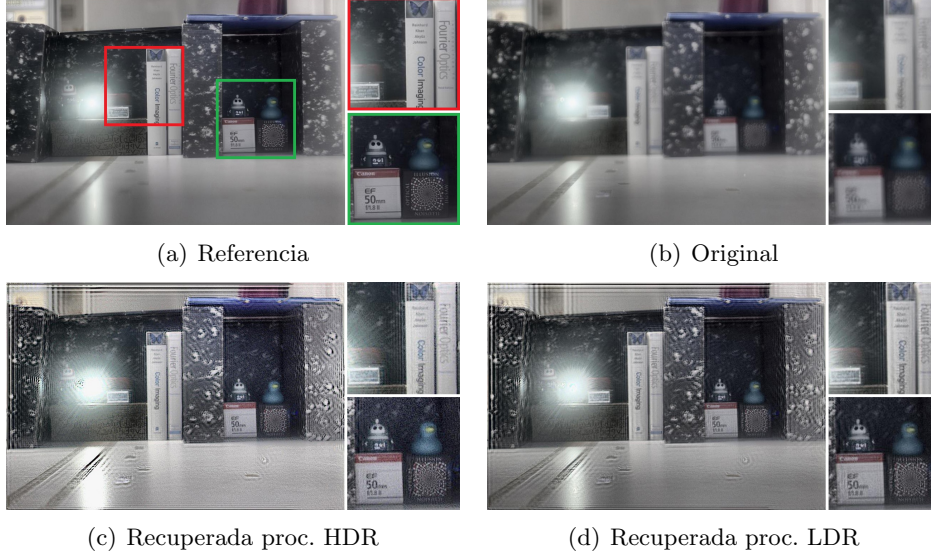


Figura 3.12: Resultados de los mejores procesados HDR y LDR en términos de Q: empleando deconvolución de Wiener sin prior. Aunque en la imagen original desenfocada es imposible distinguir ciertos detalles, éstos se aprecian correctamente en las imágenes recuperadas, tanto con procesado HDR como con procesado LDR, a costa de un ligero *ringing*.

su calidad, a fin de establecer una comparativa entre los patrones obtenidos con el prior HDR y el prior LDR.

3.6.1. Función objetivo

Para obtener patrones óptimos es necesario conocer una función que evalúe la calidad de los mismos para guiar la optimización. Zhou [19] propone una métrica que permite evaluar, dado un patrón K , la degradación de la imagen recuperada en función de la desviación típica σ del ruido. Viene dada por la siguiente ecuación:

$$R(K) = \sum_{\xi} \frac{\sigma^2}{|K_{\xi}|^2 + \sigma^2/A_{\xi}}, \quad (3.6)$$

donde la matriz A es un prior que promedia la potencia espectral de una serie de imágenes naturales, como los que se han empleado en apartados anteriores.

Una apertura K será tanto mejor cuanto más pequeño sea el valor de $R(K)$. La métrica se obtiene a partir del método de deconvolución de Wiener, cuya formulación, como ya se ha comentado, responde a la ecuación:

$$\hat{F}_0 = \frac{F \cdot \bar{K}}{|K|^2 + |C|^2},$$

donde \hat{F}_0 es la imagen recuperada, \bar{K} es el complejo conjugado de K , $|K|^2 = K \cdot \bar{K}$ y $|C|^2 = |\sigma/F_0|^2$ es la matriz de relación ruido a señal NSR (*Noise to Signal Ratio*) de la imagen original. Esta matriz nos es desconocida, por lo que se sustituye por el prior antes mencionado $|C|^2 = |\sigma/A|^2$.

Para obtener una apertura óptima es necesario resolver el problema de optimización dado por la Ecuación 3.6. Para un patrón de apertura dado por una matriz binaria de tamaño $N \times N$ el número de soluciones posibles es de $2^{N \times N}$, por lo que se propone el uso de un algoritmo genético que emplea la Ecuación 3.6 como función de evaluación.

3.6.2. Algoritmo genético

El algoritmo genético es un algoritmo de optimización iterativo inspirado en la evolución biológica, y que funciona de la siguiente forma:

1. Inicialización: Se genera aleatoriamente una población inicial de S secuencias binarias de longitud L , siendo $L = N \times N$ el tamaño de las aperturas a generar.
2. Selección: Se evalúa cada una de las S secuencias empleando la métrica ya comentada:

$$R(K) = \sum_{\xi} \frac{\sigma^2}{|K_{\xi}|^2 + \sigma^2/A_{\xi}}$$

De ellas se seleccionan las M mejores, las que tienen menor $R(K)$, y se eliminan todas las demás.

3. Reproducción: a partir de las M secuencias seleccionadas, se obtienen $S - M$ nuevas secuencias para completar la población hasta S secuencias. Estas secuencias se obtienen de la siguiente forma:
 - Recombinación: se escogen dos de las M secuencias al azar y se intercambian sus bits con una probabilidad $c1$, obteniendo dos nuevas secuencias.
 - Mutación: para cada nueva secuencia generada mediante recombinación, se cambia el valor de todos sus bits con una probabilidad $c2$.
4. Repetición: se repiten los pasos 2 y 3 hasta que se alcanza la última generación G .
5. Finalización: se evalúan todas las secuencias restantes y la mejor de ellas se selecciona como salida del algoritmo.

Los parámetros de ejecución del algoritmo genético fueron los siguientes:

- $L=121$ (para un tamaño de aperturas de 11×11).
- Población inicial: $S=4000$.
- Población final: $M=400$.
- Probabilidad de recombinación: $c1=0.2$.
- Probabilidad de mutación: $c2=0.05$.
- Número de generaciones: $G=40$.

Se emplearon dos niveles de ruido diferentes, $\sigma = 0,0005$ y $\sigma = 0,001$ y se ejecutó el algoritmo tres veces por cada combinación de prior (HDR y LDR) y nivel de ruido, para un total de 12 aperturas.

El tiempo total de ejecución para la obtención de los 12 patrones fue de 54296.32 segundos (15 horas aproximadamente), dando una media de 4524.69 segundos por patrón (1 hora y cuarto)⁷.

En la Figura 3.13 se muestran los patrones de las aperturas obtenidas.

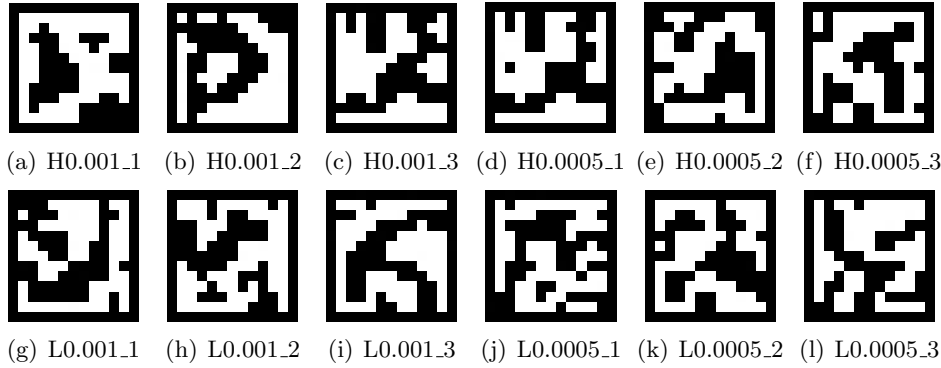


Figura 3.13: Patrones de aperturas obtenidos. Nótese que el tamaño de los patrones es de 11×11 píxeles, pero se les ha añadido un marco negro para una mejor visualización. Notación de las aperturas: $PRIOR\sigma_i$ (Prior empleado: $PRIOR=L(DR), H(DR)$; σ empleada: $\sigma=0.001, 0.0005$; Número de ejecución del algoritmo para cada combinación: $i=1, 2, 3$).

3.6.3. Evaluación de los patrones obtenidos

Para evaluar los 12 patrones que se han obtenido tras la ejecución del algoritmo genético se simula para cada uno de ellos el proceso de captura⁸ y la posterior recuperación con una serie de imágenes HDR y se emplea la

⁷Procesador: Intel Core i7-950, 3.06GHz. Memoria RAM: 6GB

⁸La σ empleada en la simulación de cada apertura es la misma que la que se usó en su obtención con el algoritmo genético.

métrica *HDR-VDP2*[12] para conocer la calidad de los resultados obtenidos. Se realizan estas simulaciones con los tipos de procesamiento HDR y LDR definidos en el Apartado 3.2. Finalmente se obtiene la media aritmética de los datos obtenidos para cada apertura. Además se evalúa de la misma forma la apertura de Zhou [19], la misma que ya se empleó anteriormente, a fin de comparar la calidad de nuestros patrones con uno válido ya conocido. Para esta apertura se realizan las simulaciones con los dos niveles de ruido empleados en el algoritmo genético. En la Figura 3.14 se muestra el valor medio de la medida de calidad Q obtenida para cada una de las aperturas, para los dos tipos de procesamiento.

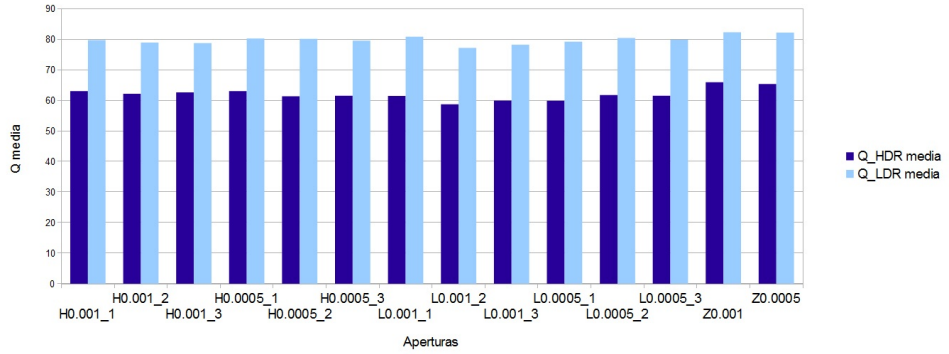


Figura 3.14: Q media obtenida para cada apertura, con procesamiento HDR (Q_{HDR}) y procesamiento LDR (Q_{LDR}). En general, las aperturas obtenidas con el prior HDR se comportan mejor que las obtenidas con prior LDR al realizar el procesamiento HDR. Z0.001 y Z0.0005 se corresponden con la apertura de Zhou empleando ruidos con $\sigma = 0,001$ y $\sigma = 0,0005$ respectivamente.

Puede observarse cómo, en general, las aperturas obtenidas con el prior de imágenes HDR se comportan mejor que las obtenidas con el prior LDR al realizar el procesamiento HDR, aunque las aperturas L0.001.1 y L0.0005.2 ofrecen resultados comparables a las primeras. En cuanto al procesamiento LDR, se observa mayor disparidad, encontrando aperturas buenas y malas indistintamente para ambos priors.

Además de esta medida de calidad de las aperturas, se debe comprobar su respuesta frecuencial para asegurar que resuelven los problemas de la apertura circular descritos anteriormente. En la Figura 3.15 se muestra la representación de los cortes centrales de las respuestas frecuenciales de las aperturas obtenidas con el prior HDR en comparación con la apertura circular. La misma representación, esta vez para las obtenidas con el prior LDR, se muestra en la Figura 3.16.

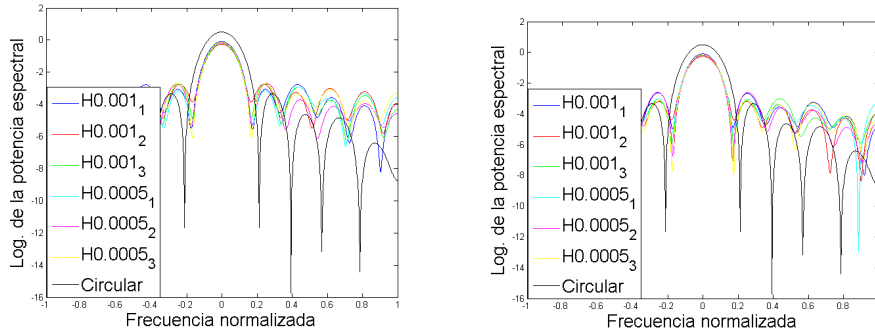


Figura 3.15: Logaritmo de la potencia espectral de las aperturas obtenidas con prior HDR frente a la apertura circular. Todos los patrones ofrecen respuestas frecuenciales más favorables al reducir la atenuación en las altas frecuencias y eliminar los cruces por cero.

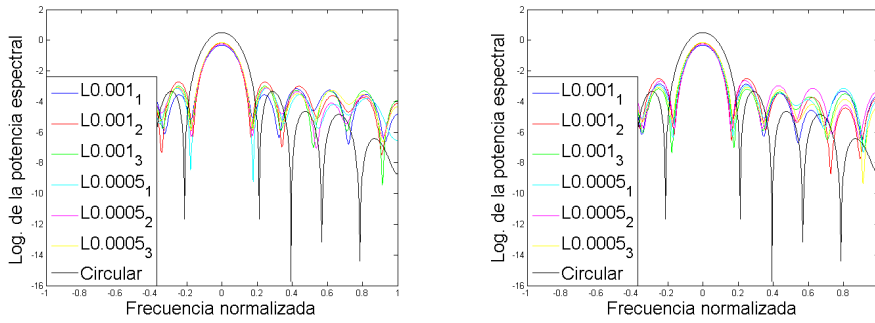


Figura 3.16: Logaritmo de la potencia espectral de las aperturas obtenidas con prior LDR frente a la apertura circular. Todas las aperturas eliminan los cruces por cero presentes en la circular, e introducen una menor atenuación en las altas frecuencias.

Puede comprobarse que todas las aperturas obtenidas tienen una respuesta frecuencial más favorable que la circular, al evitar los cruces por cero e introducir una menor atenuación en las frecuencias más altas.

