

UNIVERSIDAD DE INGENIERÍA Y TECNOLOGÍA

CARRERA DE INGENIERÍA ELECTRÓNICA



**RECONSTRUCCIÓN DE IMÁGENES A COLOR
UTILIZANDO MUESTRAS INCOMPLETAS DE
NUEVE BANDAS ESPECTRALES**

TESIS

Para optar el título profesional de Ingeniero Electrónico

AUTOR:

Juan Carlos Barbaran Meza 

ASESOR

Victor Murray 

Lima - Perú

2021

Dedicatoria:

Dedico este trabajo con todo mi corazón a mi madre, quien siempre hizo todo lo que estaba a su alcance para poder brindarme una educación de calidad. Y le agradezco por apoyarme siempre, por creer en mí y por brindarme los mejores consejos día a día. Así mismo a mi familia que me ha enseñado a no rendirme nunca y de cada uno de ellos he aprendido algo que estoy seguro me hará cumplir todas mis metas.

Agradecimientos:

En primer lugar quiero agradecer a mi asesor y mentor, Victor Murray, quien con sus conocimientos, consejos y apoyo me guió a través de cada una de las etapas de este proyecto para alcanzar los resultados obtenidos.

También quiero agradecer al equipo del laboratorio de Nanophotonics de la universidad de Harvard dirigido por Fawwaz Habbal, quienes iniciaron este proyecto y que sin su apoyo este trabajo no habría podido llevarse a cabo.

Por último quiero agradecer a mi alma mater UTEC, que ha sido la plataforma que me ha permitido desarrollar no solo mi habilidades técnicas, sino también mis habilidades blandas. Estoy seguro que lo aprendido me permitirá alcanzar cualquier meta que me proponga.

Índice general

	Pág.
RESUMEN	1
ABSTRACT	2
CAPÍTULO 1 INTRODUCCIÓN	3
CAPÍTULO 2 ANTECEDENTES	6
2.1 Filtros MSFA basados en CFA	6
2.1.1 Filtros MSFA basados en <i>nanowires</i>	7
2.1.2 Prototipo a utilizar	8
2.2 Métodos de clasificación y regresión	9
2.2.1 Algoritmo <i>support-vector machine</i>	10
2.2.2 Algoritmo <i>k-nearest neighbors</i>	11
2.2.3 Regresión lineal multivariable	11
2.2.4 Regresión no lineal multivariable	12
2.3 Cartilla de Macbeth	13
2.4 Imágenes de alto rango dinámico	14
CAPÍTULO 3 MÉTODOS	16
3.1 Base de datos	16
3.2 Separación de muestras	17
3.3 Reconstrucción de color	19
3.3.1 Preprocesamiento	19

3.3.2 Aprendizaje supervisado	21
3.3.3 Posprocesamiento	23
3.4 Reconstrucción utilizando HDR	24
3.4.1 Primera forma de obtener HDR (Tipo 1)	24
3.4.2 Segunda forma de obtener HDR (Tipo 2)	25
3.4.3 Tercera forma de obtener HDR (Tipo 3).	27
CAPÍTULO 4 RESULTADOS Y DISCUSIÓN	29
4.1 Pruebas de error numéricas	29
4.1.1 Pruebas cuantitativas sin utilizar HDR	30
4.1.2 Pruebas cuantitativas utilizando HDR	36
4.2 Pruebas cualitativas	41
4.2.1 Descripción de la prueba	41
4.2.2 Prueba de percepción de imágenes de la Cartilla de Macbeth	42
4.2.3 Prueba de percepción de imágenes de botellas	45
4.3 Prueba de ampliación de rango dinámico	47
4.4 Trabajos de investigación publicados	49
CONCLUSIONES	50
RECOMENDACIONES	51
ANEXOS	52

Índice de tablas

4.1	MSE promedio de las tres bandas de color (R, G, B) usando (4.1). Se incluyen todos los tiempos de exposición de imágenes obtenidas al capturar la cartilla de Macbeth usando el prototipo para cada uno de los métodos usados y un tamaño de 120×150 píxeles.	31
4.2	SSIM promedio de las tres bandas de color (R, G, B) usando (4.2). Se incluyen todos los tiempos de exposición de imágenes obtenidas al capturar la cartilla de Macbeth usando el prototipo para cada uno de los métodos usados y un tamaño de 120×150 píxeles.	31
4.3	MSE promedio de las tres bandas de color (R, G, B) usando (4.1). Se incluyen los dos mejores y el peor tiempo de exposición basado en los resultados de las Tablas 4.1 y 4.2 para cada tipo de separación de muestras.	32
4.4	SSIM promedio de las tres bandas de color (R, G, B) usando (4.2). Se incluyen los dos mejores y el peor tiempo de exposición basado en los resultados de las Tablas 4.1 y 4.2.	36
4.5	Resultados de la primera forma de usar imágenes de alto rango dinámico en la reconstrucción de color para seis diferentes combinaciones de tiempos de exposición en cada uno de los métodos. Valores promedio de las tres bandas de color (R, G, B).	38
4.6	Resultados de la segunda forma de usar imágenes de alto rango dinámico en la reconstrucción de color para seis diferentes combinaciones de tiempos de exposición en cada uno de los métodos. Valores promedio de las tres bandas de color (R, G, B).	39

4.7	Resultados de la tercera forma de usar imágenes de alto rango dinámico en la reconstrucción de color para seis diferentes combinaciones de tiempos de exposición en cada uno de los métodos. Valores promedio de las tres bandas de color (R, G, B).	40
-----	--	----

Índice de figuras

2.1	Esquema de construcción de los filtros de <i>nanowires</i> . En la primera etapa se construyen los <i>nanowires</i> en un wafer de silicon para luego ser cubiertos con un polímero transparente que permite el ingreso de la luz entre los <i>nanowires</i> . Finalmente los filtros son retirados utilizando una cuchilla y son adheridos a cada uno de los píxeles del sensor [8].	8
2.2	Distribución de los filtros de <i>nanowires</i> en el sensor. Cada uno de los colores representa un tipo de filtro agrupado en secciones de 3×3 píxeles llamado <i>spectral band group</i> (SBG). La unión de 9 diferentes SBG da a lugar a un <i>macro pixel</i> (MP) que tiene un tamaño de 9×9 píxeles. Finalmente la agrupación de MP se le conoce como array filter area (AFA), en el caso del prototipo propuesto tiene un tamaño de 1080×1350 píxeles.	9
2.3	Esquema de clasificación utilizando SVM. El algoritmo busca el hiperplano óptimo que separa las diferentes clases optimizando el margen máximo entre estas.	10
2.4	Esquema de clasificación utilizando <i>k-Nearest Neighbors</i> . Con un modelo de características entrenado, se predice la clase de una nueva muestra (representada con el color verde y la variable X) asumiendo la menor diferencia de valor en esas características.	11
2.5	Representación sintética de la cartilla de Macbeth. Este es un verificador de 24 colores distribuidos en 4 filas y 6 columnas.	13

2.6	Imágenes obtenidas con diferente tiempo de exposición y su respectiva imagen HDR generada a partir de esta.	14
3.1	Ejemplos de imágenes de la cartilla de Mcbeth. A la izquierda una imagen capturada con una cámara a color y a la derecha, una imagen capturada por el prototipo.	17
3.2	Ejemplos de imágenes de botellas. A la izquierda una imagen capturada con una cámara a color y a la derecha, una imagen capturada por el prototipo.	17
3.3	Esquema de separación de muestras. Al inicio, se separan los píxeles pertenecientes a un mismo tipo de filtro, seguido de ello se toma el píxel central de cada SBG para finalmente unir todos los píxeles restantes (A) o interpolar para obtener tamaños mayores de la imagen (B-C).	18
3.4	Esquema de preprocesamiento. La imagen original es multiplicada por una matriz de corrección y luego de ello se separan las muestras de los filtros para obtener 9 imágenes. A cada una de estas imágenes se le aplica un filtro mediana y una corrección espacial para mejorar los resultados. . .	20
3.5	Esquema de la obtención del modelo. Utilizando la información de los píxeles en forma matricial y los valores ideales en color de estos se entrenó un modelo en los casos de k -NN y SVM. En el caso de	22
3.6	Esquema de posprocesamiento. Se realizan dos procedimientos adicionales a la imagen a color obtenida: eliminación de valores atípicos y un filtro mediana con una ventana de 3×3 píxeles.	23
3.7	Esquema de la primera forma de introducir HDR en la reconstrucción de color. En este procedimiento se utilizan imágenes de diferente tiempo de exposición capturadas con el prototipo para generar una de alto rango dinámico. Con esta nueva imagen se obtiene una a color por el procedimiento ya explicado en secciones anteriores.	25

3.8	Esquema de la segunda forma de introducir HDR en la reconstrucción de color. Utilizando diferentes imágenes a color obtenidas a partir de otras de diferente tiempo de exposición se obtiene una de alto rango dinámico. A esta imagen a color se le aplica el posprocesamiento para obtener el resultado final.	26
3.9	Esquema de la tercera forma de introducir HDR en la reconstrucción de color. Al inicio se separan las muestras de imágenes de diferente tiempo de exposición y las bandas del mismo tipo son utilizadas para generar una imagen HDR. Al final de este procedimiento se obtienen 9 imágenes de alto rango dinámico que son utilizadas para obtener una a color con los procedimientos ya explicados.	28
4.1	Imágenes de la cartilla de Macbeth para diferentes métodos. (a) Mejor resultado en SVM. (b) Mejor resultado en k -NN. (c) Mejor resultado en regresión lineal. (d) Mejor resultado en regresión de segundo orden. (e) Mejor resultado en regresión de tercer orden. (f) Peor resultado en SVM. (g) Peor resultado en k -NN. (h) Peor resultado en regresión lineal. (i) Peor resultado en regresión de segundo orden. (j) Peor resultado en regresión de tercer orden.	34
4.2	Imágenes de la cartilla de Macbeth para diferentes métodos usando el segundo tipo de interpolación (360x450 píxeles). (a) Mejor resultado en SVM. (b) Mejor resultado en k -NN. (c) Mejor resultado en regresión lineal. (d) Mejor resultado en regresión de segundo orden. (e) Mejor resultado en regresión de tercer orden. (f) Peor resultado en SVM. (g) Peor resultado en k -NN. (h) Peor resultado en regresión lineal. (i) Peor resultado en regresión de segundo orden. (j) Peor resultado en regresión de tercer orden.	35

4.3	Imágenes de la cartilla de Macbeth para diferentes métodos usando la primera forma de aplicar HDR. (a) Mejor resultado en SVM. (b) Mejor resultado en k -NN. (c) Mejor resultado en regresión lineal. (d) Mejor resultado en regresión de segundo orden. (e) Mejor resultado en regresión de tercer orden.	38
4.4	Imágenes de la cartilla de Macbeth para diferentes métodos usando la segunda forma de aplicar HDR. (a) Mejor resultado en SVM. (b) Mejor resultado en k -NN. (c) Mejor resultado en regresión lineal. (d) Mejor resultado en regresión de segundo orden. (e) Mejor resultado en regresión de tercer orden.	39
4.5	Imágenes de la cartilla de Macbeth para diferentes métodos usando la tercera forma de aplicar HDR. (a) Mejor resultado en SVM. (b) Mejor resultado en k -NN. (c) Mejor resultado en regresión lineal. (d) Mejor resultado en regresión de segundo orden. (e) Mejor resultado en regresión de tercer orden.	40
4.6	Captura parcial de la prueba de percepción con las imágenes de la cartilla de Macbeth. Se visualiza la mitad del total de imágenes presentadas a los voluntarios.	42
4.7	Imágenes de la cartilla de Macbeth usados en la primera prueba de percepción visual ordenadas de mejor a peor (a-h) según los resultados de la prueba de percepción. (a) Imagen a color. (b) No lineal de tercer orden HDR tipo 3. (c) No lineal de segundo orden HDR tipo 3. (d) Lineal HDR tipo 3. (e) No lineal de tercer orden. (f) No lineal de segundo orden. (g) SVM. (h) Lineal.	43

4.8	Gráfica de barras de los puntajes promedio de cada uno de los métodos presentados en la primera prueba de percepción visual. En azul los resultados para las personas con experiencia en procesamiento de imágenes o fotografía. En naranja los que no tienen experiencia. En gris el puntaje tomando en cuenta el total de participantes: los dos anteriores grupos más los que no indicaron ninguna categoría.	44
4.9	Imágenes de botellas usados en la segunda prueba de percepción visual ordenadas de mejor a peor (a-h) según los resultados de la prueba de percepción. (a) No lineal de tercer orden HDR tipo 3. (b) No lineal de segundo orden HDR tipo 3. (c) Lineal HDR tipo 3. (d) Lineal. (e) Imagen a color. (f) No lineal de segundo orden. (g) No lineal de tercer orden. (h) SVM.	46
4.10	Gráfica de barras de los puntajes promedio de cada uno de los métodos presentados en la segunda prueba de percepción visual. En azul los resultados para las personas con experiencia en procesamiento de imágenes o fotografía. En naranja los que no tienen experiencia. En gris el puntaje tomando en cuenta el total de participantes: los dos anteriores grupos más los que no indicaron ninguna categoría.	47
4.11	Imágenes de frutas reconstruidas en color usando regresión no lineal de 3er orden con HDR de tercer tipo y ampliando el rango dinámico. (a) Plátanos. (b) Manzanas. (c) Uvas y una manzana verde. (d) Mandarinas y una naranja.	48

RESUMEN

En los últimos años, el uso de cámaras multiespectrales ha ido en aumento debido a su capacidad para analizar diferentes características de los materiales. Así, múltiples métodos para desarrollarlas están siendo estudiados. Un nuevo prototipo de una cámara multiespectral de tipo *multispectral filter array*, que utiliza *nanowires* como filtros, se ha desarrollado en el grupo de investigación de *Nanophotonics* de la universidad de Harvard. En este tipo de cámara, se sobrepone un filtro en cada uno de los píxeles del sensor para capturar una banda específica de la luz. Debido a que el prototipo está en fase de desarrollo, no cuenta con todas las funciones que una cámara multiespectral debería tener. Este trabajo se centró en desarrollar una de estas funciones, la cual es la reconstrucción de imágenes a color a partir de sus nueve bandas capturadas

El desarrollo de la función de reconstrucción de color se realizó en tres etapas: preprocesamiento, adquisición del modelo y posprocesamiento. Para la adquisición del modelo se usaron los métodos de *support vector machine* (SVM), *k-nearest neighbor* (*k*-NN) y regresión multivariable. Así mismo en el proceso de reconstrucción se introdujo el uso de imágenes de alto rango dinámico (HDR, proveniente de *high dynamic range*). El algoritmo utiliza imágenes de muestra del verificador de color Macbeth y de objetos reales o naturales. Fue verificado utilizando pruebas cuantitativas con error cuadrático medio y diferencia estructural, así como pruebas cualitativas con una prueba de percepción visual. Los resultados obtenidos muestran que el mejor método para obtener el modelo de color es la regresión no lineal de tercer orden. Así mismo, el uso de imágenes HDR en la reconstrucción de color mejora en 29.3 por ciento los resultados cuantitativos en la métrica de error cuadrático medio. Por último, se demostró que con el algoritmo propuesto se puede obtener una imagen a color a partir de una multiespectral de nueve bandas.