3.7. Validación física de los patrones obtenidos

Una vez evaluada la calidad de los patrones se escoge el mejor de los obtenidos con cada prior para realizar experimentos con las aperturas físicas. Concretamente se eligen las aperturas H0.001_1 y L0.001_1 por ofrecer un buen comportamiento para ambos tipos de procesado.

Se imprimen ambas aperturas junto con la apertura de Zhou [19] (que ya ha sido validada en el Apartado 3.4) en papel de transparencia, como se muestra en la Figura 3.17. Con cada una de estas aperturas y con la apertura

circular se repite el proceso de captura ya descrito en el apartado 3.4. Se llevan a cabo nuevamente los procesados HDR y LDR, esta vez empleando tan sólo la deconvolución de Wiener sin ningún prior de imágenes, ya que de esta manera se obtienen mejores resultados, tal como se ha visto antes, y se obtiene una vez más la métrica $HDR-VDP2$ [12]. Igualmente, se simulan los mismos procesos para comparar los resultados reales con los de simulación.

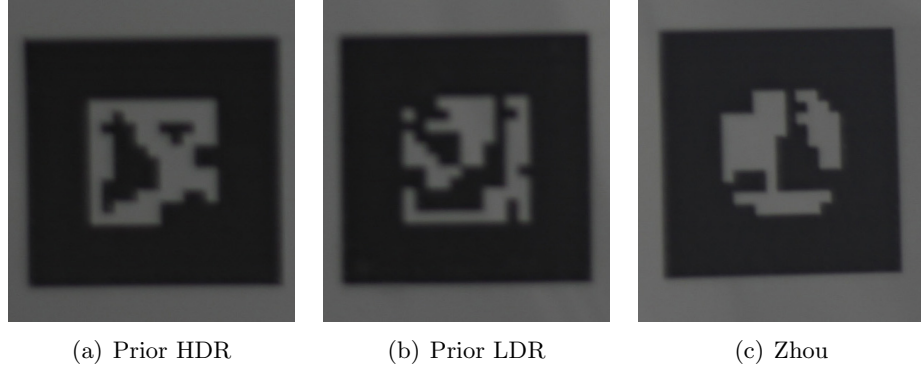


Figura 3.17: Aperturas impresas en papel de transparencia.

3.8. Resultados de la validación física

En la Figura 3.18(a) se puede ver la imagen HDR de referencia de la escena bien enfocada. Pueden distinguirse dos zonas desafiantes, la zona de la carta de ajuste fotográfico a la izquierda y la zona de los libros a la derecha. Como ya se ha comentado, la apertura circular presenta cruces por cero y atenuaciones en las frecuencias altas. En una imagen las frecuencias altas se corresponden con gradientes altos entre píxeles vecinos, ya sean bordes, cambios de color o de intensidad.

En las dos zonas señaladas se producen este tipo de gradientes, a la izquierda, la carta de ajuste presenta muchas variaciones de blanco a negro, y a la derecha, en la zona de los libros, se puede apreciar una clara diferencia de exposición entre el pato amarillo, que está recibiendo una luz directa, y el fondo, que está en penumbra. En las Figuras 3.18(c) y 3.18(d) pueden apreciarse los resultados obtenidos con la apertura circular. Se observa que en las zonas de la imagen donde se producen estos gradientes aparece una fuerte distorsión en forma de *ringing*. Esto se debe, como se ha comentado, a la respuesta frecuencial desfavorable de la apertura circular.

Los resultados para el detalle derecho con cada apertura testeada se muestran en la Figura 3.19, mientras los del detalle de la izquierda se mues-

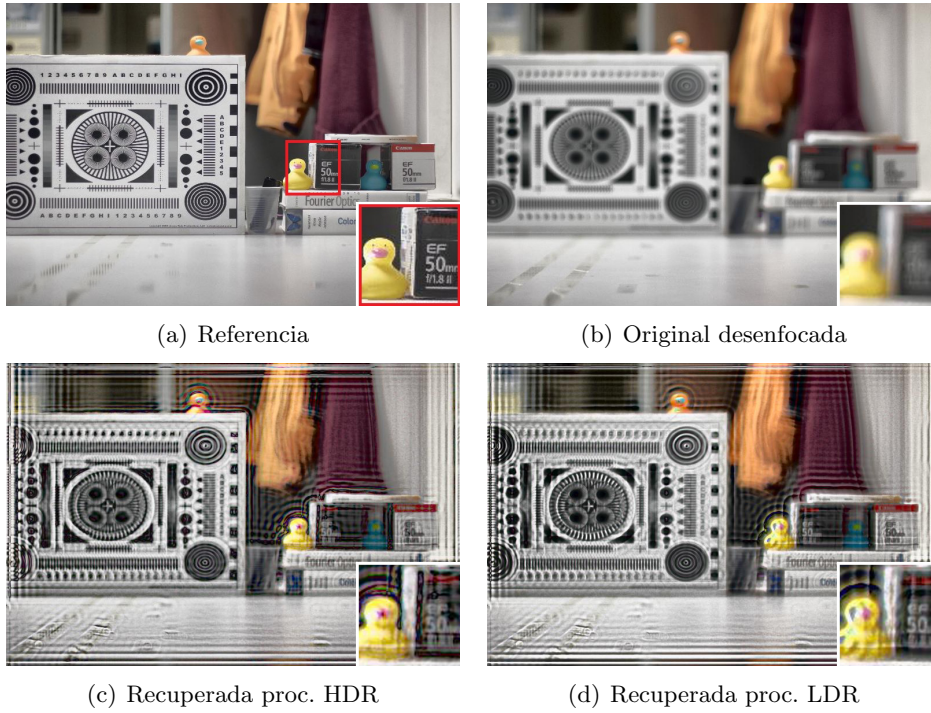
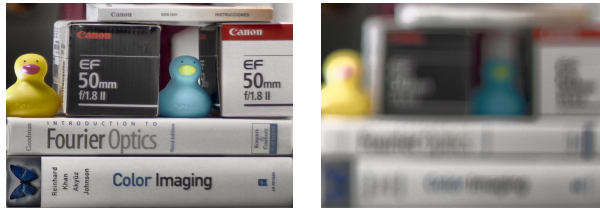


Figura 3.18: Resultados con la apertura circular. Todas las imágenes se muestran tras una reproducción de tono. Nótese la incapacidad de recuperar correctamente las imágenes enfocadas al emplear una apertura circular.

tran en la Figura 3.20. Además, en la Figuras 3.21(a) y 3.21(b) se detallan los resultados obtenidos al aplicar localmente la métrica $HDR-VDP2$ [12] en ambos detalles.

A la vista de las métricas se confirma una vez más, como ya se indica en el Apartado 3.5, que aunque la simulación del procesado LDR ofrece unos resultados mucho mejores que la del procesado HDR, en los experimentos reales ofrecen resultados de calidad muy similar.

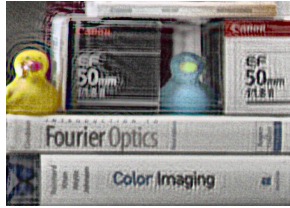
Observando detenidamente las imágenes obtenidas con uno y otro procesado, se puede comprobar cómo, empleando un procesado HDR, aparece un *ringing* mayor que el que aparece al emplear el procesado LDR. Este problema está estrechamente ligado al rango dinámico de la imagen. Como ya se ha comentado, este *ringing* aparece en zonas donde las diferencias entre píxeles vecinos son muy acentuadas. Una imagen de alto rango dinámico representa la imagen con más bits de profundidad, por lo que estas diferencias entre píxeles vecinos correspondientes a bordes o a gradientes de intensidad se acentúan. Dado que en el procesado HDR se está realizando la deconvolución sobre una imagen HDR, el *ringing* que aparece es mayor, mientras



(a) Ideal

(b) Original

PROCESADO HDR



(c) Apertura Zhou

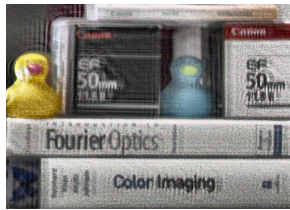


(d) Apertura con prior HDR

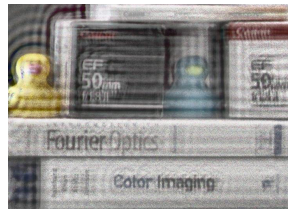


(e) Apertura con prior LDR

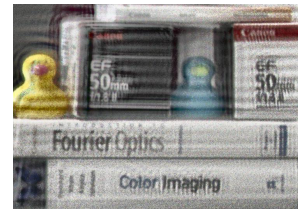
PROCESADO LDR



(f) Apertura Zhou



(g) Apertura con prior HDR



(h) Apertura con prior LDR

Figura 3.19: Resultados de ambos procesados para cada una de las aperturas, comparadas con la imagen original y la ideal, para el detalle de la derecha. Todas las imágenes se muestran tras una reproducción de tono. La imagen original que se muestra es la capturada con la apertura circular. Por simplicidad y similitud, no se muestran las originales de las demás aperturas.

que en el procesado LDR, al estar deconvolucionando las exposiciones de bajo rango dinámico el *ringing* que aparece es menor.

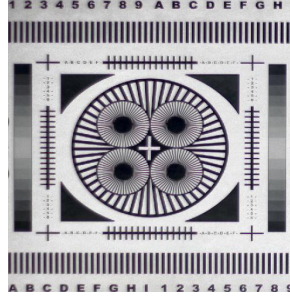
Sin embargo, también puede comprobarse cómo el procesado LDR genera resultados más ruidosos que el procesado HDR, motivo por el cual la medida de calidad que ofrece la métrica *HDR-VDP2*[12] para ambos tipos de procesado es similar. Esto se debe a que la deconvolución es un proceso ruidoso per se, y el proceso de fusión de varias exposiciones LDR en una imagen HDR amplifica el ruido de éstas.

El motivo por el cual el procesado LDR se comporta mejor que el HDR

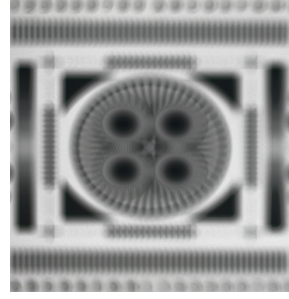
en simulación es, probablemente, que el que el modelado del ruido que se está haciendo no es del todo realista, ya que, aunque generalmente se suele considerar válida la aproximación gaussiana, en la realidad entran en juego otros muchos factores. Las simulaciones del procesado HDR ofrecen resultados más realistas, ya que los gradientes altos entre píxeles vecinos presentes en la realidad, y que son la mayor causa de distorsión en el resultado final, también aparecen en la simulación, reduciendo la calidad ofrecida por la métrica.

En cuanto a la comparación de las tres aperturas, como era de esperar, en el procesado LDR se comportan mejor aquellas que están optimizadas para imágenes LDR: la de Zhou [19] y la que se ha obtenido con el algoritmo genético con el prior LDR descrito anteriormente. Sin embargo, vemos que en el procesado HDR, la apertura obtenida con el prior HDR no parece comportarse mejor que las otras dos, ya que aunque en el detalle de la izquierda sí que ofrece un mejor resultado, en el de la izquierda no es así. No obstante, también es cierto que la apertura elegida de entre las obtenidas con el prior LDR ya indicaba en las simulaciones un comportamiento válido con el procesado HDR, como se vio en la Figura 3.14.

Finalmente, cabe destacar que ninguna de las aperturas obtenidas mediante el algoritmo genético del Apartado 3.5, con ninguno de los dos priors, supera a la apertura de Zhou [19] para ninguno de los dos procesados. Es importante también señalar, en relación con esto, que la métrica que se ha empleado en el algoritmo genético está basada en la simulación de la captura de una imagen desenfocada (convolución entre la imagen enfocada y el kernel, y ruido gaussiano aditivo) y su posterior deconvolución con el método de Wiener. Ya se explicó en el Apartado 3.2 que esta aproximación es estrictamente cierta para el modelo de procesado 1, mientras que para los dos modelos aquí empleados sólo es una aproximación.

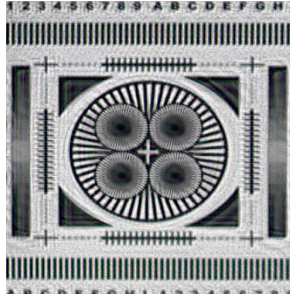


(a) Ideal

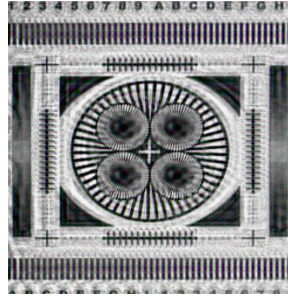


(b) Original

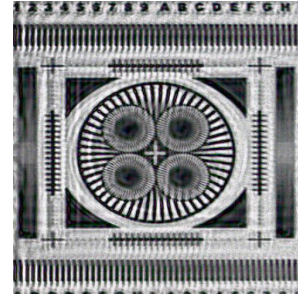
PROCESADO HDR



(c) Apertura Zhou

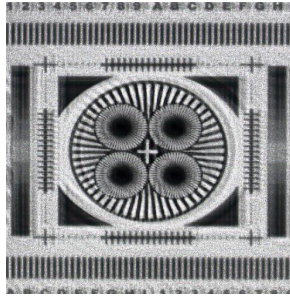


(d) Apertura con prior HDR

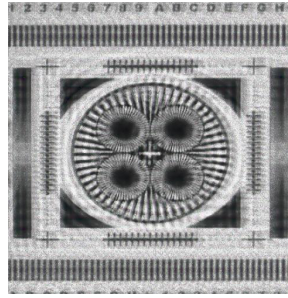


(e) Apertura con prior LDR

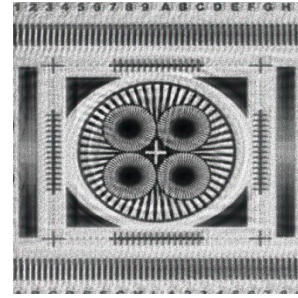
PROCESADO LDR



(f) Apertura Zhou

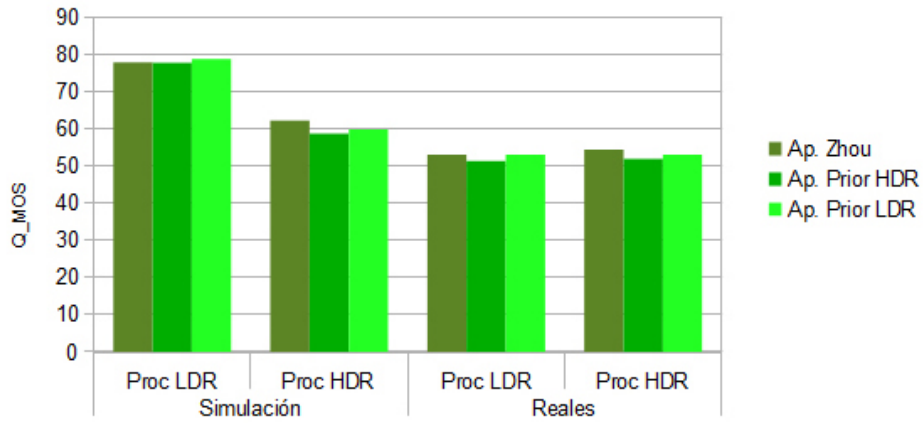


(g) Apertura con prior HDR

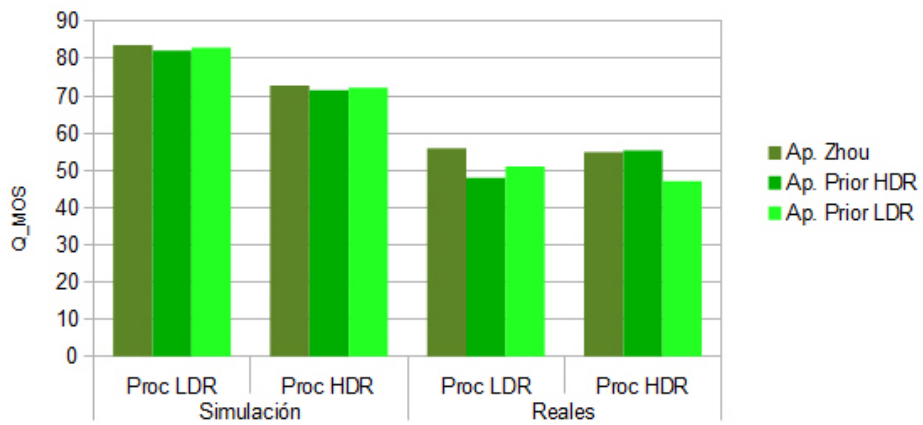


(h) Apertura con prior LDR

Figura 3.20: Resultados del ambos procesados para cada una de las aperturas, comparadas con la imagen original y la ideal, para el detalle de la izquierda. Todas las imágenes se muestran tras una reproducción de tono. La imagen original que se muestra es la capturada con la apertura circular. Por simplicidad y similitud, no se muestran las originales de las demás aperturas.



(a) Detalle de la derecha



(b) Detalle de la izquierda

Figura 3.21: Métricas obtenidas para ambos detalles con cada una de las aperturas, tanto las reales como las de simulación, empleando *HDR-VDP2*. Se comprueba una vez más como ambos modelos de procesamiento ofrecen resultados similares con experimentos reales a pesar de que las simulaciones indican lo contrario. Además, ninguna de las aperturas obtenidas supera en calidad a la apertura de Zhou.

Capítulo 4

Aperturas múltiples

4.1. Introducción

En esta parte del proyecto se estudia la técnica de generación y reenfoque de imágenes HDR a partir de distintas exposiciones capturadas con tamaño de apertura variable. Esta es una técnica ya existente, desarrollada por Hasinoff *et al.* [9, 10], y de utilidad probada, por lo que el objetivo se limita a implementarla, aunque de forma simplificada, y estudiar sus propiedades.

En primer lugar, se muestra un modelo de formación de imágenes en la cámara a partir de una división en capas de profundidad de la escena y de la luminancia de la misma. Este modelo se emplea para plantear un problema que puede resolverse con métodos de optimización para recuperar los parámetros de la imagen: luminancia HDR perfectamente enfocada y tamaño del desenfoque de cada capa.

Una vez que se conoce la función objetivo del problema, se comprueba su funcionamiento con imágenes sintetizadas manualmente a partir del modelo de formación de imágenes mencionado anteriormente. Finalmente, se valida el funcionamiento del sistema con imágenes reales.

4.2. Fotografía HDR con apertura variable

Como ya se ha comentado anteriormente, a la hora de obtener una imagen de alto rango dinámico se necesita capturar previamente una serie de fotografías de la misma escena con distinta exposición. Habitualmente, esto se consigue variando la velocidad del obturador en cada una de las capturas, dejando fijos el resto de parámetros, de modo que el tiempo de exposición, y por tanto el rango de luminancias que quedarán bien representadas, será diferente en cada fotografía.

Esto puede conseguirse también variando el tamaño de la apertura en cada fotografía en lugar del tiempo de exposición, dejando fijo éste y el resto de parámetros, de modo que con un tamaño de apertura pequeño obtendremos una imagen en la que las zonas más oscuras de la escena estarán bien expuestas, mientras que con un tamaño de apertura grande serán las zonas más luminosas las que quedarán bien representadas.

Sin embargo, la variación de la apertura no sólo afecta a la exposición de la imagen, sino también a la profundidad de campo (*depth of field*), de forma que una apertura mayor tiene menor profundidad de campo que una más pequeña, como puede verse en la Figura 4.1. Ésto produce un efecto adverso ya que, si la profundidad de la escena que se quiere fotografiar es muy grande, las zonas muy alejadas del punto de enfoque aparecerán desenfocadas en las fotografías tomadas con las aperturas de mayor tamaño, lo cual resulta problemático a la hora de fusionarlas para obtener la imagen HDR. Por este motivo suele ser preferible capturar las distintas exposiciones variando la velocidad del obturador en lugar del tamaño de la apertura.



Figura 4.1: Variación de la profundidad de campo y de la exposición en función del tamaño de la apertura. Puede verse que para la apertura más pequeña la imagen presenta una exposición menor a pesar de emplear el mismo tiempo de exposición. También puede comprobarse cómo con la apertura más pequeña toda la imagen aparece bien enfocada al encontrarse dentro de la profundidad de campo, mientras que para la más grande sólo ocurre ésto con la parte central de la escena. (Adaptada de Hasinoff *et al* [7]).

No obstante, lo que a priori puede parecer un problema se convierte en una ventaja, ya que capturando las imágenes con apertura variable y mediante un procesamiento correcto se puede extraer información de luminancia y desenfoque de las imágenes capturadas. Esta información puede emplearse para crear imágenes de la escena con alto rango dinámico y que pueden ser reenfocadas post captura.

Esto se consigue gracias a un modelo de formación de imagen robusto a partir de las imágenes de entrada que se empleará para formular un problema de optimización que permite recuperar la luminancia totalmente enfocada de la escena y el tamaño del desenfoque en cada capa según el tamaño de la apertura.

4.3. Modelo de formación de imágenes

El proceso de formación de imágenes en la cámara, fijando la velocidad del obturador, depende del tamaño de la apertura, de la distancia de los objetos a la cámara y de la distancia focal. Por sí solo, el tamaño de la apertura determinará la exposición de la imagen, y en conjunción con los otros parámetros definirá el desenfoque.

4.3.1. Modelo de exposición

La exposición de la imagen, dejando fijo el tiempo de exposición, o lo que es lo mismo, la velocidad del obturador, dependerá exclusivamente del tamaño de la apertura. Así, se define un factor de exposición e_a que aumenta conforme aumenta el diámetro de la apertura, de modo que, para una escena L , la imagen \bar{L} formada en la cámara se podrá aproximar por $\bar{L} = e_a L$.

Además de esto, hay que tener en cuenta que cuando el valor de un píxel $\bar{L}(x, y) = e_a L(x, y)$ sobrepase al valor máximo permitido por el sensor dará lugar a un píxel sobreexpuesto, cuyo valor será precisamente este valor máximo. Para simular correctamente el proceso de captura hay que tener en cuenta esta sobreexposición. El modelo completo de exposición responde a la ecuación siguiente:

$$\bar{L}(x, y) = \min[e_a L(x, y), 1] \quad (4.1)$$

4.3.2. Modelo de desenfoque

Como se ha visto en el Apartado 2.3.1, el desenfoque (*blur*) de una imagen es mayor conforme más alejada está la imagen del plano focal, ya que el círculo de confusión aumenta. Suponiendo una escena situada en un plano paralelo al plano focal, siendo L la imagen perfectamente enfocada, puede expresarse la imagen resultante \bar{L} , para un tamaño de apertura dado, como:

$$\bar{L} = L * B_\sigma \quad (4.2)$$

Donde B_σ es la PSF del desenfoque, dependiente del llamado *blur diameter* o diámetro de blur σ , que viene dado por el diámetro de la apertura D , la

distancia focal d y la distancia de la cámara a la escena d' , siendo:

$$\sigma = \frac{|d' - d|}{d} D \quad (4.3)$$

En la práctica, se asume que B_σ es una función Gaussiana 2D, siendo σ la desviación típica de la PSF:

$$B_\sigma(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}} \quad (4.4)$$

En la Figura 4.2 puede verse una imagen en la que se ha simulado un desenfoque por profundidad con distintos diámetros de blur, siguiendo el modelo aquí descrito.



Figura 4.2: Simulación del desenfoque por profundidad con diferentes tamaños de blur.

4.3.3. Modelo de luminancia por capas

Como se ha visto en el apartado anterior, es sencillo simular el desenfoque en una escena situada en un sólo plano. Sin embargo, cuando la escena está situada en varios planos o capas a diferentes profundidades, el diámetro de blur para cada una de ellas será distinto, y probablemente existirán oclusiones.

Para solucionar esto, en primer lugar se divide la escena en una serie de K capas binarias empleando una máscara $A = \{A'_k\} (k = 1, 2, \dots, K)$ para seleccionar los píxeles no ocluidos en cada uno de los planos de profundidad. Esta máscara permite calcular por separado el desenfoque de cada una de las capas, que luego pueden sumarse para obtener la luminancia de la imagen desenfocada como se muestra en la siguiente ecuación:

$$\bar{L} = \sum_{k=1}^K [(A_k \cdot L) * B_{\sigma_k}] \quad (4.5)$$

Además, se define una segunda máscara $M = \{M_k\} (k = 1, 2, \dots, K)$ que representa la oclusión presente en cada capa y producida por las que se

encuentran situadas más cerca de la cámara. Para obtener esta máscara se emplea la ecuación que se detalla a continuación:

$$M_k = \prod_{j=k+1}^K (1 - A_j * B_{\sigma_j}) \quad (4.6)$$

Esta máscara se aplica sobre cada una de las capas después de ser desenfocadas individualmente y antes de sumarlas para obtener la luminancia total de la imagen. El proceso completo se muestra en la Figura 4.3, y responde a la siguiente ecuación:

$$\bar{L} = \sum_{k=1}^K [(A_k \cdot L) * B_{\sigma_a}] \cdot M_k \quad (4.7)$$

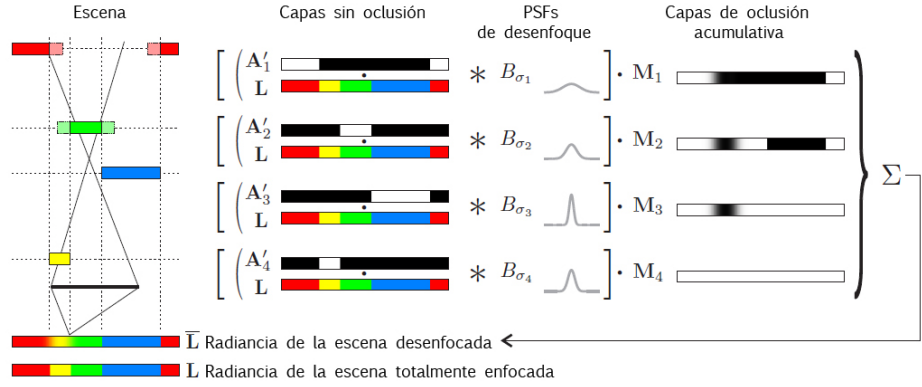


Figura 4.3: Esquema del modelo de creación de la luminancia de la imagen por capas de profundidad. La escena se divide en capas de profundidad y éstas son desenfocadas por separado. (Adaptada de Hasinoff *et al* [10])

4.3.4. Modelo completo de escena

Resumiendo los apartados anteriores, a partir de la luminancia de la escena perfectamente enfocada L , la segmentación de la misma en capas de profundidad A y el diámetro de blur por capa σ_a , se obtiene la imagen que se forma en la cámara como:

$$\bar{L} = \min \left[e_a \sum_{k=1}^K [(A_k \cdot L) * B_{\sigma_a}] \cdot M_k, 1 \right] \quad (4.8)$$

4.4. Optimización

Una vez descrito el modelo de formación de las imágenes en la cámara, éste se emplea para formular el problema inverso, mediante una función objetivo que permite, tras una optimización, recuperar la luminancia HDR perfectamente enfocada L de la escena y el tamaño de blur de cada capa σ_k , a partir de una serie de n fotografías tomadas con distinto tamaño de apertura y de la división en capas A de la escena.