Palabras clave:

Reconstrucción de color, Cámaras multiespectrales, Filtros multiespectrales

ABSTRACT

In recent years, the use of multispectral cameras has been increasing due to their ability to analyze different materials' characteristics. Thus, multiple methods to develop them are being studied. A new prototype of a multispectral filter array type multispectral camera, which uses nanowires as filters, has been developed at the Nanophotonics research group at Harvard University. In this type of camera, a filter is superimposed on each of the sensor's pixels to capture a specific band of light. Since the prototype is under development, it does not have all the functions that a multispectral camera should have. This work is focused on developing one of these functions, which is the reconstruction of color images from its nine captured bands.

The development of the color reconstruction function will be carried out in three stages: pre-processing, model acquisition, and post-processing. For the model's acquisition, the methods of support vector machine (SVM), k-nearest neighbor (k-NN), linear and multivariate nonlinear regression were used. Likewise, in the reconstruction process, high dynamic range (HDR) images were introduced. The results obtained show that the best method to obtain the color model is third-order nonlinear regression. On the other hand, it is essential to note that the k-NN and SVM methods showed better results in a mean squared error (MSE), but not in structural similarity because the color reconstruction in these methods is limited to the 24 colors that are featured in Macbeth's grid. Finally, the use of HDR in color reconstruction improves the quantitative results in the MSE metric by 29.3 percent.

Keywords:

Color reconstruction, Multispectral cameras, Multispectral Filter Array

Capítulo 1

INTRODUCCIÓN

Según el diccionario de la Real Academia Española [1], un espectrómetro se define como un dispositivo utilizado para medir el comportamiento de la luz. Los seres humanos han pasado siglos estudiando la espectrometría para obtener más información del mundo que los rodea de forma analítica. Una consecuencia de esto es el inicio del estudio de imágenes multiespectrales las cuales se definen en [2] como un grupo de imágenes en el que cada una de ellas representa la captura de parte de la luz en un rango de longitud de onda específico a través del espectro electromagnético, de forma matricial podría definirse como un arreglo tridimensional de tamaño $M \times N \times P$, donde M es el número de filas, N el número de columnas y P el número de bandas espectrales. La investigación de imágenes multiespectrales tuvo su auge en 1970 con el inicio del programa Landsat [3] y desde esa fecha sigue en incremento constante. La importancia del estudio de este tipo de imágenes radica en la información que contienen, ya que las características químicas o físicas de los materiales con los que está hecho un objeto permiten la absorción o reflexión de bandas espectrales específicas [4], haciendo que las imágenes multiespectrales contengan esas características que de otro modo solo serían identificables con procesos más complejos.

A lo largo de los años se han desarrollado diferentes métodos de adquisición de imágenes multiespectrales, cada uno con sus características propias. Uno de estos métodos es el *multispectral filter array* (MSFA), que consiste en colocar un filtro con una banda específica del espectro electromagnético en cada píxel o sensor de luz de una cámara y así obtener la información de ese espectro en ese píxel. El mismo principio se usa en la mayoría de las cámaras digitales comerciales, en estas el nombre de la malla de filtro es *Color Filter Array* (CFA). La diferencia entre MSFA y CFA es que la última, solo tiene

tres tipos de filtros del espectro visible: rojo (R, proveniente de *red*), verde (G, proveniente de *green*) y azul (B, proveniente de *blue*) mientras MSFA puede tener diferentes filtros en diferentes bandas [5] [6]. La principal ventaja de los sensores basados en MSFA o CFA es que, dependiendo de la naturaleza de los filtros, pueden adoptar tamaños mucho más pequeños que otras cámaras, permitiendo la portabilidad y otros beneficios específicos del tamaño. Asimismo, el costo de producción es menor y su principio de operación no cuenta con partes móviles lo que otorga al diseño más robustez y tolerancia al fallo.

Teniendo en cuenta los trabajos realizados en este campo [7], se puede considerar que existen muchas posibilidades para diseñar un sensor MSFA tanto en diseño de los filtros como en la distribución de los mismos en los píxeles. Las implementaciones de filtros multispectrales ofrecen soluciones con algunos problemas, como la falta de sensibilidad o la pérdida de luz entre los píxeles, por esta razón, es esencial tener en cuenta el posprocesamiento y considerar este problema para diseñar técnicas apropiadas que dependan del método de fabricación, diseño y distribución de las bandas en la malla del filtros.

En la Universidad de Harvard, en Estados Unidos, se ha implementado un prototipo de una cámara multispectral MSFA que usa *nanowires* para la elaboración de los filtros utilizando la misma técnica de [8]. Este prototipo posee funciones básicas debido a que se encuentra en la fase de desarrollo, es decir la investigación de esta técnica esta aun vigente. Una de las características que aun no se ha implementado es la captura de imágenes en color, es decir, la reconstrucción de las tres bandas de color R, G y B utilizando las nueve bandas presentes en el prototipo, este paso es importante para comprobar que las capacidades de este nuevo sensor sean equiparables a los sensores convencionales del mercado. Por ello, este trabajo se enfocará en diseñar un algoritmo que use la información de la cámara implementada, basada en nueve frecuencias, para estimar la información de color de manera similar a una cámara RGB convencional.

El objetivo principal de este trabajo es diseñar un algoritmo que compruebe que es posible obtener una imagen a color a partir de nueve bandas espectrales captadas por

un prototipo basado en *nanowires*. Así mismo, comparar diferentes metodologías usando algoritmos de *machine learning* para proponer una adecuada secuencia de procedimientos que obtenga mejores resultados desde perspectivas cualitativas y cuantitativas. Por último, este trabajo busca ayudar en el desarrollo de mejoras al prototipo de cámara multiespectral basada en *nanowires*, en el área de captura de imágenes a color.

Como alcances de este trabajo, el requerimiento fue priorizar el algoritmo en la calidad de reconstrucción y la cercanía de los colores a los reales y no la eficiencia, es decir, el tiempo de procesamiento no es relevante para este propósito. Si bien, el tiempo de procesamiento es un parámetro importante para un producto final de cámara multiespectral, en la fase de desarrollo del prototipo lo más importante es la correcta obtención del color.

El diseño del algoritmo está limitado a la base de datos obtenida en la universidad de Harvard, así como las limitaciones propias del prototipo, tales como: desorientación de píxeles, saturación o capacidad fotográfica. En las siguientes capítulos, primero se describe los antecedentes y conceptos necesarios para entender este trabajo, en el capítulo 3, se presenta la metodología utilizada la cual se divide en 3 etapas: preprocesamiento, obtención del modelo y posprocesamiento. Luego, en el capítulo 4, se presentan los resultados numéricos y cualitativos con su respectivo análisis. Finalmente, en el capítulo 5 se presenta las conclusiones y recomendaciones.

Capítulo 2

ANTECEDENTES

Este capítulo presenta los principales conceptos necesarios para entender esta tesis. En la primera sección se presentan los trabajos realizados en MSFA basados en CFA y se describe el prototipo con el que se realizó este trabajo. En la segunda sección se detallan los métodos de clasificación y regresión utilizados. Así mismo es importante explicar el modelo de color que se utilizó para contrastar la fidelidad de los resultados, esta información es descrita en la tercera sección. Por último, al final del capítulo se explica el concepto de imágenes de alto rango dinámico (HDR, por sus siglas en inglés).

2.1 Filtros MSFA basados en CFA

La amplia aplicación de CFA [6] en imágenes en color ha provocado un interés académico e industrial en la extensión de CFA a MSFA mediante el desarrollo de nuevos métodos para desarrollar estos filtros. Similar a un sistema de imagen en color basado en CFA, las imágenes multiespectrales que usan MSFA también necesitan el proceso de interpolación y pueden verse afectadas por aberraciones cromáticas, esta es la razón por la cual se han desarrollado diferentes técnicas de interpolación para cada modelo de MSFA que se ha planteado [7].

En [9] se presentó por primera vez un CFA modificado para la adquisición de imágenes multiespectrales. En este modelo, se utilizaron siete filtros colocados de manera hexagonal, de modo que cada píxel de una banda espectral dada esté rodeado por seis píxeles de diferentes bandas. Así mismo, se utilizó la interpolación con el algoritmo *k-Nearest Neighbors* [10]. Otro ejemplo de MSFA se presenta en [11], donde se propone un

método de diseño genérico de MSFA donde la probabilidad de que cada banda espectral aparezca en la matriz se puede presentar en un árbol binario, igualmente se propuso un método de interpolación para este método [12]. En contraste con un patrón complejo como el mencionado anteriormente, en [13], se propone un MSFA de seis bandas organizadas en 3×2 píxeles que representan grupos de rectángulos, utilizando este método, el autor utiliza la interpolación lineal por su bajo tiempo de ejecución. Otro MSFA es presentado en [14], donde se distribuyó cuatro filtros en franjas diagonales. Finalmente, en el trabajo [15], el equipo presenta un nuevo método para construir filtros multiespectrales usando *nanowires* verticales. El modo de funcionamiento de este tipo de filtro se detalla a continuación.

2.1.1 Filtros MSFA basados en *nanowires*

En [16] se demostró que los *nanowires* pueden ser utilizados como filtros de luz para obtener diferentes colores del espectro electromagnético. Un segundo estudio [17] demostró que también podían absorber colores fuera del espectro visible (NIR). Es decir, poseen la propiedad de refractar la luz en diferentes longitudes de onda. Esta característica se debe a la dependencia de la longitud de onda para atravesar espacios determinados, en este caso, el espacio que hay entre cada *nanowire*. Desde otro punto de vista, se puede plantear que la refracción depende del radio de los mismos. Agrupando *nanowires* de un mismo radio, en un área determinada, se puede generar un filtro capaz de absorber una longitud de onda específica. En la Figura 2.1 se aprecia el método utilizado en [8] para la construcción de estos filtros. En primer lugar se fabrica un grupo de *nanowires* en un wafer de silicon y se le cubre con un polímero transparente, seguido de ello se retira con una cuchilla los *nanowires* encapsulados en el polímero. Finalmente, toda la lámina creada es superpuesta encima de sensores de luz monocromáticos, permitiendo así que cada sensor filtre la frecuencia diseñada. Esta técnica permite fabricar cámaras multiespectrales de un tamaño similar al de las cámaras digitales convencionales, es decir de un tamaño mucho menor a las cámaras multiespectrales en el mercado.

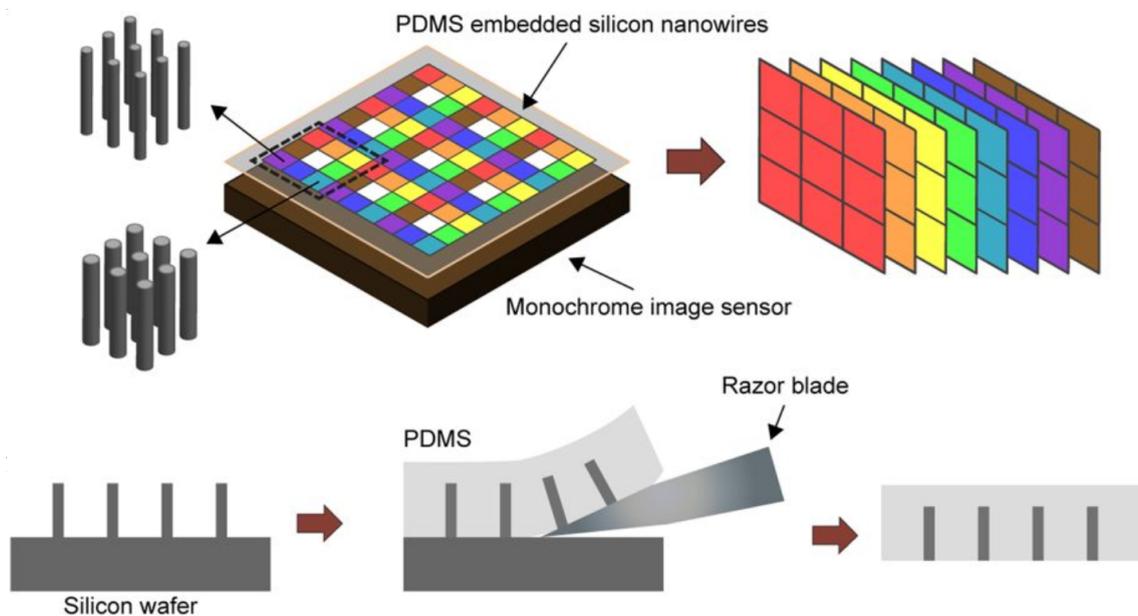


FIGURA 2.1: Esquema de construcción de los filtros de *nanowires*. En la primera etapa se construyen los *nanowires* en un wafer de silicon para luego ser cubiertos con un polímero transparente que permite el ingreso de la luz entre los *nanowires*. Finalmente los filtros son retirados utilizando una cuchilla y son adheridos a cada uno de los píxeles del sensor [8].

2.1.2 Prototipo a utilizar

En la universidad de Harvard se ha construido un prototipo de cámara multiespectral utilizando *nanowires* y el principio de [17]. El prototipo presenta nueve bandas multiespectrales, en la Figura 2.2 se aprecia un esquema de la distribución de los nueve tipos de filtros, cada uno es representado con un color diferente. A la derecha se aprecia una agrupación de 9 píxeles formando 3 filas y 3 columnas, esta región es conocida como *spectral band group* (SBG) y se caracteriza porque los 9 píxeles pertenecen al mismo tipo de filtro. En la parte central se aprecia una región de 9×9 conocida como *macropixel* (MP) y se caracteriza por ser la unidad básica de información en la cámara multiespectral, ya que presenta todos los tipos de filtros. Dependiendo de los requerimientos, una cámara multiespectral MSFA puede tener más o menos MP. En el caso del prototipo propuesto posee 18000 MP distribuidos en 1080×1350 píxeles. A esta región total de filtros que cubre el sensor se le conoce como *array filter area* (AFA).