4.4.1. Función objetivo

Definimos el problema como la estimación de las L y σ que mejor reproducen las imágenes de entrada, para la división en capas A . Así, la función objetivo queda como:

$$O(L, A, \sigma) = \frac{1}{2} \sum_{a=1}^n \|\Delta_a\|^2 \quad (4.9)$$

Donde Δ_a es el error residual existente entre las imágenes sintetizadas mediante la Ecuación 4.8, a partir de L , A y σ , y las imágenes de entrada I_a tomadas con cada una de las n aperturas:

$$\Delta_a = I_a - \bar{L}(L, A, \sigma) \quad (4.10)$$

Juntando las ecuaciones 4.8, 4.9 y 4.10 se obtiene la ecuación completa de la función objetivo:

$$O(L, A, \sigma) = \frac{1}{2} \sum_{a=1}^n \left\| I_a - \min \left[e_a \sum_{k=1}^K [(A_k \cdot L) * B_{\sigma_a}] \cdot M_k, 1 \right] \right\|^2 \quad (4.11)$$

4.4.2. Método de optimización

Para resolver el problema de optimización se utiliza la función *lsqnonlin*¹. de Matlab, por tratarse de un problema de mínimos cuadrados no lineales. Dado el tamaño del problema se le suministra al *solver* no sólo la función objetivo sino también la estructura de la matriz jacobiana asociada.

4.4.3. Inicialización

- **Inicialización de la luminancia de la escena:** La aproximación inicial de la luminancia de la escena se obtiene directamente de las imágenes de entrada, seleccionando sus píxeles y escalándolos por la inversa de su factor de exposición $1/e_a$. El objetivo es seleccionar el

¹<http://www.mathworks.es/help/toolbox/optim/ug/lsqnonlin.html>

mayor número de píxeles posible de la imagen obtenida con la apertura más pequeña, ya que es la que menos desenfoque presenta de todas. Sin embargo, no se pueden escoger los píxeles que presenten valores muy bajos de intensidad (correspondientes a zonas mas oscuras), ya que muy probablemente estarán dominados por el ruido de la imagen. Definimos un umbral $k=0.1$ de modo que para cada píxel se selecciona la apertura más pequeña para la cual la intensidad de la imagen está por encima del mismo, o la más grande de todas si no se cumple para ninguna de ellas.

- **Asignación de las capas:** Aunque el método original [9, 10] incluye un método automático de asignación de capas, por simplicidad en este proyecto se asignan las capas manualmente, como input del usuario.

4.5. Definición de los experimentos

Se realizan una serie de experimentos para comprobar la validez del método. En primer lugar se trata de recuperar la luminancia L a partir de un input de imágenes sintéticas, con división en capas y tamaño de blur conocidos, como se muestra en la Figura 4.4. A continuación se emplearán imágenes reales con luminancia y tamaño de blur por capas desconocido y se tratará de recuperar ambas a partir de las imágenes obtenidas con distintos tamaños de apertura y de una división por capas, como se ve en la Figura 4.5.

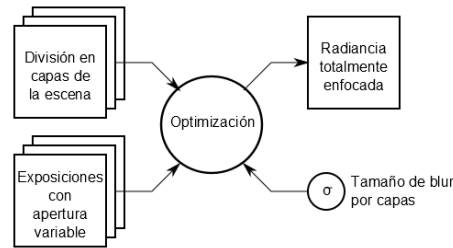


Figura 4.4: Esquema de la optimización en el experimento con imágenes sintéticas. Se obtiene la luminancia totalmente enfocada de la escena a partir de las exposiciones obtenidas con distinto tamaño de apertura, la división en capas de la escena y el tamaño de blur por capas.

4.5.1. Recuperación de luminancia L en imágenes sintéticas

Para crear la imagen sintética, se escoge una imagen LDR en color de tamaño 256×256 píxeles y se le simula el alto rango dinámico dividiéndola en tres bandas verticales y exponiéndolas artificialmente, para obtener la

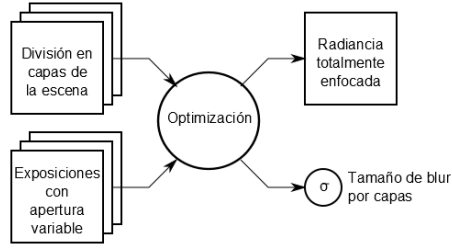


Figura 4.5: Esquema de la optimización en el experimento con imágenes reales. Se obtiene la luminancia totalmente enfocada de la escena y el tamaño de blur por capas a partir de las exposiciones obtenidas con distinto tamaño de apertura y la división en capas de la escena.

imagen L que se muestra en la Figura 4.6(a). Además, se divide la imagen en tres capas asignando diferentes profundidades en tres bandas horizontales (Figuras 4.6(b)-4.6(d)).

Finalmente, se aplica el modelo de formación de imagen para simular la captura de la escena con tres diferentes tamaños de apertura, situando el plano de enfoque en la capa central. Las imágenes que se obtienen se muestran en la Figura 4.7.

Puede observarse cómo la exposición y el desenfoque en las capas posterior y frontal aumentan con el tamaño de la apertura. Para obtener estas imágenes se ha empleado un diámetro de blur $\sigma_A = 3$ píxeles en las capas desenfocadas para la apertura mayor. El diámetro de blur σ_a para cada una de las otras dos aperturas se obtiene a partir del de la mayor y del factor de exposición de la apertura e_a como:

$$\sigma_a = \sigma_A \sqrt{\frac{e_a}{e_A}}, \quad (4.12)$$

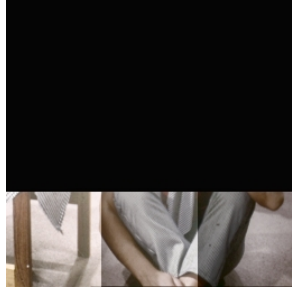
donde e_A es el factor de exposición correspondiente a la mayor de las aperturas. Se emplean factores de exposición de 1, 4 y 16 para la apertura más pequeña, la media y la más grande, respectivamente. A partir de estas tres imágenes se obtiene una inicialización de la luminancia como se ha indicado en el Apartado 4.4.3 y se procede a la optimización para recuperar la luminancia de la escena.

4.5.2. Recuperación de luminancia L y tamaño de blur en imágenes reales

Para validar el método con escenas reales, se monta un escenario compuesto por tres capas a distintas profundidades y se capturan tres imágenes



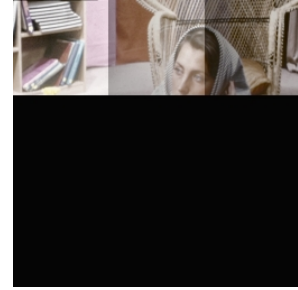
(a) Imagen sintética



(b) Capa posterior



(c) Capa central



(d) Capa frontal

Figura 4.6: Imagen sintética con exposición artificial para simular el rango dinámico y división en capas de profundidad. Las tres bandas verticales se exponen artificialmente, y las tres bandas horizontales se corresponden con capas de distinta profundidad. Todas las imágenes se muestran tras una reproducción de tono.

de la escena, con tamaños de apertura de $f/8$, $f/4$ y $f/2$ stops. Para el resto de parámetros de captura se escoge $ISO=100$ y tiempo de exposición de $1/20$ segundos. Las imágenes se capturan en formato RAW con un tamaño de 4752×3168 píxeles y se elimina la corrección gamma de las mismas para hacerlas lineales. Posteriormente se recortan y se les aplica un *resize* para obtener imágenes de 256×256 , con el objetivo de hacerlas más manejables y reducir el coste computacional. El factor de exposición para estas tres imágenes es aproximadamente de 1, 1.8 y 3.5 respectivamente.

Esta vez se desconoce el tamaño de blur de cada capa, por lo que se recuperará simultáneamente junto con la luminancia HDR enfocada de la escena, empleando las tres imágenes capturadas con los tamaños de apertura y la división en capas obtenida manualmente. Para ello se alterna la opti-

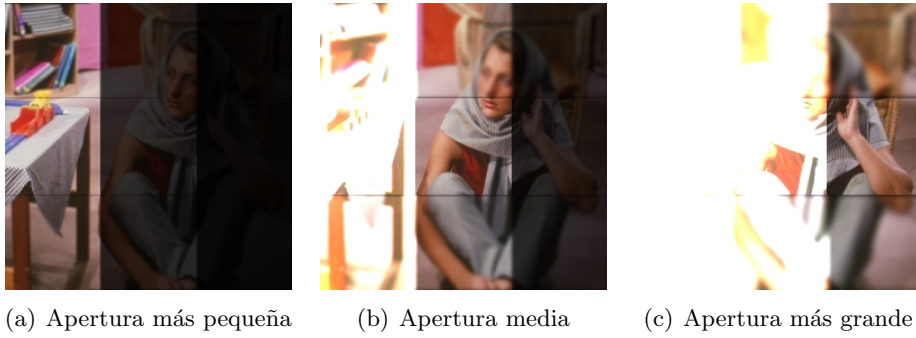


Figura 4.7: Simulación de captura con tres tamaños de apertura diferentes. Puede observarse una variación tanto en la exposición en toda la imagen como en el desenfoque de las capas frontal y posterior.

mización para ambos parámetros: en primer lugar se realizan 10 iteraciones dejando fijo el tamaño de blur inicial σ y optimizando para la luminancia L , y a continuación otras 10 dejando fija la luminancia L y optimizando para σ . Este proceso se repite hasta la convergencia.

4.6. Resultados

A continuación, se exponen y discuten los resultados de los experimentos propuestos en el apartado anterior.

4.6.1. Imágenes sintéticas

Empleando como entrada las imágenes sintéticas mostradas en la Figura 4.7, fijando el tamaño de blur al mismo valor empleado en su obtención ($\sigma = 3$ píxeles) y empleando la división en capas de la Figura 4.6, se ejecuta la optimización de la función objetivo. En la Figura 4.8 se muestran tanto la inicialización de la luminancia como la recuperada al final del proceso.

A la vista del resultado, se observa cómo el algoritmo ha sido capaz de recuperar a la perfección la luminancia de la escena sintética a partir de las tres imágenes de entrada (Figura 4.7) a pesar de los errores claros presentes en la inicialización. Aplicando la norma $L2$ a la luminancia recuperada con respecto a la luminancia sintética ideal se obtiene un resultado en tanto por ciento de $L2 = 7,4521e - 008$, lo que se traduce en un error insignificante.

4.6.2. Imágenes reales

Se prueba el algoritmo en un escenario real, tratando de recuperar simultáneamente la luminancia y el tamaño de blur σ por capa, como se indica



Figura 4.8: Resultado de la optimización para recuperar la luminancia de la imagen sintética. Ésta se obtiene correctamente y perfectamente enfocada, a pesar de una inicialización claramente distorsionada.

en el Apartado 4.4.2.

El algoritmo converge a partir de la séptima repetición, y emplea un tiempo de 35373 segundos (aproximadamente diez horas) de media en cada una de ellas². En la Figura 4.9 se muestra la luminancia resultante junto a las imágenes de entrada y la división en capas de la escena. Se aprecia claramente cómo el algoritmo es capaz de recuperar la escena totalmente enfocada, aunque aparece una leve distorsión en los bordes entre las capas debida a la oclusión. El método original resuelve este problema aplicando procedimientos de *inpainting*, pero por encontrarse fuera de los objetivos marcados se optó por no implementarlo.

El tamaño de blur recuperado tras la optimización para la imagen con apertura más grande es de 1.797 píxeles en la capa frontal, 0 en la capa intermedia (ya que está en foco) y 1.678 en la posterior. Nótese que conociendo los valores de σ se puede obtener una estimación de la profundidad relativa entre capas.

Una vez recuperada la luminancia totalmente enfocada y los tamaños de blur, aplicando el modelo de formación de imágenes se pueden manipular en postcaptura los parámetros de captura de la escena de la forma que se desee. En el CD adjunto a la memoria se incluye un vídeo en el cual se muestra una resíntesis completa de los parámetros de la cámara. En la Figura 4.10

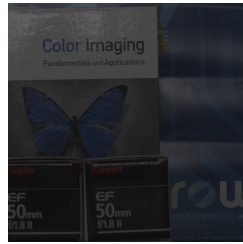
²Procesador: Intel Core i5, 2.26GHz. Memoria RAM: 6GB

se muestra la escena con todo su rango dinámico, reenfocada en cada una de las capas. También se puede manipular la exposición con toda la escena perfectamente enfocada, como se muestra en la Figura 4.11, o modificar al mismo tiempo el enfoque y la exposición como si se estuviera manipulando el tamaño de la apertura, tal como se muestra en la Figura 4.12.

Como puede verse en estos resultados, el sistema implementado permite recuperar correctamente la luminancia enfocada de la escena con ligeras distorsiones, así como el tamaño de blur correspondiente a cada capa, ofreciendo una gran versatilidad en postcaptura al permitir simular sintéticamente la variación de todos los parámetros de la cámara.



(a) Luminancia recuperada



(b) $f/8$



(c) $f/4$



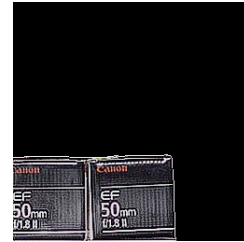
(d) $f/2$



(e) Capa posterior



(f) Capa intermedia



(g) Capa frontal

Figura 4.9: Luminancia recuperada de la imagen real, junto a las imágenes de entrada y la división en capas. El algoritmo permite recuperar la escena totalmente enfocada a costa de una ligera distorsión en los bordes entre capas debido a la oclusión.



(a) Capa posterior



(b) Capa central



(c) Capa frontal

Figura 4.10: Luminancia HDR recuperada reenfocada en postcaptura a cada una de las capas.



(a) Menor exposición



(b) Exposición intermedia



(c) Mayor exposición

Figura 4.11: Manipulación de la exposición en postcaptura con la escena totalmente enfocada.



(a) f8



(b) f4



(c) f2

Figura 4.12: Manipulación del tamaño de apertura en postcaptura.

Capítulo 5

Conclusiones

A continuación se exponen las conclusiones del trabajo realizado, así como el grado de consecución de los objetivos que se plantearon. Además se sugieren algunas líneas de investigación futuras que surgen a raíz del estudio realizado, y finalmente se exponen una serie de conclusiones personales del autor del proyecto.