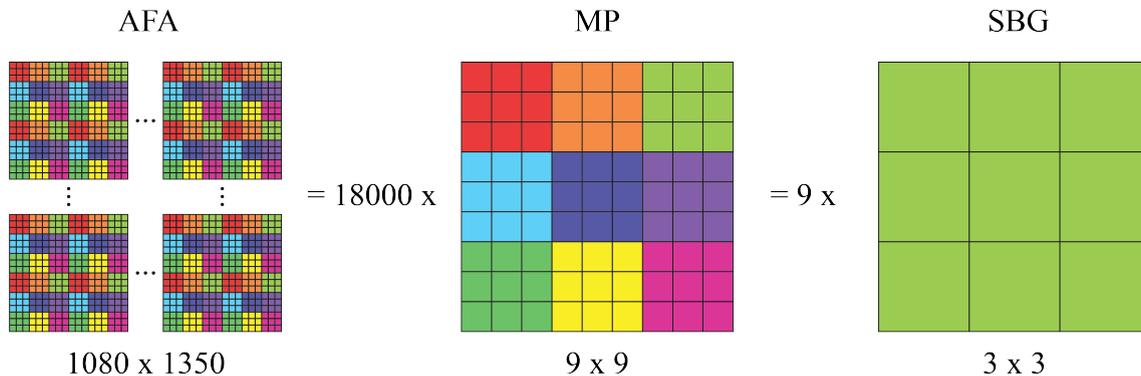


FIGURA 2.2: Distribución de los filtros de *nanowires* en el sensor. Cada uno de los colores representa un tipo de filtro agrupado en secciones de 3×3 píxeles llamado *spectral band group* (SBG). La unión de 9 diferentes SBG da a lugar a un *macro pixel* (MP) que tiene un tamaño de 9×9 píxeles. Finalmente la agrupación de MP se le conoce como array filter area (AFA), en el caso del prototipo propuesto tiene un tamaño de 1080×1350 píxeles.

El AFA del prototipo fue superpuesto en una cámara monocromática modelo DMK-27AUP031, la cual cuenta con un sensor CMOS MT9P031 de una resolución de 2592×1944 píxeles (5 MP), en donde cada uno de ellos mide $2.2 \times 2.2 \mu m$. Aunque el sensor tiene una mayor cantidad de píxeles que el AFA, los datos a analizar en este trabajo se limita por la zona multispectral capturada, es por ello que las imágenes en la base de datos tienen un tamaño de 1080×1350 píxeles. Así mismo, cabe resaltar que la resolución de las mismas es de 8 bits.

2.2 Métodos de clasificación y regresión

A continuación se explica brevemente los conceptos de los métodos de clasificación (*support vector machine* y *k-nearest neighbors*) y regresión multivariable (lineal y no lineal) que se utilizaron en este trabajo. Estos algoritmos son frecuentemente usados para entrenar modelos de clasificación y regresión de datos.

2.2.1 Algoritmo *support-vector machine*

Este es un método de aprendizaje supervisado que clasifica los datos utilizando un hiperplano. El algoritmo encuentra el hiperplano óptimo que separa los datos correctamente. En la Figura 2.3 podemos ver un esquema de este proceso, el hiperplano está representado por una línea, sin embargo se puede variar la naturaleza de este utilizando un *kernel*, en este trabajo se ha utilizado el *kernel* Gaussiano presentado en la siguiente ecuación:

$$G(x_i, x_k) = \exp(-\|x_i - x_k\|^2),$$

siendo $G(x_i, x_k)$ el valor del kernel, x_i y x_j muestras de la caracterización del modelo.

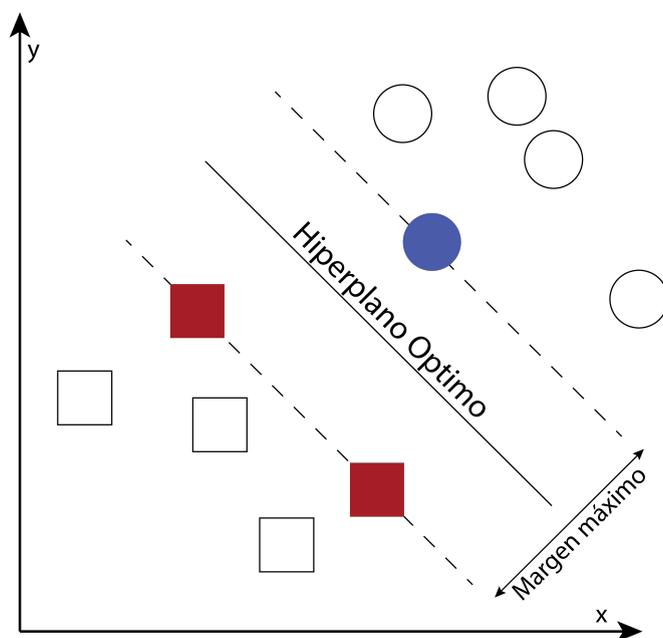


FIGURA 2.3: Esquema de clasificación utilizando SVM. El algoritmo busca el hiperplano óptimo que separa las diferentes clases optimizando el margen máximo entre estas.

2.2.2 Algoritmo *k-nearest neighbors*

Este es un método de clasificación supervisado que utiliza data previamente clasificada para predecir la clase de una nueva muestra. El algoritmo calcula la función de densidad de probabilidad para ello, es decir calcula la probabilidad de que la muestra pertenezca a una clase específica y elige la que tenga mayor valor. En la Figura 2.4 se observa un esquema de este dalgoritmo.

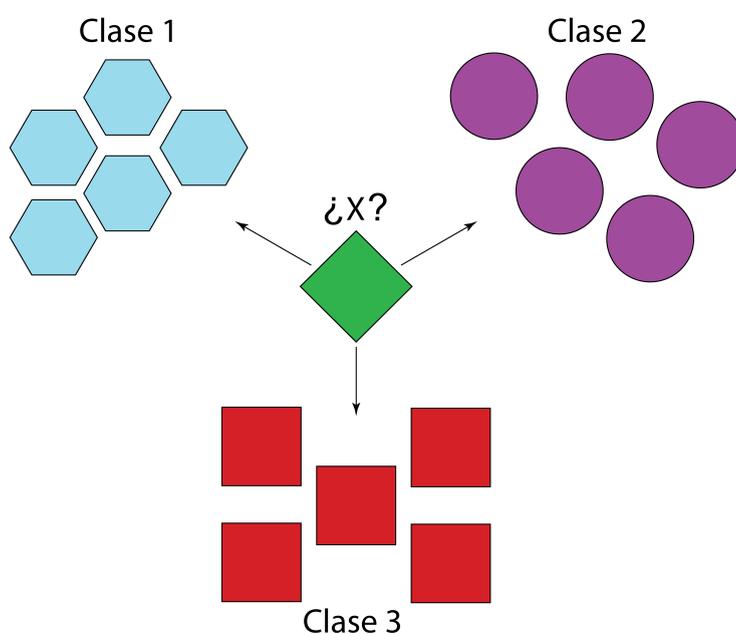


FIGURA 2.4: Esquema de clasificación utilizando *k-Nearest Neighbors*. Con un modelo de características entrenado, se predice la clase de una nueva muestra (representada con el color verde y la variable X) asumiendo la menor diferencia de valor en esas características.

2.2.3 Regresión lineal multivariable

La regresión lineal multivariable se utiliza para predecir el valor de uno o más respuestas dado un determinado grupo de datos, así mismo se puede utilizar para estimar la asociación lineal entre dos grupos de variables. La regresión lineal multivariable tiene

la siguiente forma:

$$y_i = b_0 + \sum_{j=1}^n b_j x_{ij} + e_{ij},$$

siendo y_i la respuesta deseada, b_j coeficientes, x_{ij} las variables dependientes o predictores y e_i el error para cada píxel ubicado en (i, j) . De manera matricial se vería de la siguiente forma:

$$Y = bX + E,$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ 1 & x_{31} & x_{32} & \dots & x_{3p} \\ \vdots & \vdots & \ddots & \vdots & \\ 1 & x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ \vdots \\ e_n \end{bmatrix}. \quad (2.1)$$

Para generar el modelo de predicción se debe resolver el sistema de la ecuación anterior obteniendo todos los valores de b_j y e_j .

2.2.4 Regresión no lineal multivariable

De manera similar a la regresión lineal, se busca predecir valores a partir de otros llamados predictores, la diferencia es que el modelo usado es una función no lineal. En el caso de este trabajo se utilizaron los siguientes modelos:

$$y_i = b_0 + \sum_{j=1}^n b_j x_{ij} + \sum_{j=1}^n b_k x_{ij}^2 + e_{ij} \quad (2.2)$$

y

$$y_i = b_0 + \sum_{j=1}^n b_j x_{ij} + \sum_{h=1}^n b_h x_{ij}^2 + \sum_{k=1}^n b_k x_{ij}^3 + e_{ij}. \quad (2.3)$$

El primer modelo indicado en (2.2) es la combinación de un polinomio lineal y uno cuadrático. En el segundo (2.3), combina un polinomio cuadrático y uno cúbico. El objetivo de este método es obtener los coeficientes b_0 , b_j , b_k y e_i en el entrenamiento.

2.3 Cartilla de Macbeth

La cartilla de Macbeth es un instrumento utilizado para calibrar todo tipo de cámaras a color. Cuenta con veinticuatro colores distribuidos en 4 filas y 6 columnas. La principal característica de esta cartilla es que los valores RGB de cada color son conocidos, por lo tanto una foto realizada a la cartilla debería tener esos mismos valores, por ello se utilizaran imágenes capturadas con el prototipo para entrenar y validar el modelo obtenido de reconstrucción de color. En la Figura 2.5 se muestra la cartilla de Macbeth y como están distribuidos los colores.

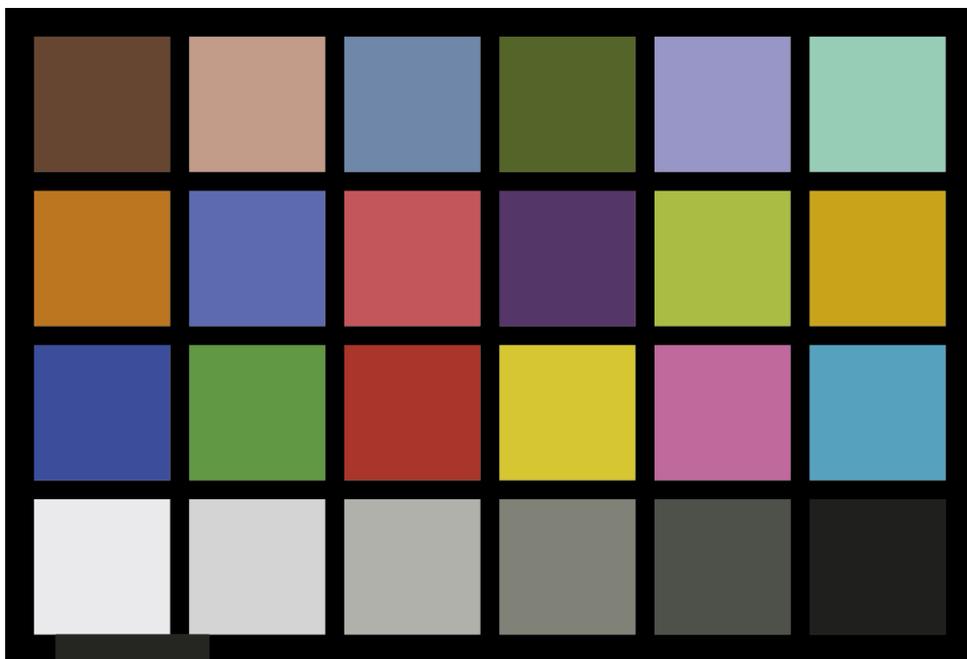


FIGURA 2.5: Representación sintética de la cartilla de Macbeth. Este es un verificador de 24 colores distribuidos en 4 filas y 6 columnas.



FIGURA 2.6: Imágenes obtenidas con diferente tiempo de exposición y su respectiva imagen HDR generada a partir de esta.

2.4 Imágenes de alto rango dinámico

Las imágenes de alto rango dinámico (HDR, por su nombre en inglés *high dynamic range*) son imágenes generadas a partir de dos o más capturas del sensor con diferente tiempo de exposición, el proceso de conversión genera que el rango dinámico aumente en la imagen, es decir, se amplía las tonalidades haciendo que las partes más oscuras y claras sean resaltadas correctamente [18]. La forma matemática en la que es obtenida cada uno

de los píxeles de una imagen HDR es representada de la siguiente manera:

$$L_{ij} = \frac{\sum_{k=1}^N \frac{f^{-1}Z_{ij}w(Z_{ij})}{\Delta t_k}}{\sum_{k=1}^N w(Z_{ij})}, \quad (2.4)$$

donde f es la función de respuesta del sensor CMOS, Z_{ij} es la intensidad en la imagen en la posición (i, j) , Δt_k es el tiempo de exposición de la imagen k , que puede tomar un valor desde 1 a N y $w(Z_{ij})$ es el factor de peso para cada imagen.

En la Figura 2.6 se aprecia un ejemplo de imagen HDR, en la izquierda, dos imágenes capturadas con diferentes tiempos de exposición que resaltan diferentes partes de la escena. La imagen superior tiene un tiempo de exposición bajo y resalta las partes más expuestas a la luz como la textura de las nubes, el edificio y parte de las plantas, en cambio la inferior tiene un tiempo de exposición alto ya que resalta las imágenes menos expuestas a la luz como por ejemplo el árbol, césped y arbustos. A la derecha se muestra la imagen HDR generada. Es evidente que la imagen de la derecha captura mejor la escena y resalta todos los elementos en una sola escena.

Capítulo 3

MÉTODOS

Este capítulo se centra en explicar detalladamente la metodología empleada para realizar la correcta reconstrucción de color de imágenes capturadas por el prototipo. Se divide en 5 secciones, en la primera sección se detalla la base de datos de imágenes disponibles, en la segunda se explica el procedimiento que se realizó a cada imagen hasta obtener su reconstrucción a color que incluye preprocesamiento, aprendizaje supervisado y posprocesamiento. Finalmente, en la última sección se explica como se utilizó las diferentes imágenes de diferentes tiempos de exposición para mejorar los resultados.

3.1 Base de datos

En este trabajo se utilizó una base de datos de imágenes capturadas por el prototipo. Esta se divide en cuatro grupos: imágenes de la cartilla de Macbeth, imágenes de fondo blanco, imágenes de botellas e imágenes de frutas. Además, en el caso de las imágenes de Macbeth, se tiene imágenes divididas en tres subgrupos: imágenes en color, imágenes en gris e imágenes capturadas con el prototipo (multiespectrales). Todas las capturas tienen un conjunto de imágenes de la misma escena con 12 tiempos de exposición diferentes que van desde 1 hasta 1/600 segundos.

En las Figuras 3.1 y 3.2 se aprecian algunos ejemplos de las imágenes de la base de datos. A la izquierda se muestran las imágenes a color obtenidas con una cámara a color con las mismas características que el prototipo. A la derecha se muestran ejemplos de imágenes obtenidas con el prototipo. Como se puede apreciar los filtros no se encuentran en toda la cámara sino en una porción de esta, la zona oscura de las imágenes es la AFA

del prototipo. Otro punto importante a mencionar es que algunas zonas de la AFA del prototipo presentan defectos ocasionados durante la unión de los filtros a los sensores de luz. Por ejemplo, el defecto más evidente se encuentra en la esquina inferior derecha, donde existen decoloraciones no uniformes en la malla de filtros. En las siguientes secciones se explicará como se lidió con esos problemas.



FIGURA 3.1: Ejemplos de imágenes de la cartilla de Mcbeth. A la izquierda una imagen capturada con una cámara a color y a la derecha, una imagen capturada por el prototipo.



FIGURA 3.2: Ejemplos de imágenes de botellas. A la izquierda una imagen capturada con una cámara a color y a la derecha, una imagen capturada por el prototipo.

3.2 Separación de muestras

Para realizar la reconstrucción, primero es necesario separar la información de cada tipo de filtro. Al inicio se tiene una AFA de longitud 1080×1350 píxeles en donde se encuentran los 9 tipos de filtros. Lo primero que se realizó es separar cada AFA en

nueve, en la que cada una solo contenga la información de un solo filtro (SBG). Como se puede apreciar el resultado son imágenes de 1080×1350 con espacios de 6 píxeles vacíos entre cada grupo de filtros de 3×3 píxeles.

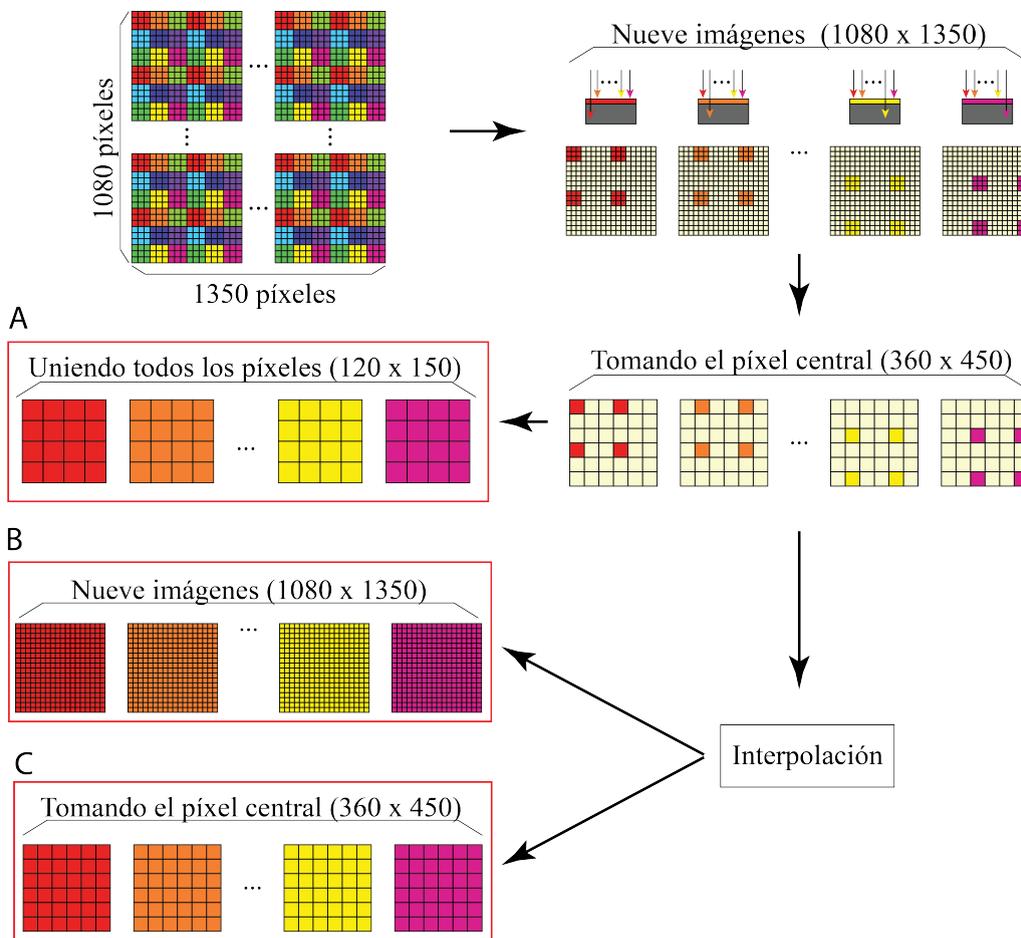


FIGURA 3.3: Esquema de separación de muestras. Al inicio, se separan los píxeles pertenecientes a un mismo tipo de filtro, seguido de ello se toma el píxel central de cada SBG para finalmente unir todos los píxeles restantes (A) o interpolar para obtener tamaños mayores de la imagen (B-C).

Debido a que el prototipo presenta un error de orientación entre los píxeles y los filtros que se originó al unir las dos partes, se optó por tomar el píxel central de cada SBG, ya que ese píxel tiene la mayor probabilidad de no presentar el error de orientación. Al realizar este procedimiento las imágenes quedan reducidas a 360×450 píxeles y los

espacios entre cada filtro ahora son dos píxeles vacíos. Existen tres maneras en las que se obtuvo las imágenes finales que se resaltan en la Figura 3.3 con un cuadro rojo. En la primera (A), se unió todos los píxeles con filtros en una imagen, eliminando los dos espacios entre cada uno, resultando una imagen de 120×150 píxeles. En las 2 siguientes, se realizó una interpolación lineal, en (B), se interpoló al tamaño original de la imagen y en (C) se interpoló al tamaño conseguido al tomar solo el píxel central de los grupos de 3×3 píxeles. El resultado final, en los tres casos, son nueve imágenes que serán denominadas como *single band array* (SBA) ya que contienen información de una sola banda.

3.3 Reconstrucción de color

El proceso de reconstrucción de color consta de tres partes: en la primera se realiza algunos procedimientos para eliminar errores generados en la construcción del prototipo, en la segunda, se utiliza los métodos de aprendizaje supervisado para calcular coeficientes que puedan representar un modelo de clasificación o regresión y por último, en la tercera, se utiliza técnicas para mejorar el resultado. Cada uno de estos pasos son explicados a detalle a continuación. Así mismo, cabe resaltar que en los esquemas y explicaciones propuestas, se representa el caso de usar la división de muestras sin interpolación (A), las otras dos formas siguen el mismo procedimiento con la excepción que el tamaño de las imágenes resultantes varía.

3.3.1 Preprocesamiento

Es necesario realizar algunos procedimientos antes de reconstruir la imagen a color para eliminar algunos problemas propios del prototipo. En la Figura 3.4 se observa un esquema de estos procedimientos y el orden en que fueron realizados. En primer lugar, la imagen con todos los filtros fue multiplicada por una matriz de corrección para corregir el valor de saturación de todos los píxeles, esta matriz también puede ser considerada la

respuesta del instrumento. Para obtenerla, se utilizó las imágenes de fondo blando de la base de datos. Se calculó la mediana de la AFA de cada una de ellas y se dividió entre cada píxel de la imagen, la matriz tiene la siguiente forma:

$$M_{ij} = \frac{\text{mediana}(W)}{W_{ij}}, \quad (3.1)$$

siendo M_{ij} la matriz de corrección que posee el mismo tamaño que la AFA del prototipo, W la AFA de la imagen perteneciente a la captura de una hoja en blanco y ij cada una de las posiciones de los píxeles en la imagen. Después de la corrección, se separó las muestras de los nueve filtros como se explicó en la sección 3.2 y se procedió a aplicar un filtro de la mediana con una ventana de 3×3 píxeles a cada una de las imágenes de diferentes frecuencias. Este procedimiento es necesario porque cada filtro no fue construido uniformemente y no todos son iguales, generando una pequeña variación entre ellos a pesar de ser del mismo tipo, este filtro eliminará valores atípicos.

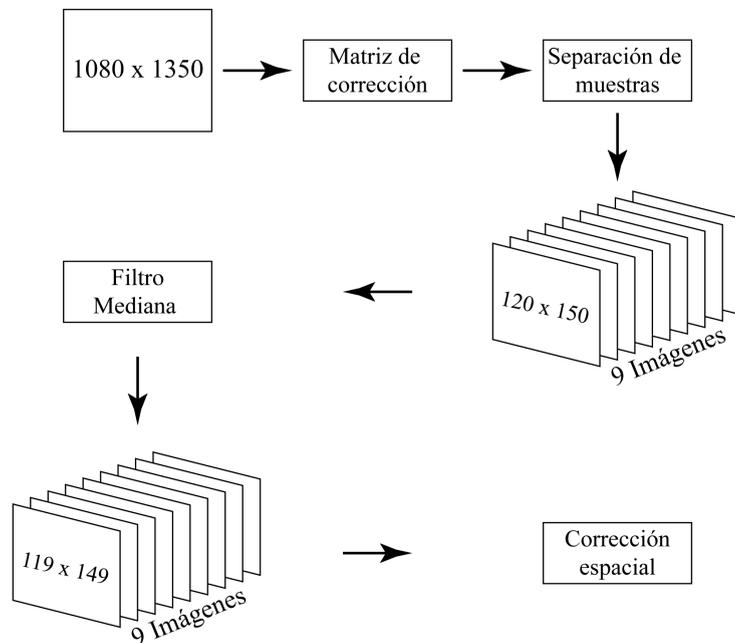


FIGURA 3.4: Esquema de preprocesamiento. La imagen original es multiplicada por una matriz de corrección y luego de ello se separan las muestras de los filtros para obtener 9 imágenes. A cada una de estas imágenes se le aplica un filtro mediana y una corrección espacial para mejorar los resultados.

El último procedimiento tiene por objetivo eliminar el error espacial que se origina al asumir que todos los píxeles de diferente frecuencia que se encuentran en una posición similar pertenecen a la una misma zona capturada. Al separar las muestras, inicialmente había 6 píxeles vacíos que separaban cada grupo de frecuencias y al unir todos esos píxeles, esos espacios omitidos generan este error. Para apaciguar los efectos de este, se procedió a mover espacialmente un píxel ortogonalmente las nueve imágenes excluyendo la quinta frecuencia ya que es la central. Este procedimiento origina que se pierda un píxel en cada lado (bordes) disminuyendo el tamaño de la imagen a 119×149 píxeles en el caso de la primera forma de separación de muestras.

3.3.2 Aprendizaje supervisado

Utilizando las nueve SBA pertenecientes a la cartilla de Macbeth obtenidas después del preprocesamiento, se entrenó los algoritmos de aprendizaje supervisado para así obtener un modelo para la reconstrucción de color. En la Figura 3.5 se observa un esquema de este procedimiento. Lo primero que se realizó fue crear una matriz de color (RGB) que contendrá los valores ideales que debería tener cada píxel, estos valores fueron obtenidos de la hoja técnica de la cartilla de Macbeth. En esta segunda matriz cada fila corresponde a un píxel y cada columna a un valor del color está definida como:

$$RGB = \begin{bmatrix} R_1 & G_1 & B_1 \\ R_2 & G_2 & B_2 \\ R_3 & G_3 & B_3 \\ \vdots & \vdots & \vdots \\ R_n & G_n & B_n \end{bmatrix},$$

donde R, G y B representan los valores rojo, verde y azul respectivamente. Lo segundo que se realizó fue poner en modo matricial la información de cada píxel, para ello se seleccionó porciones rectangulares de la imagen que contengan un color de los veinticuatro

disponibles. Cada píxel con la misma posición (i, j) en cada SBA, representado matricialmente como S^n , donde $n \in \{1, 2, 3, \dots, 9\}$, son agrupados en la misma fila para crear una matriz de entrenamiento (T) definida como:

$$T = \begin{bmatrix} S_{11}^1 & S_{11}^2 & S_{11}^3 & \dots & S_{11}^8 & S_{11}^9 \\ S_{12}^1 & S_{12}^2 & S_{12}^3 & \dots & S_{12}^8 & S_{12}^9 \\ S_{13}^1 & S_{13}^2 & S_{13}^3 & \dots & S_{13}^8 & S_{13}^9 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ S_{ij}^1 & S_{ij}^2 & S_{ij}^3 & \dots & S_{ij}^8 & S_{ij}^9 \end{bmatrix}.$$

Estas dos matrices se utilizaron para entrenar un modelo, en el caso de k -NN y SVM. Por otro lado, en el caso de regresión lineal, se agregó una columna de “1” a la matriz T, esta representa el coeficiente independiente.

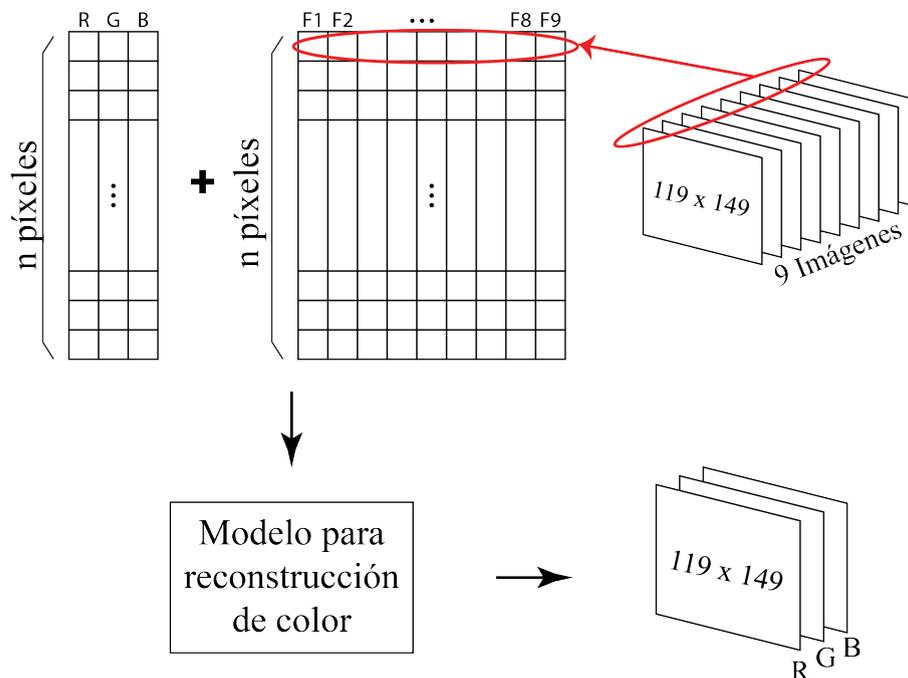


FIGURA 3.5: Esquema de la obtención del modelo. Utilizando la información de los píxeles en forma matricial y los valores ideales en color de estos se entrenó un modelo en los casos de k -NN y SVM. En el caso de

Esta nueva matriz junto con la matriz RGB se utilizaron para obtener los coeficientes presentados en la ecuación 2.1. Por último para regresión no lineal se utilizó un modelo cuadrático (ecuación 2.2) y un modelo cúbico (ecuación 2.3) para la obtención de los coeficientes. Utilizando el modelo o los coeficientes se procedió a obtener los valores R,G y B de cada píxel y así obtener tres imágenes de tamaño 119×149 .