5.1. Conclusiones del trabajo realizado

Se han estudiado dos técnicas distintas de reenfoque de imágenes HDR. La primera de ellas consiste en aplicar las técnicas conocidas de aperturas codificadas para corrección de desenfoques a imágenes de alto rango dinámico. El objetivo que se planteó respecto a este punto fue realizar un estudio exhaustivo del empleo de estas técnicas para comprobar cómo se comportaban en este tipo de imágenes. A la vista de los resultados obtenidos se puede concluir lo siguiente:

- Se ha demostrado la validez y la viabilidad del empleo de aperturas codificadas para corrección de desenfoques en imágenes HDR. Además se han propuesto tres modelos de procesamiento diferentes para conseguir reenfoque de una imagen HDR desenfocada y se ha establecido una comparación de la calidad que ofrece cada uno de ellos, si bien el primero sólo ha podido estudiarse en simulación al no contar con el material necesario para realizar experimentos reales.
- En cuanto a los otros dos modelos, llamados procesamiento HDR y procesamiento LDR, se ha comprobado que ambos ofrecen resultados de calidad similar, a pesar de que las simulaciones indican que el primero se comporta peor que el segundo. Además, el procesamiento HDR resulta menos costoso computacionalmente al requerir de una sola deconvolución, mientras el procesamiento LDR necesita una deconvolución por cada exposición.

- Se ha comprobado cómo a la hora de realizar una deconvolución de una imagen HDR se obtienen mejores resultados empleando un prior estadístico obtenido a partir de imágenes naturales HDR que con otro obtenido a partir de imágenes LDR. A pesar de ello, en los resultados aquí obtenidos no se aprecia una mejora en el empleo de priors en la deconvolución con respecto al empleo de NSR constante ya que los priors aquí empleados no son del todo óptimos, haciéndose necesaria una mejora en este campo.
- Se han obtenido aperturas codificadas para HDR empleando la métrica de Zhou [19] con priors de imágenes HDR, sin obtener una mejora de resultados con respecto a las aperturas obtenidas con priors LDR.

La segunda técnica estudiada consiste en la obtención de la luminancia HDR perfectamente enfocada de una escena, así como el grado de desenfoque existente en cada capa de profundidad, a partir de una serie de imágenes de la misma capturadas con distintos tamaños de apertura. De acuerdo con los resultados, se obtienen las siguientes conclusiones:

- La técnica implementada permite obtener correctamente la luminancia HDR enfocada de una escena, así como el tamaño de blur por capas de profundidad.
- Esta técnica ofrece una gran versatilidad, ya que a partir de la división en capas de la escena, de la luminancia HDR enfocada y del tamaño de blur por capas de profundidad se pueden obtener imágenes de la escena variando los parámetros de exposición y enfoque de la forma que se desee.
- A pesar de las ventajas que presenta, se trata de una técnica excesivamente costosa en tiempo de computación, al requerir de una optimización que precisa de varias horas para alcanzar la convergencia.

5.2. Trabajo futuro

A raíz de la investigación desarrollada se proponen algunas líneas de investigación futuras:

- Dado que no ha sido posible realizar experimentos reales con el primero de los procesados propuestos al no contar con el material necesario para ello, se propone un estudio más a fondo del mismo empleando cámaras fotográficas que permitan obtener imágenes HDR con una sola captura.

- Debido a la inexistencia de priors HDR válidos para la resolución de problemas HDR, basándose en el trabajo de Pouli *et al.* [16] y en los resultados aquí obtenidos, se propone un estudio a fondo sobre la construcción de priors basados en estadísticas de imágenes HDR.
- Aunque no se han logrado obtener aperturas óptimas para HDR, esto no implica que realmente no existan. La métrica que se ha empleado en el algoritmo genético está basada en la deconvolución de Wiener simulando la captura de una imagen desenfocada y su posterior deconvolución. Esta simulación no es válida del todo para HDR ya que no tiene en cuenta la fusión de las exposiciones. Así, se propone un estudio a fondo del problema para obtener una métrica más apropiada para imágenes HDR, que tenga en cuenta tanto la captura de las exposiciones como su fusión.
- En relación con el punto anterior y con el trabajo de Masiá *et al.* [14], se propone también la posibilidad de emplear métricas perceptuales en la obtención de patrones óptimos para HDR.

5.3. Conclusiones personales

Gracias a este proyecto he podido adentrarme en el mundo de la investigación y descubrir la exigencia de explorar un campo abierto y desconocido, y la satisfacción final al descubrir que el trabajo da resultado después de tanto tiempo y esfuerzo.

Durante este tiempo he tenido la oportunidad de trabajar dentro del Graphics and Imaging Lab del Grupo de Informática Gráfica Avanzada (GIGA), lo que me ha permitido experimentar desde dentro el día a día de un grupo de investigación de estas características, y atesorar una serie de valiosas experiencias y aprendizajes que me serán muy útiles el día de mañana.

Además, la investigación correspondiente a la primera parte del proyecto ha dado lugar a un artículo que ha sido enviado al Congreso Español de Informática Gráfica (CEIG 2012). Aunque a fecha de hoy todavía está pendiente de aceptación, escribirlo fue una nueva experiencia positiva.

Bibliografía

- [1] Aggarwal M., Ahuja N.: Split Aperture Imaging for High Dynamic Range. *Int. J. Comp. Vis.* (2004).
- [2] Akyüz A.O., Reinhard E.: Noise reduction in high dynamic range imaging. *JVCIR* (2007).
- [3] Caroli E., Stephen J., Cocco G., Natalucci L., Spizzichino A.: Coded aperture imaging in X- and Gammaray astronomy. *Space Science Reviews* (1987).
- [4] Debevec P. E., Malik J. Recovering high dynamic range radiance maps from photographs. *SIGGRAPH* (1997).
- [5] Gottesman S., Fenimore E.: New family of binary arrays new family of binary arrays for coded aperture imaging. *Applied Optics* (1989).
- [6] Grossberg M. D., Nayar S. K.: High Dynamic Range from Multiple Images: Which Exposures to Combine. In *Proc. ICCV Workshop CPMCV* (2003).
- [7] Hasinoff S.: Variable-Aperture Photography. PhD Thesis, Dept. of Computer Science, University of Toronto (2008).
- [8] Hasinoff S.W., Durand F., Freeman W.T.: Noise-Optimal Capture for High Dynamic Range Photography. In *Proc. IEEE CVPR* (2010).
- [9] Hasinoff S.W., Kutulakos K.N.: A Layer-Based Restoration Framework for Variable-Aperture Photography. *ICCV* (2007).
- [10] Hasinoff S.W., Kutulakos K.N.: Multiple-Aperture Photography for High Dynamic Range and Post-Capture Refocusing. *IEEE Trans. Pattern Analysis and Machine Intelligence* (submitted).
- [11] Levin A., Fergus R., Durand F., Freeman W.: Image and depth from a conventional camera with a coded aperture. *ACM Transactions on Graphics* (2007).

- [12] Mantiuk R., Kim K.J., Rempel A.G., Heidrich W.: HDR-VDP-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions. SIGGRAPH (2011).
- [13] Masiá B., Corrales A., Presa L., Gutiérrez D.: Coded apertures for defocus deblurring. Symposium Iberoamericano de Computación Gráfica (2011).
- [14] Masiá B., Presa L., Corrales A., Gutiérrez D.: Perceptually-Optimized Coded Apertures for Defocus Deblurring. Computer Graphics Forum (2012).
- [15] Nayar S.K., Branzoi V.: Adaptive Dynamic Range Imaging: Optical Control of Pixel Exposures over Space and Time. ICCV (2003).
- [16] Pouli T., Cunningham D., Reinhard E.: Statistical regularities in low and high dynamic range images. ACM Symposium on Applied Perception in Graphics and Visualization (APGV) (2010).
- [17] Veeraraghavan A., Raskar R., Agrawal A., Mohan A., Tumblin J.: Dappled photography: mask enhanced cameras for heterodyned light fields and coded aperture refocusing. ACM Trans. Graph. (2007).
- [18] Zhou C., Lin S., Nayar S.: Coded aperture pairs for depth from defocus. ICCV (2009).
- [19] Zhou C., Nayar S. K.: What are Good Apertures for Defocus Deblurring? IEEE International Conference on Computational Photography (2009).

Apéndice A

Software y procedimientos de fusión de imágenes HDR

En este apéndice se muestran algunos ejemplos de software comercial que permiten fusionar imágenes HDR a partir de una serie de exposiciones, así como el procedimiento a seguir en cada uno de ellos para realizar esta fusión.

A.1. Adobe Photoshop CS5

Precio: desde 1001.82 €(Actualización desde: 293.82 €)

Descarga: <http://www.adobe.com/es/products/photoshop.html>

Procedimiento de fusión de imágenes HDR:

1. Seleccionar: "Archivo → Automatizar → Combinar para HDR Pro..."
2. Hacer clic en "Explorar..." y seleccionar las exposiciones correspondientes. Si se desea que el programa trate de alinear las exposiciones automáticamente, seleccionar la opción "Intentar alinear automáticamente las imágenes de origen". A continuación, hacer clic en "OK".
3. Si la imagen está en formato RAW o en algún otro que conserve la información de exposición, saltar al paso siguiente. En caso contrario aparecerá una ventana para introducir los parámetros de exposición de cada fotografía. Escribirlos y hacer clic en "OK".
4. Seleccionar "Modo: 32 bits" y hacer clic en "OK".

A.2. Luminance HDR

Precio: Gratuito.

Descarga: <http://qtpfsgui.sourceforge.net/>

Procedimiento de fusión de imágenes HDR:

1. Hacer clic en "New HDR image".
2. Seleccionar "Cargar imágenes" y escoger las exposiciones correspondientes. Hacer clic en "Abrir".
3. Introducir el valor de exposición en EVs de cada fotografía y hacer clic en "Siguiente". Si la imagen está en un formato que conserve los parámetros de exposición no será necesario introducirlos.
4. Alinear las imágenes si es preciso y hacer clic en "Siguiente".
5. Seleccionar uno de los perfiles de creación que determinan la función de pesos, la curva de respuesta y el modelo de creación, o seleccionar cada uno por separado escogiendo la opción "Use custom configuration". Hacer clic en "Finalizar".

A.3. Photomatix Pro

Precio: desde 39 \$.

Descarga: <http://www.hdrsoft.com/download.html>

Procedimiento de fusión de imágenes HDR:

1. Hacer clic en "Load Bracketed Photos".
2. Hacer clic en "Browse" y escoger las exposiciones correspondientes.
3. Seleccionar la diferencia de EVs entre imágenes consecutivas con "Specify the E.V. spacing" o introducir manualmente el valor de EV de cada fotografía. Hacer clic en "OK".
4. Si se desea que el programa trate de alinear las imágenes automáticamente, seleccionar "Align source images" en la ventana "Preprocessing Options". Asimismo, se ofrecen opciones de reducción de ruido, aberraciones cromáticas y efectos de ghosting. Seleccionarlas las que se deseen y hacer clic en "OK".
5. Seleccionar "Process: Exposure Fusion" y hacer clic en el botón "Process".

Apéndice B

Inserción de aperturas codificadas en el objetivo *Canon EF 50mm f/1.8 II*

En este apéndice se detalla el procedimiento de inserción de una apertura codificada en el objetivo *Canon EF 50mm f/1.8 II*.

Material necesario (ver Figura B.1)

- Destornillador pequeño
- Paño para limpiar lentes
- Contenedor para las partes pequeñas
- Cinta aislante



Figura B.1: Objetivo *Canon EF 50mm f/1.8 II* junto a los materiales empleados.

Es importante trabajar en un espacio ordenado y limpio, y ser cuidadoso a la hora de manipular el objetivo, pues consta de varias piezas pequeñas

susceptibles de extraviarse, así como partes delicadas que pueden dañarse irreversiblemente.

Metodología:

1. Desenroscar la tapa de atrás. Como se muestra en la Figura B.2, aparecerán dos pequeños tornillos que deberán ser retirados. A continuación, poner el objetivo en enfoque manual y girar hacia abajo el disco de enfoque. Empujar suavemente la parte metálica hacia abajo con el destornillador, y hacer palanca contra la cubierta de plástico para sacarla. Es muy importante tener cuidado de no dañar la circuitería interior, la lente o los cables.



Figura B.2: Objetivo montado, con la tapa trasera desenroscada.

2. Presionar desde el interior del objetivo para despegar la tapa lateral. Una vez que ésta ceda, hacer palanca con el destornillador para soltarla del todo, como se muestra en la Figura B.3.



Figura B.3: Objetivo con la tapa lateral despegada.

3. Quitar el último tornillo y la pieza metálica que sujeta, indicados en la Figura B.3. Una vez hecho esto, girar el disco de enfoque hacia arriba y el objetivo quedará desmontado en las tres piezas que se muestran en la Figura B.4.

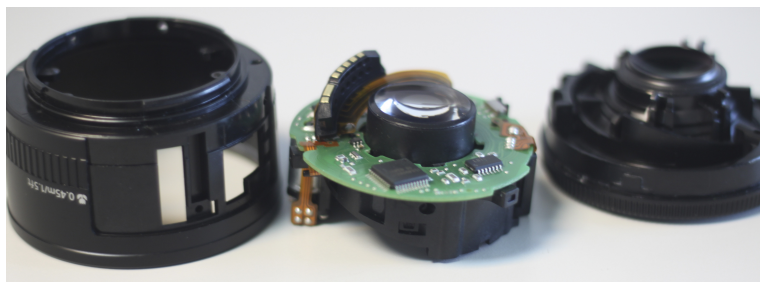


Figura B.4: Objetivo desmontado en sus tres piezas principales.

4. En la pieza de la lente, colocar la apertura codificada como se muestra en la Figura B.5. Es importante colocarla bien centrada para que todo el patrón esté dentro del círculo de la lente. Sujetarla con pequeñas tiras de cinta aislante, con cuidado de no tapar el patrón con ella.