3.3.3 Posprocesamiento

El último paso en la reconstrucción de color es el posprocesamiento para mejorar los resultados. En primer lugar, se eliminó los valores atípicos de cada banda de color. Debido a que la cámara tiene una resolución de 8 bits, los valores oscilan en un rango entre 0 y 255, por tal motivo valores negativos o mayores a 255 son considerados atípicos o saturados. Para eliminar estos valores se reemplazó cualquier píxel negativo por 0 y cualquier valor mayor a 255 por este mismo valor. El segundo procedimiento que se realizó fue un filtro de la mediana con una ventana de 3×3 píxeles. Este filtro tiene por objetivo suavizar la imagen y que no haya cambios bruscos de color entre píxeles adyacentes. Un esquema de todo el posprocesamiento se muestra en la Figura 3.6.

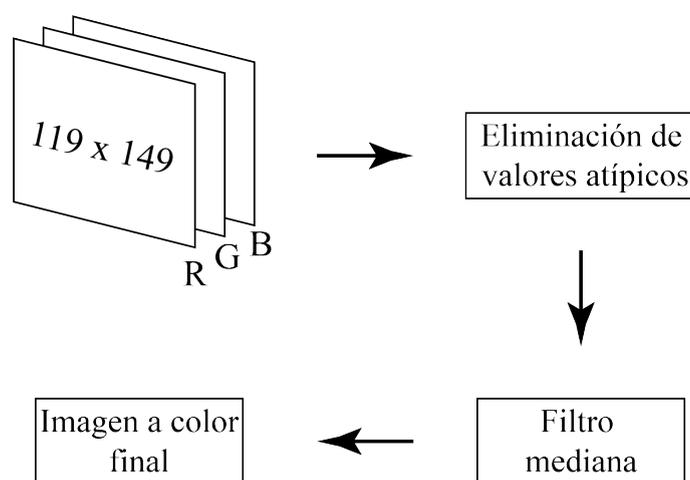


FIGURA 3.6: Esquema de posprocesamiento. Se realizan dos procedimientos adicionales a la imagen a color obtenida: eliminación de valores atípicos y un filtro mediana con una ventana de 3×3 píxeles.

3.4 Reconstrucción utilizando HDR

Una variación de la reconstrucción de color es la introducción de la utilización de imágenes de alto rango dinámico (HDR, por su nombre en inglés *high dynamic range*) en el procedimiento. Esto se puede realizar de tres maneras diferentes, en la primera se utiliza HDR antes de separar las muestras en SBA, en la segunda se utiliza después de la reconstrucción de color y finalmente en la última se utiliza HDR después de separar las muestras. El procedimiento realizado en cada caso se explica en las siguientes secciones.

3.4.1 Primera forma de obtener HDR (Tipo 1)

En la Figura 3.7 se encuentra un esquema de la primera forma en la que se introdujo las imágenes HDR a la reconstrucción de color. En este caso se realiza antes de la separación de filtros. Se tomó AFA de diferentes tiempos de exposición (representados matricialmente por A^n donde n es el número de tiempos de exposición seleccionados) para utilizar HDR y así obtener una sola AFA con un mayor rango dinámico.

Algoritmo 1: Primera forma de obtener HDR (Tipo 1)

Input: AFA de las imágenes multiespectrales representadas por A^n , donde n es el número de tiempos de exposición de una muestra de imágenes.

Output: Imagen a color.

Start

HDR \leftarrow Obtener HDR usando A^n ;

HDR \leftarrow Preprocesamiento{HDR};

C \leftarrow Obtener imagen a color usando HDR ;

C \leftarrow Posprocesamiento{C};

Retorna C;

End

Esta nueva AFA es procesada con normalidad con los métodos explicados en secciones anteriores, es decir, preprocesamiento, entrenamiento y posprocesamiento. En el algoritmo 1 se encuentra en forma de pseudocódigo la implementación de este proceso.

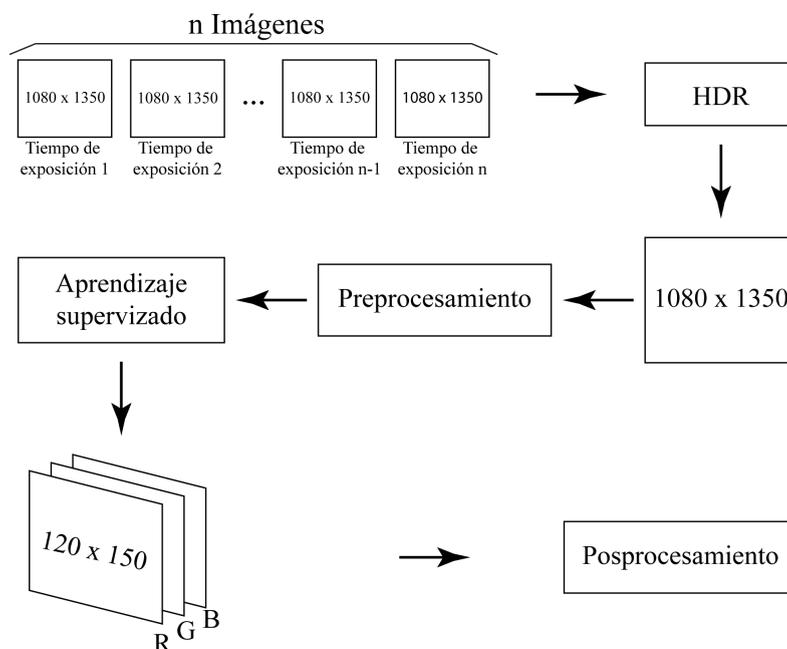


FIGURA 3.7: Esquema de la primera forma de introducir HDR en la reconstrucción de color. En este procedimiento se utilizan imágenes de diferente tiempo de exposición capturadas con el prototipo para generar una de alto rango dinámico. Con esta nueva imagen se obtiene una a color por el procedimiento ya explicado en secciones anteriores.

3.4.2 Segunda forma de obtener HDR (Tipo 2)

La segunda forma de introducir HDR es hacerlo al final de la reconstrucción de color, para ello, se obtiene el color de AFAs de diferentes tiempos de exposición y con todos los resultados se obtiene una nueva imagen utilizando HDR, finalmente a esa imagen se le realiza posprocesamiento. En la Figura 3.8 se aprecia un esquema de este procedimiento. En el algoritmo 2 se encuentra en forma de pseudocódigo la implementación de este proceso.

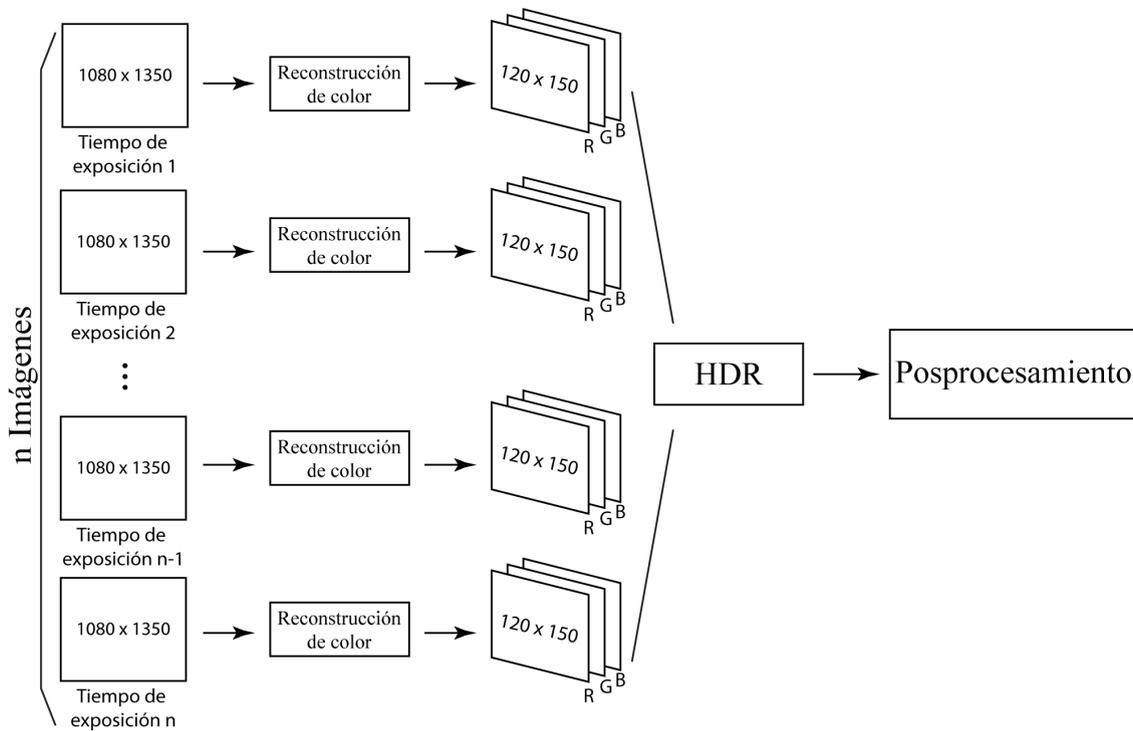


FIGURA 3.8: Esquema de la segunda forma de introducir HDR en la reconstrucción de color. Utilizando diferentes imágenes a color obtenidas a partir de otras de diferente tiempo de exposición se obtiene una de alto rango dinámico. A esta imagen a color se le aplica el posprocesamiento para obtener el resultado final.

Algoritmo 2: Segunda forma de obtener HDR (Tipo 2)

Input: Imágenes multispectrales representadas por A^n , donde n es el número de tiempos de exposición de una muestra de imágenes.

Output: Imagen a color.

Start

para i en n **hacer**

$p \leftarrow$ Preprocesamiento $\{A^i\}$;
 $c \leftarrow$ Obtener imagen a color usando p ;
 $C^i \leftarrow$ Almacenar c ;

$HDR \leftarrow$ Calcular HDR usando C^n ;

$HDR \leftarrow$ Posprocesamiento $\{HDR\}$;

Retorna HDR;

End

3.4.3 Tercera forma de obtener HDR (Tipo 3).

Esta última forma es la más compleja de las tres debido a que se aplica HDR después de la separación de los filtros en SBAs. En la Figura 3.9 se encuentra un esquema de los pasos realizados. En primer lugar, se obtienen los SBAs de cada una de las imágenes de diferente tiempo de exposición, seguido de ello se ordenan por tipo de filtro y se realiza HDR en cada grupo para obtener así una imagen de alto rango dinámico por filtro. Utilizando estas nueve imágenes HDR, se procedió a la reconstrucción de color y el posprocesamiento con normalidad. En el algoritmo 3 se encuentra en forma de pseudocódigo la implementación de este proceso.

Algoritmo 3: Tercera forma de obtener HDR (Tipo 3)

Input: Imágenes multiespectrales representadas por A^n , donde n es el número de tiempos de exposición de una muestra de imágenes; matriz RGB.

Output: Imagen a color.

Start

para i en n **hacer**

$M_m^i \leftarrow$ Separar muestras de $\{A^i\}$, donde M_m^i contiene los nueve SBA representados por $m \in \{1, 2, \dots, 9\}$;

para j en m **hacer**

$Z_m \leftarrow$ Almacenar los SBA del mismo tipo m de M_m^i ;

$HDR^m \leftarrow$ Calcular HDR usando Z_m ;

$HDR^m \leftarrow$ Filtro y corrección espacial de HDR^m para terminar preprocesamiento ;

$C \leftarrow$ Obtener imagen a color usando HDR^m ;

$C \leftarrow$ Posprocesamiento $\{C\}$;

Retorna C ;

End

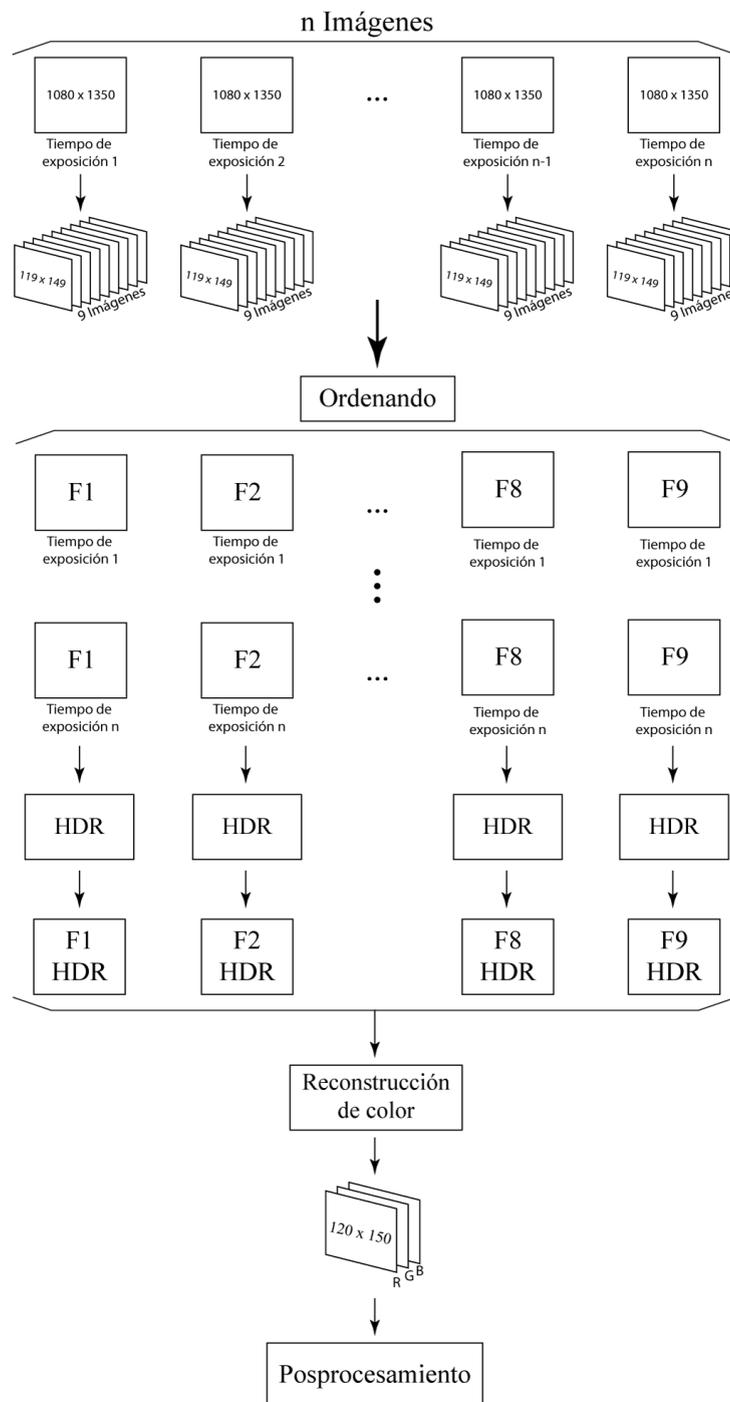


FIGURA 3.9: Esquema de la tercera forma de introducir HDR en la reconstrucción de color. Al inicio se separan las muestras de imágenes de diferente tiempo de exposición y las bandas del mismo tipo son utilizadas para generar una imagen HDR. Al final de este procedimiento se obtienen 9 imágenes de alto rango dinámico que son utilizadas para obtener una a color con los procedimientos ya explicados.

Capítulo 4

RESULTADOS Y DISCUSIÓN

En el siguiente capítulo se detalla los resultados obtenidos utilizando la metodología ya explicada, para ello se realizaron dos tipos de prueba: cuantitativas y cualitativas. En la primera, se utilizó error cuadrático medio y similitud estructural, por otro lado en la segunda, prueba de percepción visual fueron empleadas. Estos dos procedimientos se detallan a continuación.

4.1 Pruebas de error numéricas

Para las pruebas cuantitativas se utilizó las imágenes de la cartilla de Macbeth comparadas con los valores ideales de esta. En este trabajo se utilizó dos métodos, el primero de ellos es el error cuadrático medio (MSE, por sus siglas en inglés). El MSE mide el promedio de la diferencia numérica que existe entre el valor reconstruido y el valor real que se debería obtener por cada píxel, al cuadrado. Matemáticamente se define:

$$e = \frac{1}{nm} \sum_{i=1}^n \sum_{j=1}^m (\hat{Y}_{ij} - Y_{ij})^2, \quad (4.1)$$

siendo e el error cuadrático medio, \hat{Y}_{ij} el valor del píxel ideal y Y_{ij} el valor estimado, n el número de filas y m el número de columnas. El segundo método numérico que se realizó fue la diferencia estructural (SSIM, por sus siglas en inglés), el cual se define en la siguiente ecuación:

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_1)}, \quad (4.2)$$

siendo y y x porciones de la imagen ideal y la reconstruida respectivamente, μ_y y μ_x representan el promedio de intensidades presentes en y y x , así mismo, σ_y y σ_x representan la varianza, por último c_1 y c_2 representan constantes. Este método mide que tan parecidas son dos imágenes en estructura y forma sin importar los valores numéricos de los píxeles. El valor de SSIM normalizado varía entre 0 y 1, siendo 0 imágenes totalmente diferentes y 1, idénticas. Una explicación más detallada del método es presentada en [19]. El MSE y SSIM es calculado por separado en cada banda de color (R, G, B) y se toma el promedio de los tres como métrica de desempeño.

Toda la metodología incluyendo las pruebas numéricas fueron implementadas en el entorno de Matlab. Estas se dividen en dos partes, en la primera se elaboran todos los test sin utilizar imágenes HDR y en la segunda parte, las pruebas son en su totalidad utilizando imágenes HDR. Por último, cabe mencionar que se utilizará una simbología para simplificar los métodos de regresión en las tablas, siendo la regresión lineal (Lineal), regresión no lineal de segundo orden (NL1) y regresión no lineal de tercer orden (NL2).

4.1.1 Pruebas cuantitativas sin utilizar HDR

Las pruebas sin utilizar HDR se dividen en dos partes: en la primera se reconstruyó el color de todas las muestras disponibles, en la segunda, se comparó los tres métodos de interpolación para los dos mejores y el peor resultado de la prueba anterior. Para la primera prueba se realizó la reconstrucción de color en cada una de las doce imágenes de diferentes tiempos de exposición para cada uno de los cinco métodos. Los resultados numéricos de esta prueba son mostrados en las Tablas 4.1 y 4.2. Según el MSE, los tiempos de exposición con mejores resultados en todos los métodos fueron 1/10 s y 1/20 s, seguidos de 1/5 s y 1/40 s. Así mismo los más bajos errores se obtuvieron con SVM y k -NN, seguido de regresión no lineal de tercer orden. En el caso de SSIM, los tiempos de exposición entre 1/10 y 1/40 mostraron los mejores resultados.

TABLA 4.1: MSE promedio de las tres bandas de color (R, G, B) usando (4.1). Se incluyen todos los tiempos de exposición de imágenes obtenidas al capturar la cartilla de Macbeth usando el prototipo para cada uno de los métodos usados y un tamaño de 120×150 píxeles.

Tiempo de exposición (s)	SVM	k -NN	Lineal	NL 1	NL 2
1	201.11	561.34	2861.2	2606.5	2463.8
1/5	16.243	91.204	1263.1	916.90	792.31
1/10	5.9621	12.230	497.48	404.26	369.67
1/20	49.8319	29.648	606.63	520.12	478.76
1/40	42.990	52.040	892.24	802.10	714.65
1/60	52.027	127.03	1158.5	1053.8	939.50
1/80	54.592	172.81	1386.8	1257.9	1142.3
1/100	97.744	287.51	1540.2	1407.3	1290.4
1/150	174.601	361.69	1922.4	1766.3	1620.4
1/200	247.893	491.66	2230.4	2022.5	1854.7
1/400	357.065	601.07	2786.1	2455.9	2273.0
1/600	375.498	729.52	2972.2	2652.4	2435.1

TABLA 4.2: SSIM promedio de las tres bandas de color (R, G, B) usando (4.2). Se incluyen todos los tiempos de exposición de imágenes obtenidas al capturar la cartilla de Macbeth usando el prototipo para cada uno de los métodos usados y un tamaño de 120×150 píxeles.

Tiempo de exposición (s)	SVM	k -NN	Lineal	NL 1	NL 2
1	0.131	0.163	0.185	0.034	0.045
1/5	0.116	0.154	0.356	0.098	0.105
1/10	0.144	0.170	0.492	0.543	0.549
1/20	0.168	0.143	0.465	0.5215	0.530
1/40	0.110	0.139	0.349	0.310	0.348
1/60	0.154	0.175	0.245	0.184	0.291
1/80	0.135	0.146	0.294	0.152	0.286
1/100	0.178	0.192	0.245	0.127	0.248
1/150	0.128	0.112	0.131	0.104	0.183
1/200	0.077	0.083	0.138	0.097	0.126
1/400	0.291	0.120	0.104	0.041	0.119
1/600	0.027	0.068	0.101	0.029	0.095

TABLA 4.3: MSE promedio de las tres bandas de color (R, G, B) usando (4.1). Se incluyen los dos mejores y el peor tiempo de exposición basado en los resultados de las Tablas 4.1 y 4.2 para cada tipo de separación de muestras.

Tiempo E. (s)	Método	120x150 píxeles	360x450 píxeles	1080x1350 píxeles
1/10	SVM	5.9621	5.291	6.285
1/20	SVM	49.831	41.954	45.941
1/600	SVM	375.49	376.12	395.24
1/10	<i>k</i> -NN	12.230	11.412	12.542
1/20	<i>k</i> -NN	29.648	18.388	18.531
1/600	<i>k</i> -NN	29.648	74.428	79.284
1/10	Lineal	497.48	489.46	552.24
1/20	Lineal	606.63	593.93	674.77
1/600	Lineal	2972.2	2985.5	3013.8
1/10	NL 1	404.26	408.05	410.28
1/20	NL 1	520.12	521.29	520.41
1/600	NL 1	2652.4	2655.6	2642.4
1/10	NL 2	369.67	375.15	381.34
1/20	NL 2	478.76	480.01	481.03
1/600	NL 2	2435.1	2412.6	2304.1

Cabe resaltar que en el caso de SVM y *k*-NN los valores SSIM son bastante bajos en todos los tiempos de exposición y esto junto con los resultados en MSE nos dan un indicio que estos métodos reconstruyen mejor el color original pero no la forma de la imagen, caso contrario de las regresiones. Por otro lado, es importante notar que los tiempos de exposición extremos no presentaron buenos resultados, esto se puede explicar debido a que esas imágenes son las más oscuras o en caso contrario, las más iluminadas de todas, así mismo se puede notar que los resultados mejoran gradualmente mientras los tiempos de exposición se acercan más a 1/10 s y 1/20 s, indicando que la intensidad de luz capturada por el sensor debido a los diferentes tiempos de exposición influyen en los resultados.

La segunda prueba se realizó para determinar como varían los resultados cuando se utilizan diferentes métodos de separación de muestras, es decir cuando el tamaño de

la imagen aumenta. Para ello se eligió los dos mejores tiempos de exposición (1/10 s y 1/20 s) y el peor (1/600 s) de la primera prueba para compararlos. En la Tabla 4.3 y 4.4 se pueden apreciar los resultados obtenidos. En el caso de MSE, las imágenes de tamaño 360×450 píxeles mostraron los mejores resultados para SVM y k -NN, algo diferente ocurre en las regresiones, ya que los resultados son similares para los 3 tamaños. Por otro lado, en SSIM, en los dos casos que la imagen incrementa de tamaño, el error aumenta, se puede inferir que esto se debe a que la interpolación genera un ruido que afecta la estructura de la imagen. Así mismo es importante mencionar que los mejores resultados en MSE siguen presentándose en SVM, k -NN y los mejores resultados en SSIM, siguen presentándose en las regresiones.

En la Figura 4.1 se aprecian las mejores y peores imágenes a color obtenidas con esta primera prueba. En ésta se puede notar que los colores reconstruidos con los métodos SVM y k -NN son casi idénticos a los de Macbeth, sin embargo la estructura de la cartilla de Macbeth se muestra un poco difusa, por otro lado, las imágenes de regresión lineal y no lineal presentan una cartilla de Macbeth más orgánica y uniforme. En la Figura 4.2, se aprecia las imágenes reconstruidas utilizando el segundo tipo de separación de muestras, los resultados son similares a los anteriores, con la excepción que las imágenes presentan mejor resolución. Estas dos figuras comprueban porqué las regresiones presentan un menor error en SSIM, pero mayor error en MSE. A pesar que SVM y k -NN presenten los colores más cercanos a los originales, la estructura de la cartilla de Macbeth es menos parecida. Por ello, en estas primeras pruebas se considera que los mejores resultados se obtuvieron utilizando regresión no lineal de segundo y tercer orden con un tamaño de 360×450 y utilizando la imagen de tiempo de exposición 1/10 s.

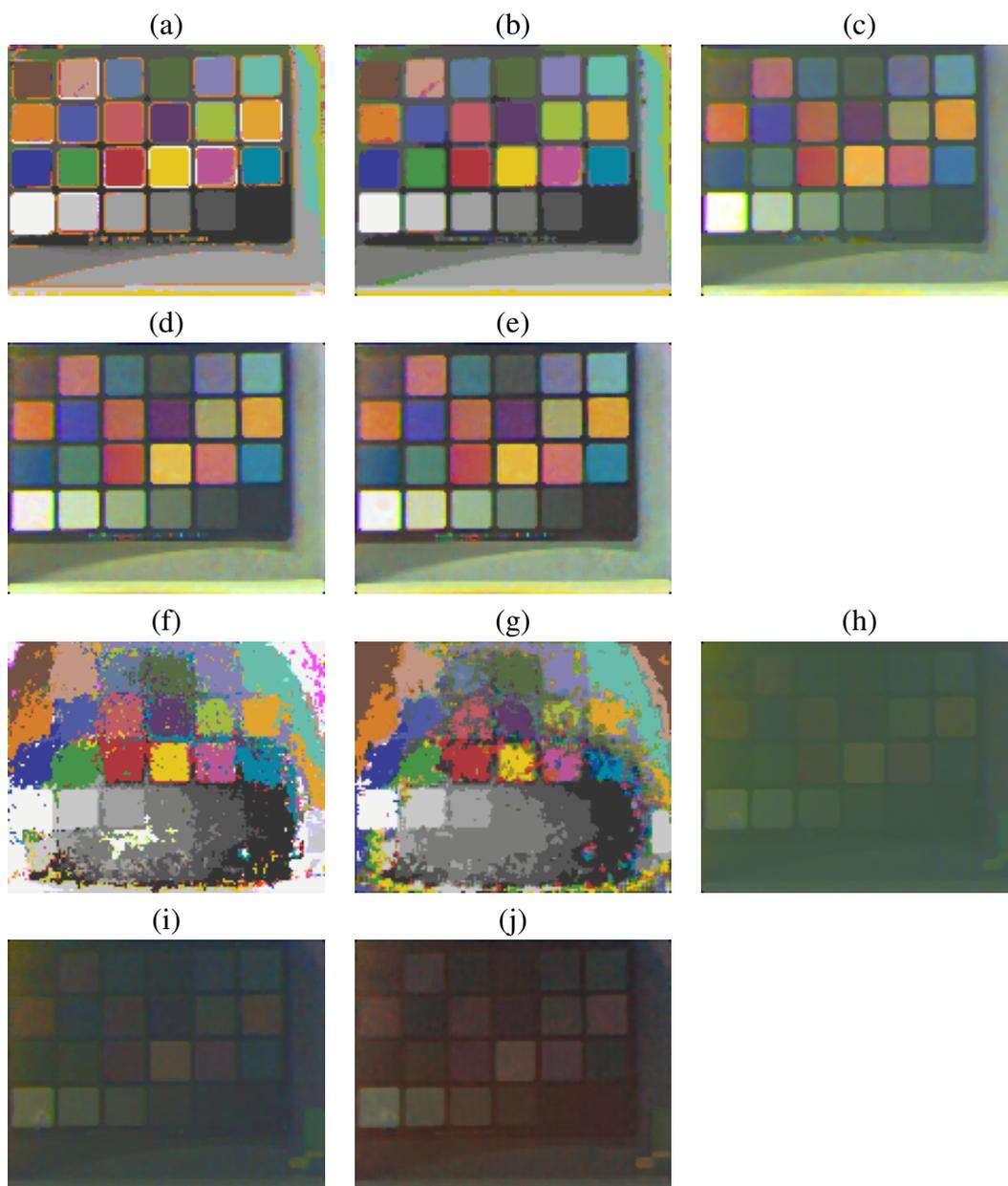


FIGURA 4.1: Imágenes de la cartilla de Macbeth para diferentes métodos. (a) Mejor resultado en SVM. (b) Mejor resultado en k -NN. (c) Mejor resultado en regresión lineal. (d) Mejor resultado en regresión de segundo orden. (e) Mejor resultado en regresión de tercer orden. (f) Peor resultado en SVM. (g) Peor resultado en k -NN. (h) Peor resultado en regresión lineal. (i) Peor resultado en regresión de segundo orden. (j) Peor resultado en regresión de tercer orden.