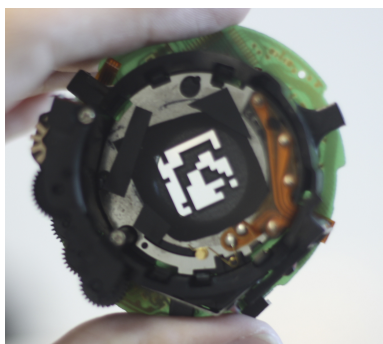


Figura B.5: Pieza central con la apertura codificada.

5. Limpiar bien las lentes y volver a ensamblar el objetivo, repitiendo a la inversa los pasos 1-4.

Apéndice C

Artículo: *Analysis of Coded Apertures for Defocus Deblurring of HDR Images*

En este apéndice se adjunta el artículo resultante de la primera parte del proyecto, correspondiente al análisis del empleo de aperturas codificadas para corrección de desenfoques en imágenes HDR. El artículo fue sometido al **Congreso Español de Informática Gráfica** (CEIG 2012), y a día de hoy se encuentra pendiente de aceptación.

Analysis of Coded Apertures for Defocus Deblurring of HDR Images

Abstract

In recent years, research on computational photography has reached important advances in the field of coded apertures for defocus deblurring. These advances are known to perform well for low dynamic range images (LDR), but nothing is written about the extension of these techniques to high dynamic range imaging (HDR).

In this paper, we focus on the analysis of how existing coded apertures techniques perform in defocus deblurring of HDR images. We present and analyse three different methods for recovering focused HDR radiances from an input of blurred LDR exposures and from a single blurred HDR radiance, and compare them in terms of the quality of their results, given by the perceptual metric HDR-VDP2. Our research includes the analysis of the employment of different statistical deconvolution priors, made both from HDR and LDR images, performing synthetic experiments as well as real ones.

Categories and Subject Descriptors (according to ACM CCS): I.4.3 [Image Processing and Computer Vision]: Enhancement—Sharpening and deblurring

1. Introduction

The field of computational photography has obtained impressive results in last years, improving conventional photography results. One well known problem that conventional cameras present is the limitation of the sensor to capture images with an extended dynamic or luminance range. In a conventional camera the luminance range is limited and parts of the scene which present luminance out of the range would not be correctly represented. In this context HDR imaging (High Dynamic Range imaging) [RWPD05] is a strategy to capture and represent the extended luminance range present in real scenes.

Also in terms of defocus deblurring computational photography has reached important advances. Since image capture can be modelled as a convolution between the focused image and the blur kernel plus a noise function, recovering a sharp image is reduced to a deconvolution problem. However, traditional circular apertures present a very poor response in frequency domain with multiple zero-crossings and attenuation in high frequencies. Thus, recovered images present poor quality. Coded apertures are designed to have an appropriate frequency response to solve this problem, placing them in the camera lens in order to code light before it reaches the camera sensor. The defocus blur is encoded and high frequencies are better preserved in the original image, obtaining better deblurred images after deconvolution.

This work turns around both approaches, analysing the use of coded aperture for defocus deblurring techniques in HDR imaging. While it is well known that the use of coded apertures for defocus deblurring offers good performance with LDR images [ZN09], we believe this is the first time that these techniques are extended to HDR imaging.

For this purpose, we rely in a coded aperture specifically designed for defocus deblurring of LDR images by Zhou et al. [ZN09] and use it to analyse this problem in HDR images. The pattern of this aperture can be seen in Figure 1 beside its power spectrum compared to that of a circular aperture. Note that this aperture offers a better frequency response for defocus deblurring than the circular aperture.

We propose and analyse three different processing models for recovering focused HDR images, one from a single blurred HDR radiance and two from an input of blurred LDR exposures, and analyse them first in simulation environment and finally in real scenarios. We also analyse the use of deconvolution statistical priors, made both from HDR and from LDR images, taking into account the work of Pouli et al. [PCR10] and following the approach that, to solve HDR problems, the use of HDR priors instead of LDR ones would lead to better results due to the existing statistical differences between both types of images.

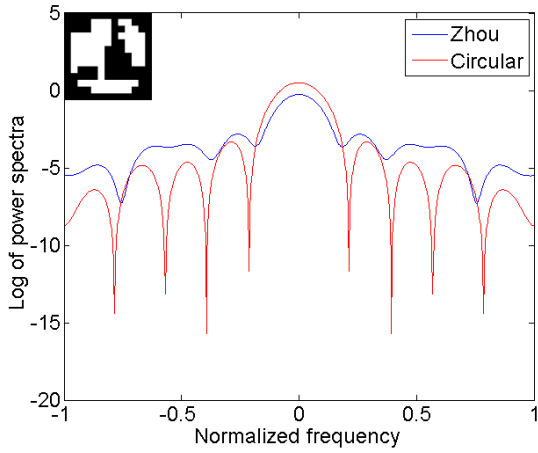


Figure 1: Power spectra of the coded aperture designed for defocus deblurring by Zhou et al. [ZN09] and a conventional circular aperture. Note how the coded aperture pattern offers better frequency response as it avoids zero-crossings and reduces the attenuation in high frequencies.

2. Previous Work

Coded apertures have been traditionally used in astronomy since 1960s to address SNR problems related to lensless imaging, coding the incoming high frequency x-rays and γ -rays. One well known pattern for this purpose is MURA pattern (Modified Uniformly Redundant Array) [GF89].

More recently, in the field of computational photography, Veeraghavan et al. [VRA*07] showed how coded apertures can be used to reconstruct 4D light fields from 2D sensor information. Also coded apertures have been used for solving defocus deblurring problem. The main idea is to obtain coded apertures with better frequency response than the conventional circular aperture. Levin et al. [LFDF07] designed an optimal aperture for depth from defocus and a novel deconvolution method in order to achieve all in-focus images and depth recovery simultaneously. Another approach to recover focus and depth information of a scene was developed by Zhou et al. [ZLN09], in this case obtaining a pair of apertures through genetic algorithms and gradient descent search. Also Zhou et al. presented a metric that evaluates the goodness of a coded aperture for defocus deblurring based on the quality of the resulting deblurred image and image statistics [ZN09]. Recently, Masia et al. [MCPG12] introduced perceptual metrics in the optimization process leading to an aperture design and proved the benefits of this perceptually optimized coded apertures.

For more information about technical details of HDR imaging we refer the reader to Reinhard's book. [RWPD05].

3. Processing methods

The capture process of an image f is given by Equation 1:

$$f = f_0 * k + \eta \quad (1)$$

where f_0 is the focused scene, η is a gaussian white noise with standard deviation σ and k is a convolution kernel determined by the aperture shape and the blur size.

In order to study the viability of the employment of coded apertures for defocus deblurring in HDR images, we simulate the capture process and attempt to recover a sharp image from the simulated blurred image.

Being f_0^{HDR} an HDR scene, we can use the approximation given by Equation 2 to simulate the capture of a High Dynamic Range radiance f^{HDR} only if we are able to capture it in one single shot.

$$f^{HDR} = f_0^{HDR} * k + \eta \quad (2)$$

Some existing cameras allow the capture of extended dynamic range, but in most cases HDR images are obtained by capturing series of LDR exposures and merging them later.

Then, being f_{0n}^{LDR} , ($n = 1, \dots, N$) a set of LDR exposures of the same focused HDR scene f_0^{HDR} , we can simulate the capture of the defocused HDR radiance by first simulating the capture of each exposure following Equation 3, and second merging them into a single HDR defocused radiance as expressed in Equation 4, being g the HDR merging operator.

$$f_n^{LDR} = f_{0n}^{LDR} * k + \eta \quad (3)$$

$$f^{HDR} = g(f_1^{LDR}, f_2^{LDR}, \dots, f_N^{LDR}) \quad (4)$$

Once f^{HDR} is obtained, we can recover the focused HDR radiance \hat{f}_0^{HDR} by performing a single deconvolution. However, as we have the LDR defocused exposures, it is possible to deblur them separately with a set of N deconvolutions and merge them later to obtain \hat{f}_0^{HDR} , following Equation 5.

$$\hat{f}_0^{HDR} = g(\hat{f}_{01}^{LDR}, \hat{f}_{02}^{LDR}, \dots, \hat{f}_{0N}^{LDR}) \quad (5)$$

According to this, we present three different methods for recovering focused HDR radiances:

1. **Processing HDR radiance obtained with a single shot:** Following Equation 2 to model the capture and recovering the focused radiance with a single deconvolution, as seen in Figure 2(a).
2. **Processing HDR radiance obtained by merging LDR exposures:** Following Equations 3 and 4 and recovering the focused HDR with a single deconvolution. The pipeline of this processing is shown in Figure 2(b).
3. **Processing LDR exposures separately before merging:** Following Equation 3 to model the N captures, recovering the focused LDR exposures with N deconvolutions and merging them as in Equation 5 to obtain the HDR focused radiance. This pipeline can be seen in Figure 2(c).

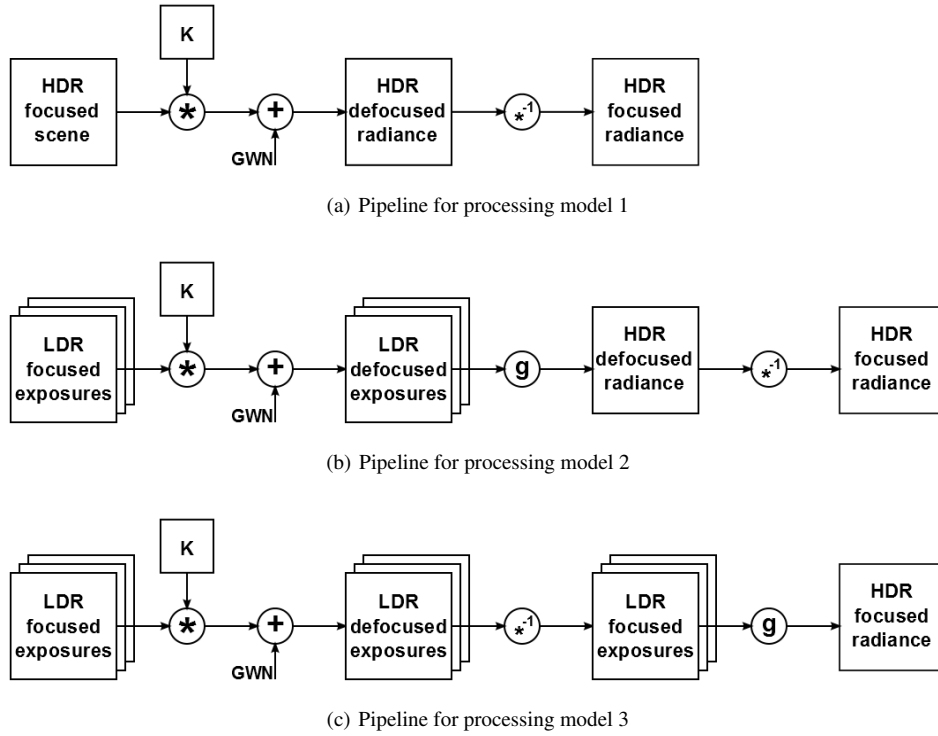


Figure 2: Pipelines for all different processing models, where k is the convolution kernel, GWN is the Gaussian White Noise, g is the HDR merging operator and $*$ is the convolution operator.

4. Simulation of processing models

First we analyse these three models by performing simulations in order to study their viability before proceeding to real experiments. To carry out these, we use one of the coded apertures developed by Zhou et al. [ZN09], which is shown in Figure 1. This aperture is known to work well for defocus deblurring LDR images.

For the simulations we use a set of seven HDR photographs with different dynamic ranges for the first model, and their three corresponding LDR exposures for the other two. One of them is shown in Figure 3. The main goal is to recover the focused HDR images with all three processing methods. We use the perceptual metric *HDR-VDP2* [MKRH11] in order to assess the quality of the results. This metric compares a reference HDR image with its distorted version, providing quality and visibility (probability of detection) measures based on a calibrated model of the human visual system. In this work we focus in obtaining the quality factor Q , a prediction of the quality degradation of the recovered HDR image with respect to the reference HDR image, expressed as a mean-opinion-score (with values between 0 and 100). This metric can not only work with HDR images, but also with their LDR counterparts.

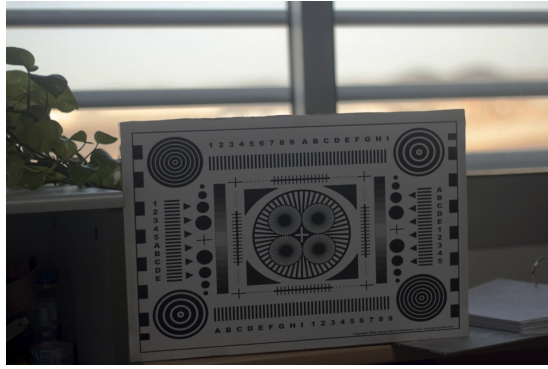
We test four different noise levels, with standard devi-

ations of $\sigma = 0.0005, 0.001, 0.005, 0.05$, and three different deconvolution models based on Wiener deconvolution, which formulation in frequency is given by Equation 6.

$$\hat{F}_0 = \frac{F \cdot \bar{K}}{|K|^2 + |C|^2} \quad (6)$$

Where \hat{F}_0 is the Fourier Transform of the recovered image, \bar{K} is the complex conjugate of K , $|K|^2 = K \cdot \bar{K}$ and $|C|^2 = |\sigma/F_0|^2$ is the Noise to Signal Ratio (NSR) matrix of the original image. From this deconvolution, we study these three different variations:

- **Wiener Deconvolution without prior, with constant NSR matrix.** Replacing $|C|^2$ in Equation 6 by a constant NSR matrix. We tested several values and found that there is a trade-off between noise and ringing in resulting images. We finally decided to set $|C|^2 = 0.005$ achieving good balance between both artifacts.
- **Wiener Deconvolution using natural HDR image prior.** Replacing $|F_0|^2$ in Equation 6 by an statistical prior matrix averaging power spectra of a series of 198 HDR images. We construct the prior employing HDR images from the database of Tania Pouli (<http://taniapouli.co.uk/research/statistics/>).
- **Wiener Deconvolution using natural LDR image prior.** Replacing $|F_0|^2$ as in the previous, using a prior of 198



(a) Tone mapped HDR



(b) Overexposed



(c) Medium exposed



(d) Sub exposed

Figure 3: Example of one of the HDR images used in simulation, with the three exposures merged to obtain it, with relative exposures of +2, 0 and -2 stops.