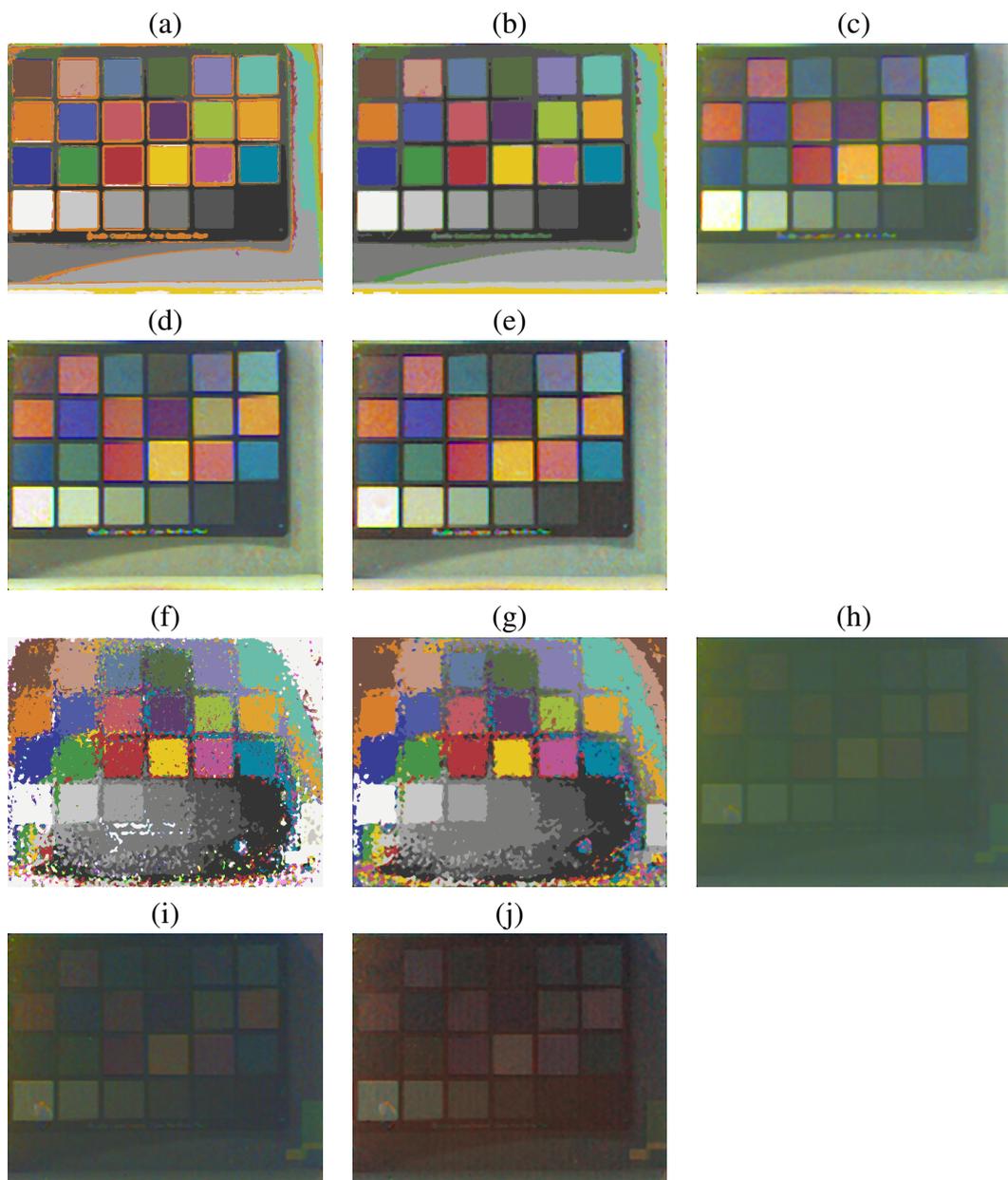


FIGURA 4.2: Imágenes de la cartilla de Macbeth para diferentes métodos usando el segundo tipo de interpolación (360x450 píxeles). (a) Mejor resultado en SVM. (b) Mejor resultado en k -NN. (c) Mejor resultado en regresión lineal. (d) Mejor resultado en regresión de segundo orden. (e) Mejor resultado en regresión de tercer orden. (f) Peor resultado en SVM. (g) Peor resultado en k -NN. (h) Peor resultado en regresión lineal. (i) Peor resultado en regresión de segundo orden. (j) Peor resultado en regresión de tercer orden.

TABLA 4.4: SSIM promedio de las tres bandas de color (R, G, B) usando (4.2). Se incluyen los dos mejores y el peor tiempo de exposición basado en los resultados de las Tablas 4.1 y 4.2.

Tiempo E. (s)	Método	120x150 píxeles	360x450 píxeles	1080x1350 píxeles
1/10	SVM	0.144	0.138	0.124
1/20	SVM	0.168	0.141	0.124
1/600	SVM	0.027	0.021	0.014
1/10	<i>k</i> -NN	0.170	0.151	0.112
1/20	<i>k</i> -NN	0.143	0.198	0.124
1/600	<i>k</i> -NN	0.068	0.019	0.037
1/10	Lineal	0.492	0.316	0.202
1/20	Lineal	0.465	0.221	0.204
1/600	Lineal	0.101	0.085	0.054
1/10	NL 1	0.543	0.307	0.298
1/20	NL 1	0.5215	0.399	0.385
1/600	NL 1	0.029	0.011	0.001
1/10	NL 2	0.549	0.311	0.312
1/20	NL 2	0.530	0.302	0.295
1/600	NL 2	0.095	0.025	0.012

4.1.2 Pruebas cuantitativas utilizando HDR

La segunda parte de las pruebas cuantitativas se divide en tres, en cada parte se comprueba cada una de las formas de utilizar HDR en la reconstrucción de color explicadas en la sección 3.4. Las pruebas se realizaron para seis combinaciones de imágenes de tiempos de exposición: [1/5 s, 1/10 s], [1/10 s, 1/20 s], [1/20 s, 1/40 s], [1/5 s, 1/10 s, 1/20 s], [1/10 s, 1/20 s, 1/40 s] y [1/5 s, 1/10 s, 1/40 s]. Estas combinaciones fueron elegidas teniendo en cuenta los mejores resultados de las primeras pruebas.

En la Tabla 4.5 se muestran los resultados de la primera forma de utilizar HDR en la reconstrucción de color, en esta se puede apreciar que en el caso de MSE, en promedio en todos los métodos se mostraron mejores resultados en la combinación de tiempo de exposición [1/5 s, 1/10 s, 1/20 s], sin embargo solo tomando en cuenta las regresiones,

mejores resultados se ven en la combinación [1/5 s, 1/10 s], destacando regresión Lineal de tercer orden con un valor de 335.98. Así mismo se puede notar que existe una mejora en promedio comparado con los resultados sin utilizar HDR. En el caso de SSIM, los resultados son muy parecidos para todas las combinaciones indicando que la exposición de la luz es un factor que influye en poca medida en la estructura de la imagen. En la Figura 4.3 se muestran las imágenes para cada uno de los métodos propuestos, como se observa, son bastante similares a las obtenidas sin HDR, sin embargo existe un cambio mínimo que se puede comprobar con los resultados numéricos.

Los resultados numéricos de la segunda forma de utilizar HDR se muestran en la Tabla 4.6. En este caso, los mejores resultados se obtuvieron en [1/10 s, 1/20 s] seguido de [1/10 s, 1/20 s, 1/40 s] para las regresiones, así mismo para SVM y k -NN los mejores resultados se obtuvieron en [1/20 s, 1/40 s] y [1/10 s, 1/20 s, 1/40 s]. Contrariamente al método anterior, existe un incremento de error en promedio en MSE comparado con la reconstrucción sin usar HDR, es decir esta segunda forma de reconstrucción no parece mostrar buenos resultados. En el caso de SSIM existen variaciones pequeñas en promedio, pero no son significativas y los mejores resultados se mantienen alrededor de 0.55. En la Figura 4.3 se muestran los resultados de los cinco métodos. En el que se puede encontrar una diferencia distinguible, comparado con los obtenidos sin utilizar HDR, es en SVM y k -NN, en los que se aprecia que los recuadros de colores presentan pequeñas manchas. En los otros casos las imágenes son bastante similares. La última prueba cuantitativa se ve plasmada en la Tabla 4.7 y la Figura 4.5. En el caso de SVM y k -NN en MSE los resultados están bastante parecidos y no hay una combinación que destaque. Por otro lado en las regresiones, los mejores resultados en MSE se obtuvieron, en las combinaciones de tiempo de exposición de [1/5 s, 1/10 s] mostrando los menores valores hasta ahora, destacando regresión no lineal de tercer orden con 261.47 mejorando en 29.3 % el mejor resultado sin usar HDR. Por otro lado en SSIM, todas las combinaciones presentan resultados similares, destacándose [1/10 s, 1/20 s].

TABLA 4.5: Resultados de la primera forma de usar imágenes de alto rango dinámico en la reconstrucción de color para seis diferentes combinaciones de tiempos de exposición en cada uno de los métodos. Valores promedio de las tres bandas de color (R, G, B).

Tiempo E. (s)	Métrica	<i>k</i> -NN	SVM	Lineal	NL1	NL2
[1/5, 1/10]	MSE	26.877	28.289	466.22	366.39	335.98
[1/10, 1/20]	MSE	21.952	19.548	501.46	413.02	380.34
[1/20, 1/40]	MSE	27.107	22.396	681.51	594.77	536.35
[1/5, 1/10, 1/40]	MSE	27.104	19.4245	569.92	471.58	438.66
[1/5, 1/10, 1/20]	MSE	28.833	19.1248	479.54	386.18	358.88
[1/10, 1/20, 1/40]	MSE	23.371	19.305	572.90	486.63	443.70
[1/5, 1/10]	SSIM	0.1800	0.1837	0.5580	0.5541	0.5248
[1/10, 1/20]	SSIM	0.1486	0.1835	0.5525	0.5392	0.5474
[1/20, 1/40]	SSIM	0.1737	0.1731	0.5510	0.5193	0.5589
[1/5, 1/10, 1/40]	SSIM	0.1720	0.1853	0.5480	0.5380	0.5444
[1/5, 1/10, 1/20]	SSIM	0.1668	0.1839	0.5529	0.4856	0.5532
[1/10, 1/20, 1/40]	SSIM	0.1854	0.1923	0.5509	0.4633	0.5387

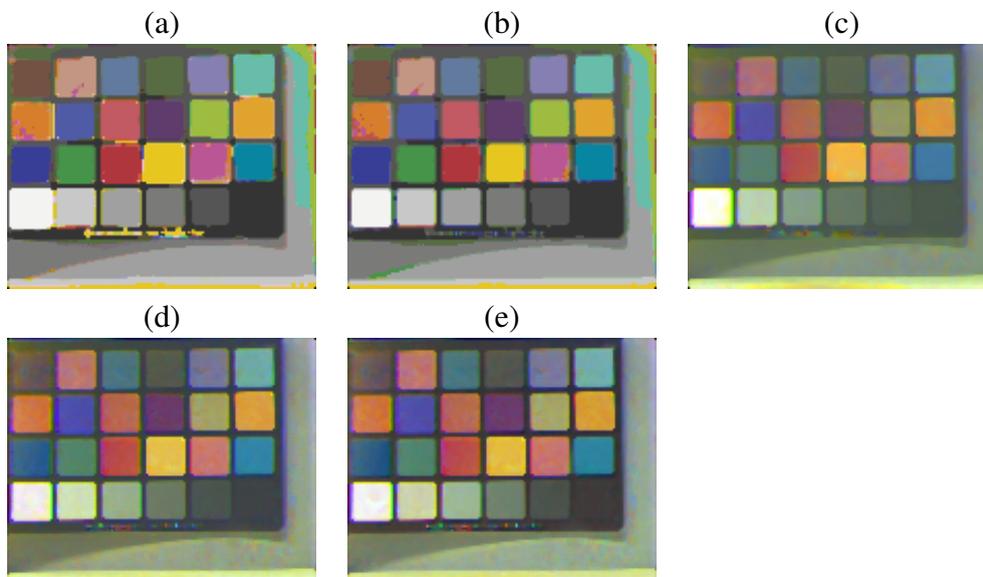


FIGURA 4.3: Imágenes de la cartilla de Macbeth para diferentes métodos usando la primera forma de aplicar HDR. (a) Mejor resultado en SVM. (b) Mejor resultado en *k*-NN. (c) Mejor resultado en regresión lineal. (d) Mejor resultado en regresión de segundo orden. (e) Mejor resultado en regresión de tercer orden.

TABLA 4.6: Resultados de la segunda forma de usar imágenes de alto rango dinámico en la reconstrucción de color para seis diferentes combinaciones de tiempos de exposición en cada uno de los métodos. Valores promedio de las tres bandas de color (R, G, B).