LDR images instead, extracted from the database of Tania Pouli.

We explore the use of HDR priors in processing models 1 and 2 as we are deconvolving an HDR radiance, inspired by Pouli et al. [PCR10]. Note that we do not perform processing model 3 with HDR prior since we are deconvolving LDR images with it. Since the aperture we are using is optimized for a noise level of $\sigma = 0.005$ we set this value for standard deviation of gaussian noise in our deconvolutions with priors.

5. Performance comparison

Once all the simulations are finished, we compute the mean quality factor Q , given by the *HDR-VDP2* metric, of the seven images obtained with the three proposed processing models shown in Figure 2. For each model we analyse four different noise levels and the three different deconvolution variations explained in Section 4. This information is collected in Figure 4.

We can see how the use of priors is strongly recommended for processing model 1 when image noise is very high. In this noisy scenario, HDR prior offers better results than LDR prior, however, when image noise decreases, all three different deconvolutions produce similar behaviours.

As expected, HDR prior outperforms LDR prior in processing model 2, but we can see how the use of Wiener deconvolution with constant NSR matrix offers similar or even better quality all along the noise variation.

For processing model 3, the use of constant NSR matrix in the deconvolution seems to offer better results than the LDR prior.

In regard to the comparison between all three processing models, we can see how processing model 1 clearly derives in better results than the other two, while processing model 2 seems to be the worst of them. Also, processing model 1 offers the lowest cost as it only requires one deconvolution, while processing model 2 requires one deconvolution and one exposure fusion, and processing model 3 requires one deconvolution for each exposure and one exposure fusion.

In Figure 5 we show the result of one of the noisy simulations ($\sigma = 0.05$) using processing model 1, with both priors. We can see how the use of HDR prior slightly reduces the recovered image noise. In Figure 6 we show an example of the same HDR scene recovered with processing model 2, with both priors, this time with $\sigma = 0.0005$. In this low noise scenario we can appreciate how the use of HDR prior instead of LDR one results in a reduction of ringing artifacts.

6. Validation in real scenarios

After performing the simulations we proceed to validate the same processes in real scenarios. We can not validate processing model 1 in real scenarios because of the lack of the required equipment: an HDR camera that allows the capture of an HDR image with a single shot. For this reason, physical validation is restricted to processing models 2 and 3, which for simplicity we call HDR method and LDR method. We use a DSLR camera Canon EOS 500D with an EF 50mm f/1.8 II lens for all the tests. The same coded aperture used in simulation (Figure 1) is printed and inserted into the camera lens.

6.1. Image capture process

We construct a scene with large luminance range and capture three images using the multi-bracketing camera option set to relative exposures of +2, 0 and -2 stops. For these captures we fix the ISO setting value at 100 and aperture size at F2.0, leading to exposure times of 1/5, 1/20 and 1/80 seconds.

We place the scene 180 cm away from the camera, and set the focus plane at 120 cm, leading to a defocus distance of 60 cm. We also take three exposures of the well focused scene to obtain a ground truth HDR image that allows comparison, using the same capture parameters described above.

All images are taken in RAW format, with a size of 4752x3168 pixels. To reduce computational time and cost we resize images by a factor of 0.2, reducing them to 951x634 pixels.

6.2. System calibration

In order to recover the focused HDR image of the scene we need to know the PSF of the capture system as the response

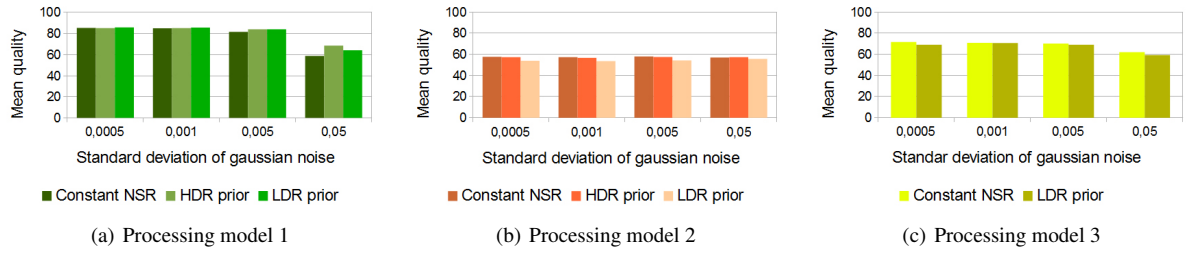


Figure 4: Mean Q obtained with the HDR-VDP2 metric for each processing model, with all different combinations of noise and deconvolution prior.

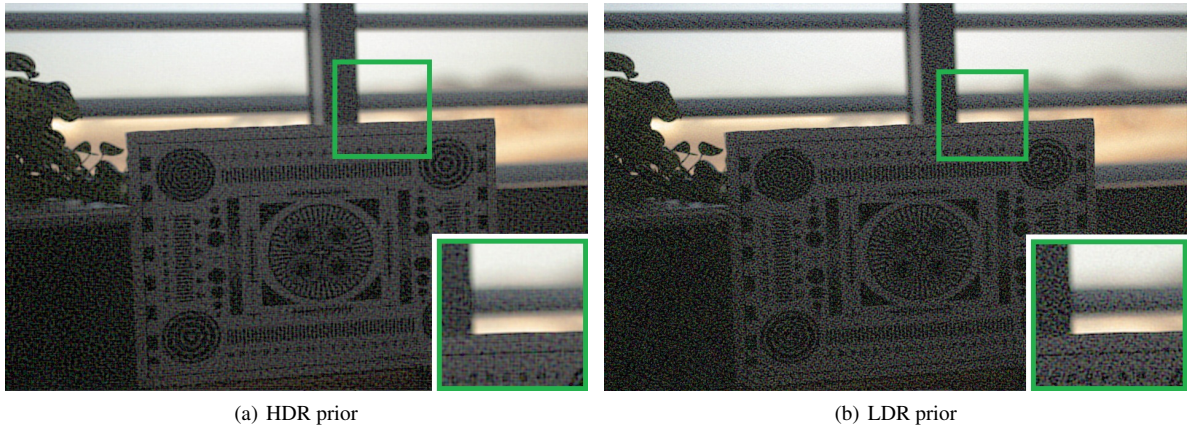


Figure 5: Comparison between images recovered after simulation of processing model 1, with HDR and LDR priors and $\sigma = 0.05$. Note how the use of the HDR prior instead of LDR slightly reduces image noise.

to an impulse, to use it as the kernel in the deconvolution process. To calibrate the PSF at the depth of interest (60 cm) we use a LED mounted on a pierced thick black cardboard in order to make a point light source. We lock the focus plane at 120 cm and place the cardboard with the LED at 180 cm. In order to be coherent with image capture, we obtain three images, one for each exposure value, with the same capture parameters used to capture the scene. The central detail of these images is shown in Figure 7. We also obtain an HDR image of the montage to obtain the PSF that we will use in the deconvolution in the HDR method.

The cropped image of the LED serves us as PSF, after thresholding it in order to eliminate residual light, and normalizing it to preserve energy in the deconvolution process. Note that the threshold value changes for each PSF, increasing with the exposure value: 0.39 for underexposed, 0.5 for well-exposed and 0.8 for overexposed. For the PSF used in the HDR method the threshold value is 0.2. The resulting PSFs are shown in Figure 8. After image resizing the kernel, its size is approximately 14x14 pixels.

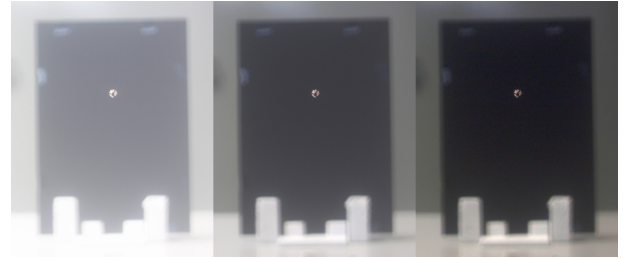


Figure 7: Central detail of the three different exposures used to recover the PSFs.

6.3. Deblurred images recovery

Once we obtain the PSFs we recover the sharp images following the HDR and LDR methods. For the HDR one we merge the defocused exposures into a defocused HDR radiance and obtain the deblurred HDR image performing a single deconvolution using the HDR kernel, as in Figure 2(b). For the LDR method we perform one deconvolution for each

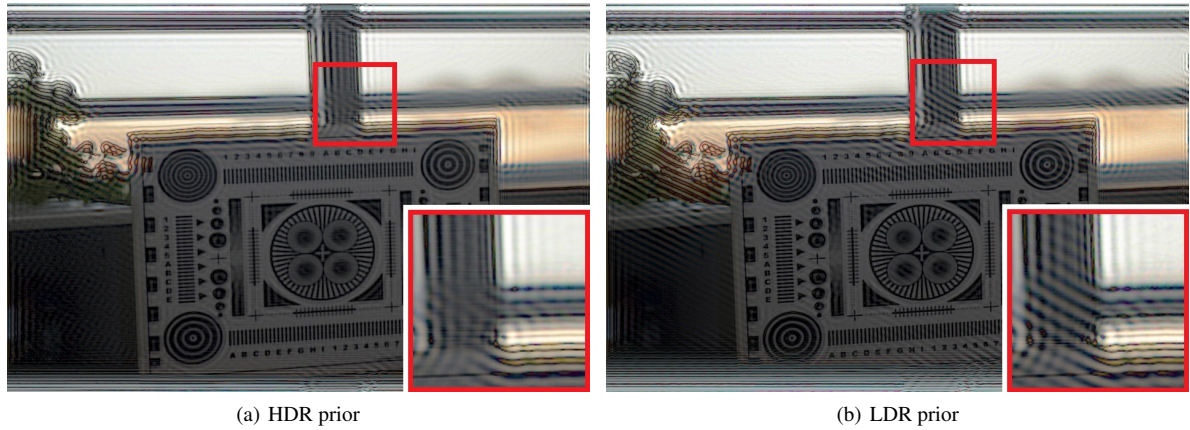


Figure 6: Comparison between images recovered after simulation of processing model 2, with HDR and LDR priors and $\sigma = 0.0005$. Note how using HDR prior instead of LDR seems to reduce image ringing.

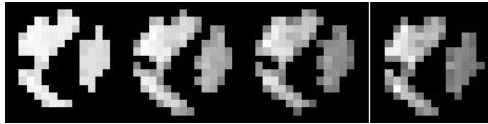


Figure 8: PSFs obtained for deconvolution. From left to right: PSF for the high, the central and the low exposure, used in the LDR method, and PSF obtained by merging the three exposures used in the HDR method.

defocused exposures, using the corresponding PSF for each one, and merge the resulting recovered exposures into the focused HDR image, as in Figure 2(c).

We carry out the same Wiener deconvolution models described in Section 4, excluding again the use of HDR prior for the LDR method.

7. Results and discussion

Once we perform all the experiments, we compare the results of both methods. We compute the quality factor Q given by the *HDR-VDP2* metric of the HDR recovered images and show the results in two different scenarios. We also check the effect of the use of the different deconvolution models, specially those which employ deconvolution priors.

7.1. Methods comparison

For our first scenario, in Figure 9 we show the quality factor Q , given by the *HDR-VDP2* metric, of the HDR images recovered with each processing. These results indicate that, while simulation results suggested that the LDR method offered better results than the HDR method (see Figure 4), real

experiments point that both methods offer very similar qualities. Note that, according to these metrics, the use of priors results in worse performance. We explore this fact further in Subsection 7.2.

We show the result of both methods, using constant NSR, in Figure 10, in order to offer a visual comparison of how both methods perform. We also show the original HDR radiance and the ground truth ideal HDR radiance. We can see how both methods offer similar results as the metrics indicate, while the HDR method requires only one deconvolution and the LDR method needs one deconvolution for each exposure. Furthermore, attending to the highlighted details and comparing recovered and original images we see how both methods are able to recover the well focused HDR radiance. These images prove that the employment of coded apertures for defocus deblurring of HDR images is viable and presents a good performance.

We test again our approximations performing the experiments in a new scenario, in order to check if the second results correlate with the first ones. In Figure 11 we show the quality factor Q given by the *HDR-VDP2* metric for this second scenario.

Again, the use of priors derives in worse results than the use of constant NSR, for both processing methods. In Figure 12 we show the HDR images of this scenario recovered with the HDR and LDR methods with constant NSR. As we can see, both methods offers good results, recovering the focused image with little ringing artifacts.

7.2. Effects of using prior

As shown in Figures 9 and 11, in real experiments we see that both HDR and LDR methods perform much better when no deconvolution prior is used. We check the images recovered with both priors in order to know why this happens.

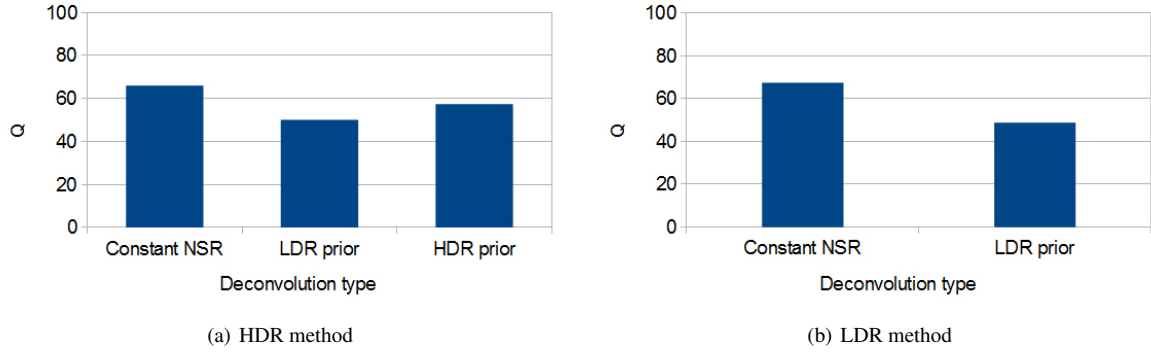


Figure 9: Quality factor Q obtained with the HDR-VDP2 metric for our first real scenario, for each processing model and deconvolution. We can observe how in the HDR method the HDR prior outperforms the LDR one, and how both LDR and HDR methods using constant NSR offer similar quality.

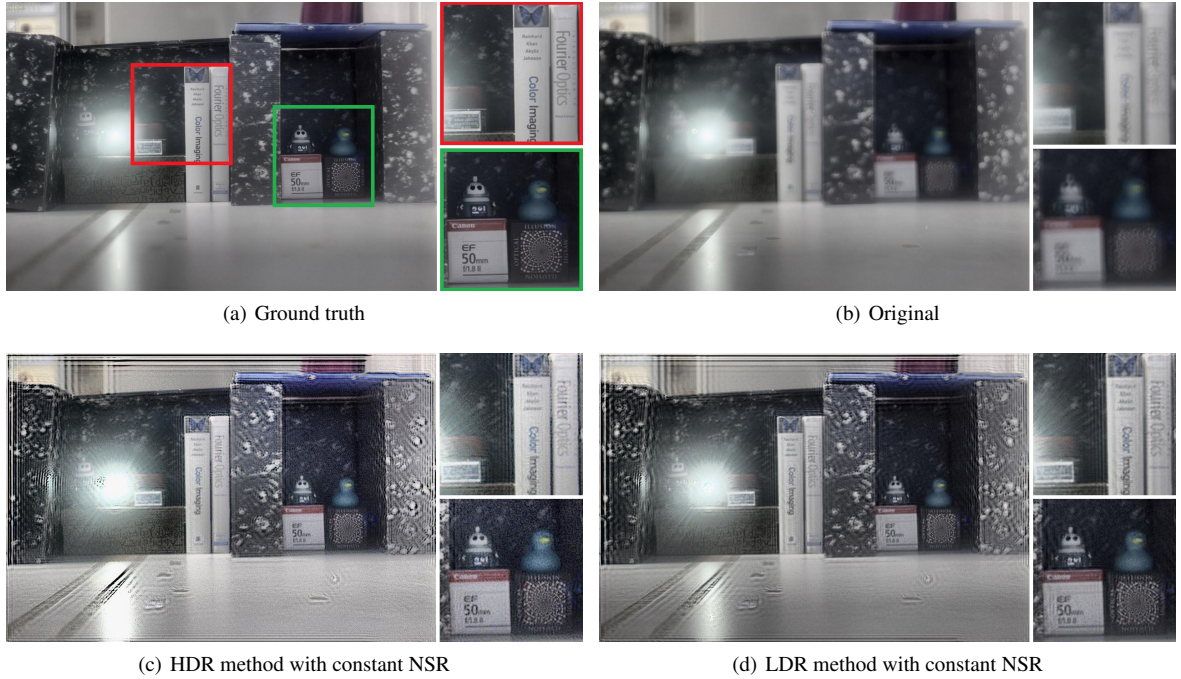


Figure 10: HDR results obtained for our first real scenario with best processing methods in terms of Q (c,d), compared to ground truth and original images, all of them tone mapped. Here we see how both methods offer good and similar results.

If we carefully observe these images we can appreciate a grid shaped distortion, as seen in Figure 13. This distortion clearly reduces the visual quality of the images recovered with deconvolution prior. Further, we notice again that HDR prior outperforms LDR prior in the HDR method, as it minimizes, but not completely removes, this distortion.

We explore the variation of σ in the deconvolution process and see the impact of this alteration in the described

distortion. This variation corresponds to a higher weight for the deconvolution prior. In Figure 14 we see some of the images obtained with different σ in the deconvolution process for the LDR method. We see how increasing this value we obtain a better reduction of prior distortion and ringing. In exchange, we find that this increase leads to less sharp results, resulting in a trade-off between both effects.

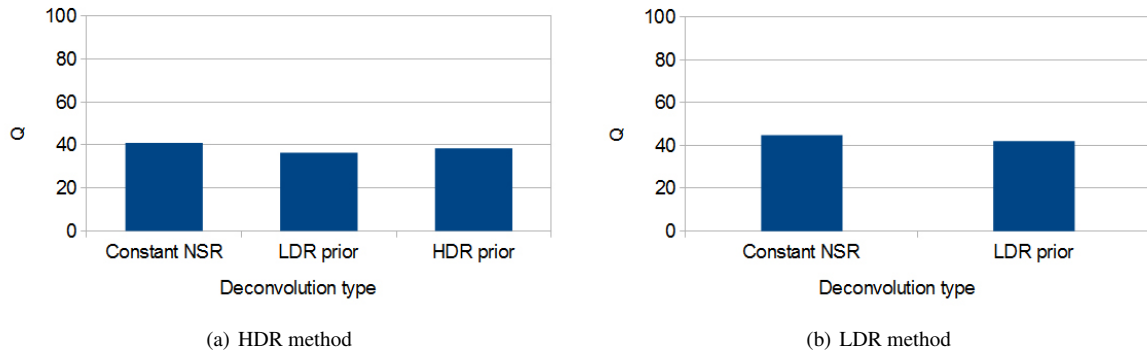


Figure 11: Quality factor Q obtained with the HDR-VDP2 metric for our second real scenario, for each processing model and deconvolution. Note that in the HDR method the HDR prior outperforms the LDR one, and that in both methods the use of constant NSR offers the best results.

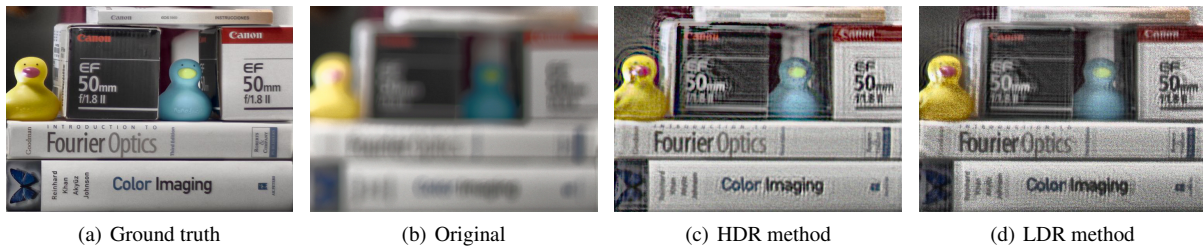


Figure 12: HDR results obtained for our second real scenario with best processing methods in terms of Q , compared to ground truth and original images, all of them tone mapped. Here we see how both methods are able to recover sharp details such as the book titles.

8. Conclusions and future work

In this paper we explore for the first time, to our knowledge, the use of coded apertures for defocus deblurring of HDR images, showing that these techniques, which used to be employed in LDR images, can be extended for HDR imaging. We implement three different processing models, either responding to an input of an HDR defocused radiance or a series of LDR defocused exposures of the same scene.

The first processing model offers the best results in simulation, but due to the limited dynamic range of our camera, we are not able to capture defocused HDR images with a single shot. Thus we could not test this model as properly as we would like to, so the first future work that follows this paper is to perform more experiments in this way, employing more advanced cameras that allow the capture of extended dynamic range images with just one shot.

The two other processing models are validated with real experiments, finding that both of them are viable and offer good quality results. We show that the proposed HDR method performs as good as the LDR one, despite the fact that simulations indicate otherwise, and reduces the computational cost as it only requires one deconvolution.

We conclude that the use of deconvolution priors made of HDR images instead of conventional LDR priors leads to better performance, but due to the fact that the prior we are employing is far from optimal the best results come when no prior is employed in the process. From this, and relying on the work of Pouli et al. [PCR10], we believe that more research related to HDR priors is needed. As many optimization problems benefit from the use of statistical regularities of the images, and taking into account the advances on HDR imaging, we may think that the construction of good HDR priors is another avenue of future work.

One of the immediate applications of these new priors, which is highly related to our work, is the design of optimal aperture patterns for defocus deblurring of HDR images. As the aperture we have employed [ZN09] is obtained by means of a genetic algorithm which uses prior information of LDR images, we may think that it is possible to obtain new specific coded apertures optimized for HDR.

Finally, we hope that the results shown in this work will encourage other researchers to further explore this line.

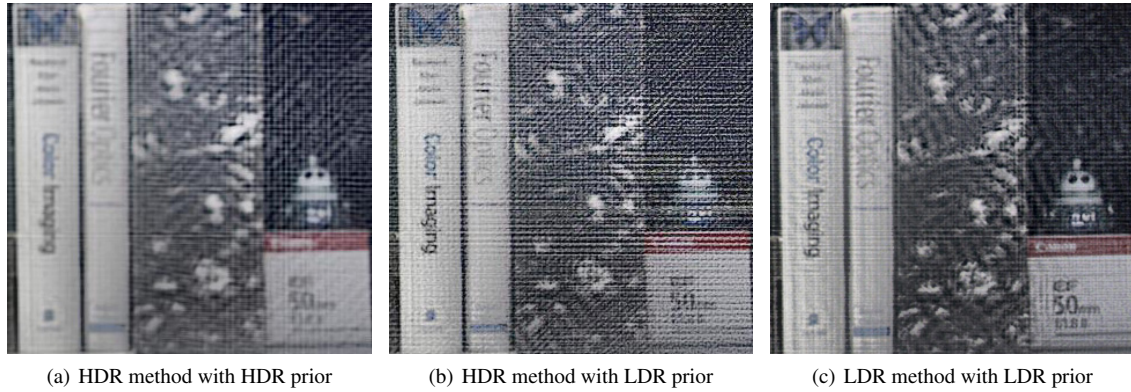


Figure 13: Detail of our recovered images of the first real scenario using priors, where we can appreciate a clear grid shape distortion. Note that, in the HDR method, using HDR prior instead of LDR one reduces this effect. All the images are tone mapped.

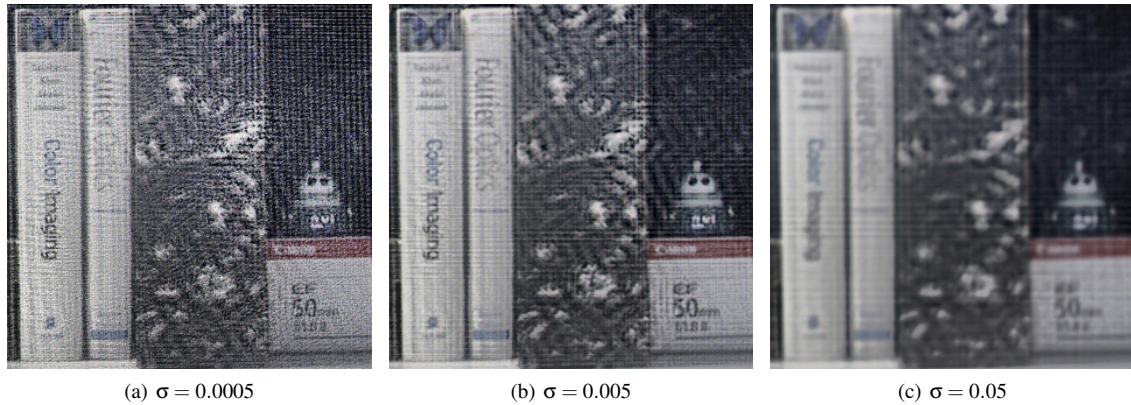


Figure 14: Effect of the variation of σ in the deconvolution seen for the LDR method. We can see a trade-off between the grid shape distortion and image sharpness. All the images are tone mapped.

References

- [GF89] GOTTESMAN S., FENIMORE E.: New family of binary arrays for coded aperture imaging. *Applied Optics*, 20 (1989), 4344–4352. [2](#)
- [LDF07] LEVIN A., FERGUS R., DURAND F., FREEMAN W.: Image and depth from a conventional camera with a coded aperture. *ACM Transactions on Graphics* 26, 3 (2007). [2](#)
- [MCPG12] MASIA B., CORRALES A., PRESA L., GUTIERREZ D.: Perceptually optimized coded apertures for defocus deblurring. *Computer Graphics Forum* (2012). To appear. [2](#)
- [MKRH11] MANTIUK R., KIM K. J., REMPEL A. G., HEIDRICH W.: HDR-VDP-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions. *ACM Transactions on Graphics* 30, 40 (2011). [3](#)
- [PCR10] POULI T., CUNNINGHAM D., REINHARD E.: Statistical regularities in low and high dynamic range images. *ACM Symposium on Applied Perception in Graphics and Visualization (APGV)* (July 2010). [1, 3, 8](#)
- [RWPD05] REINHARD E., WARD G., PATTANAİK S., DEBEVEC P.: *High Dynamic Range Imaging*. Morgan Kaufmann Publishers, 2005. [1, 2](#)
- [VRA*07] VEERARAGHAVAN A., RASKAR R., AGRAWAL A., MOHAN A., TUMBLIN J.: Dappled photography: mask enhanced cameras for heterodyned light fields and coded aperture refocusing. *ACM Transactions on Graphics* 26 (July 2007). [2](#)
- [ZLN09] ZHOU C., LIN S., NAYAR S.: Coded aperture pairs for depth from defocus. In *IEEE International Conference on Computer Vision (ICCV)* (Kyoto, Japan, 2009). [2](#)
- [ZN09] ZHOU C., NAYAR S. K.: What are Good Apertures for Defocus Deblurring? In *IEEE International Conference on Computational Photography* (San Francisco, CA, USA, 2009). [1, 2, 3, 8](#)