Tiempo E. (s)	Métrica	<i>k</i> -NN	SVM	Lineal	NL1	NL2
[1/5, 1/10]	MSE	211.80	201.56	1059.9	767.10	702.70
[1/10, 1/20]	MSE	63.080	56.855	528.14	512.98	473.26
[1/20, 1/40]	MSE	35.919	27.036	562.80	654.70	593.04
[1/5, 1/10, 1/40]	MSE	118.33	107.48	972.50	707.90	640.30
[1/5, 1/10, 1/20]	MSE	131.88	122.38	900.10	647.60	595.56
[1/10, 1/20, 1/40]	MSE	44.332	38.483	696.85	578.51	526.18
[1/5, 1/10]	SSIM	0.1453	0.1633	0.5533	0.5257	0.5511
[1/10, 1/20]	SSIM	0.1500	0.1770	0.5217	0.5441	0.5293
[1/20, 1/40]	SSIM	0.1534	0.1090	0.5467	0.5033	0.5404
[1/5, 1/10, 1/40]	SSIM	0.1497	0.1221	0.6576	0.5294	0.5324
[1/5, 1/10, 1/20]	SSIM	0.1359	0.1651	0.4575	0.5344	0.5557
[1/10, 1/20, 1/40]	SSIM	0.1631	0.1360	0.5577	0.5451	0.5527

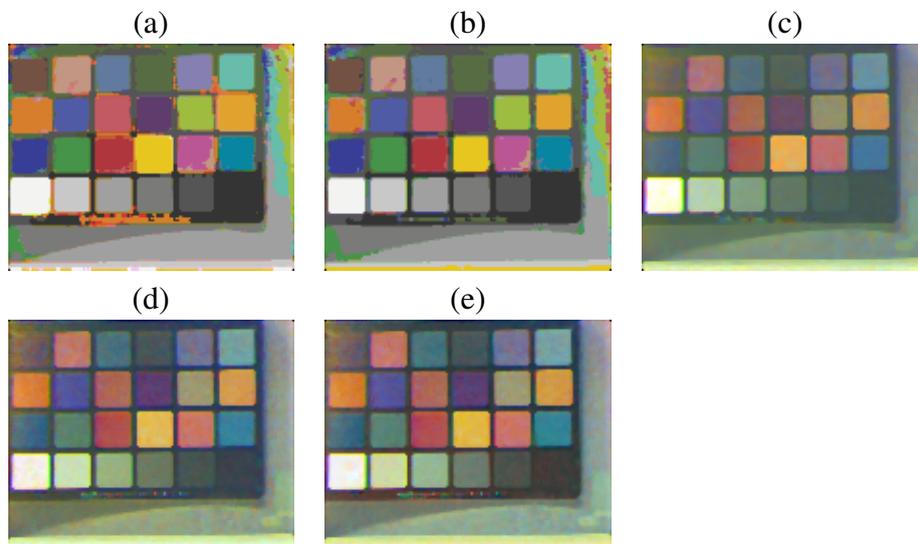


FIGURA 4.4: Imágenes de la cartilla de Macbeth para diferentes métodos usando la segunda forma de aplicar HDR. (a) Mejor resultado en SVM. (b) Mejor resultado en *k*-NN. (c) Mejor resultado en regresión lineal. (d) Mejor resultado en regresión de segundo orden. (e) Mejor resultado en regresión de tercer orden.

TABLA 4.7: Resultados de la tercera forma de usar imágenes de alto rango dinámico en la reconstrucción de color para seis diferentes combinaciones de tiempos de exposición en cada uno de los métodos. Valores promedio de las tres bandas de color (R, G, B).

Tiempo E. (s)	Métrica	<i>k</i> -NN	SVM	Lineal	NL1	NL2
[1/5, 1/10]	MSE	12.709	14.298	349.47	276.56	261.47
[1/10, 1/20]	MSE	15.044	12.226	456.07	370.92	350.11
[1/20, 1/40]	MSE	21.436	16.109	714.80	618.06	509.80
[1/5, 1/10, 1/40]	MSE	21.436	16.588	714.80	618.06	559.11
[1/5, 1/10, 1/20]	MSE	15.044	12.160	456.07	370.92	350.11
[1/10, 1/20, 1/40]	MSE	21.435	16.518	714.80	618.06	535.03
[1/5, 1/10]	SSIM	0.2947	0.2992	0.5484	0.5617	0.5626
[1/10, 1/20]	SSIM	0.2807	0.2926	0.5151	0.5329	0.5365
[1/20, 1/40]	SSIM	0.2480	0.2606	0.5325	0.4946	0.4306
[1/5, 1/10, 1/40]	SSIM	0.2480	0.2556	0.5325	0.4946	0.5044
[1/5, 1/10, 1/20]	SSIM	0.2807	0.2926	0.5151	0.5329	0.5365
[1/10, 1/20, 1/40]	SSIM	0.2480	0.2524	0.5325	0.4946	0.5098

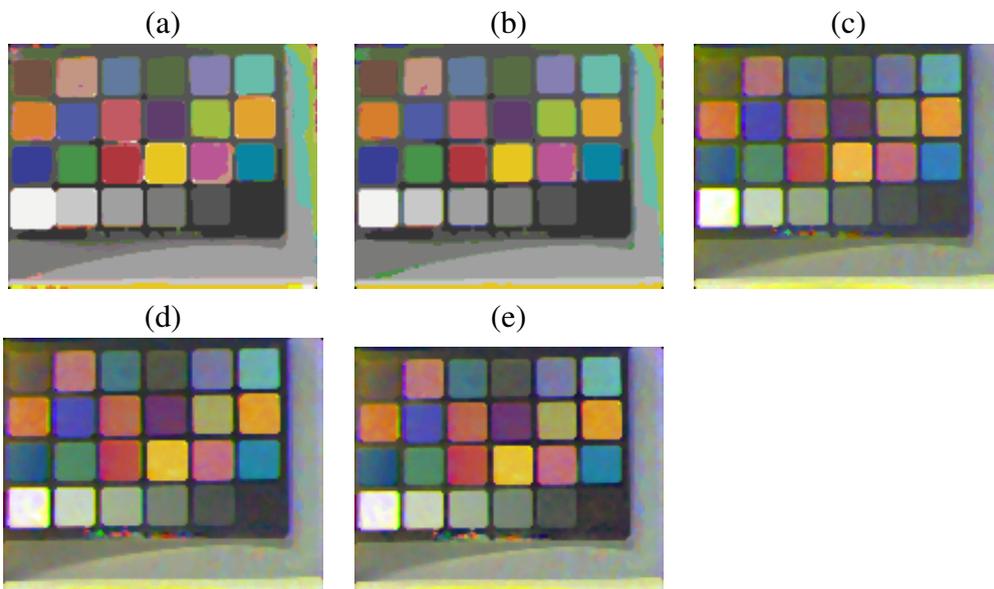


FIGURA 4.5: Imágenes de la cartilla de Macbeth para diferentes métodos usando la tercera forma de aplicar HDR. (a) Mejor resultado en SVM. (b) Mejor resultado en *k*-NN. (c) Mejor resultado en regresión lineal. (d) Mejor resultado en regresión de segundo orden. (e) Mejor resultado en regresión de tercer orden.

4.2 Pruebas cualitativas

Para las pruebas cualitativas se utilizó la reconstrucción de color de las imágenes de botellas y la cartilla de Macbeth para realizar una prueba de percepción visual, esto es explicado a detalle en las subsecciones siguientes. En la primera, se describe la estructura de la prueba así como la muestra. En la segunda se muestran los resultados de la prueba con las imágenes de Macbeth y por último se explica los resultados de la prueba utilizando las imágenes de botellas.

4.2.1 Descripción de la prueba

Esta prueba tiene como objetivo comparar las imágenes de color desde la perspectiva del ojo humano, comúnmente este tipo de pruebas es conocida como prueba de percepción. La prueba se realizó creando una encuesta en la plataforma digital *SurveyMonkey* y se dividió en tres preguntas: en la primera se buscaba saber si el encuestado estaba familiarizado con fotografía profesional o procesamiento de imágenes. La segunda y tercera pedían ordenar de mayor a menor calidad un grupo de ocho imágenes pertenecientes a la cartilla de Macbeth, o botellas. En la Figura 4.6 se presenta una captura de como se ve la segunda y tercera pregunta de la prueba realizada. Esta fue llenada por 70 personas, de las cuales el 57 % aseguró tener conocimientos en fotografía profesional o procesamiento de imágenes. Por otro lado 26 % marcó la opción de no tener ningún conocimiento en el área, el resto de participantes omitió esta pregunta. El conjunto de imágenes elegido para la segunda y tercera prueba pertenecen a todos los métodos utilizados en este trabajo, menos k -NN, ya que sus resultados son muy similares a SVM. Así mismo se incluyó sus versiones utilizando HDR de tercer tipo y una imagen capturada con una cámara a color con similares características que el prototipo.

Prueba de percepción visual Visual perception test

Página 2 de 3
Page 2 of 3

* 4. ¿Qué imagen se ve mejor? Ordena de mayor a menor calidad (puedes arrastrar las imágenes con el mouse o poner los números correspondientes en las casillas.)
Which image looks better? Order from highest to lowest quality (you can drag the images with the mouse or put the corresponding numbers in the boxes.)

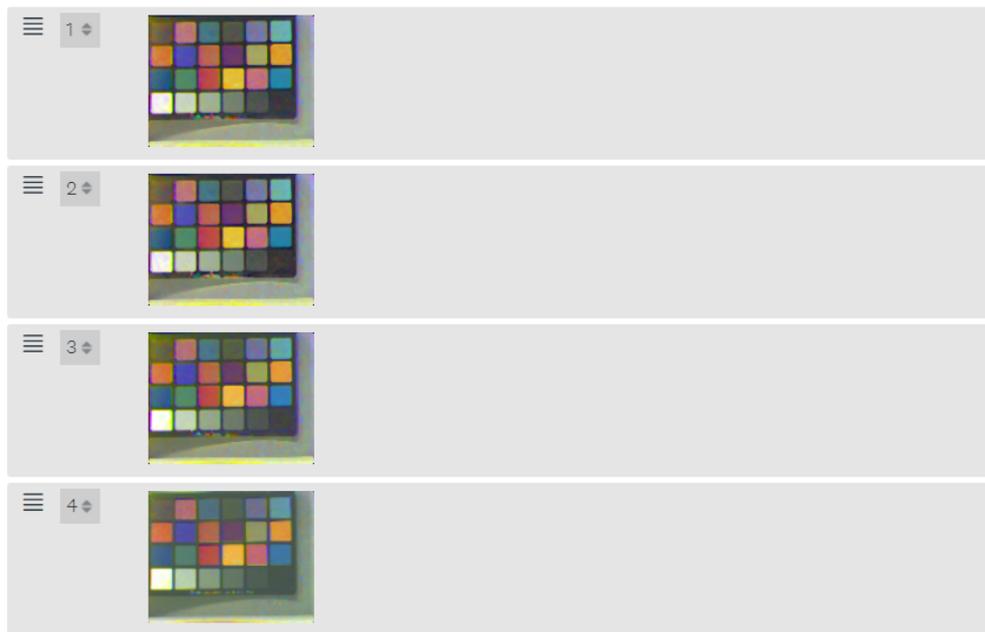


FIGURA 4.6: Captura parcial de la prueba de percepción con las imágenes de la cartilla de Macbeth. Se visualiza la mitad del total de imágenes presentadas a los voluntarios.

4.2.2 Prueba de percepción de imágenes de la Cartilla de Macbeth

Las imágenes que se muestran en la Figura 4.7 son las que se presentaron en la primera prueba de percepción visual. Según el orden que cada encuestado propuso se calificó cada imagen con un puntaje entre 8 y 1, siendo 8 la mejor y 1 la peor. Los datos obtenidos fueron analizados, en primer lugar, dividiéndolos en dos grupos: los que consideran que tienen más conocimientos en fotografía o procesamiento de imágenes y los que no. Así mismo también se analizaron de manera global con todas las respuestas obtenidas.

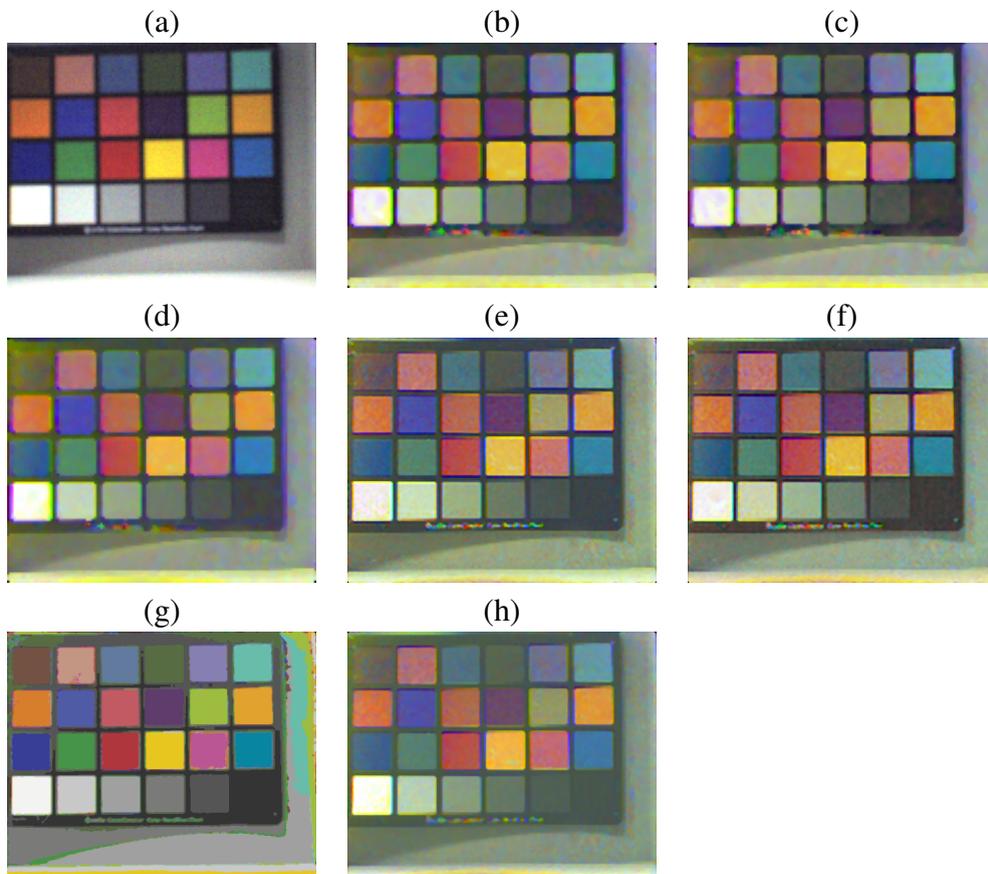


FIGURA 4.7: Imágenes de la cartilla de Macbeth usados en la primera prueba de percepción visual ordenadas de mejor a peor (a-h) según los resultados de la prueba de percepción. (a) Imagen a color. (b) No lineal de tercer orden HDR tipo 3. (c) No lineal de segundo orden HDR tipo 3. (d) Lineal HDR tipo 3. (e) No lineal de tercer orden. (f) No lineal de segundo orden. (g) SVM. (h) Lineal.

En la Figura 4.8, en color azul se puede apreciar un gráfico de barras con las calificaciones promedio obtenidas para cada una de las imágenes tomando en cuenta solo las personas que consideran tener más conocimiento en fotografía o procesamiento de imágenes. Como se observa, el mejor resultado es obtenido por la imagen a color con un puntaje promedio de 6.73, seguido de regresión no lineal de tercer orden con un puntaje de 5.81 casi empatado con la regresión lineal usando HDR con un puntaje de 5.83. Era de esperarse que la imagen a color esté en primer lugar debido a que presenta los colores

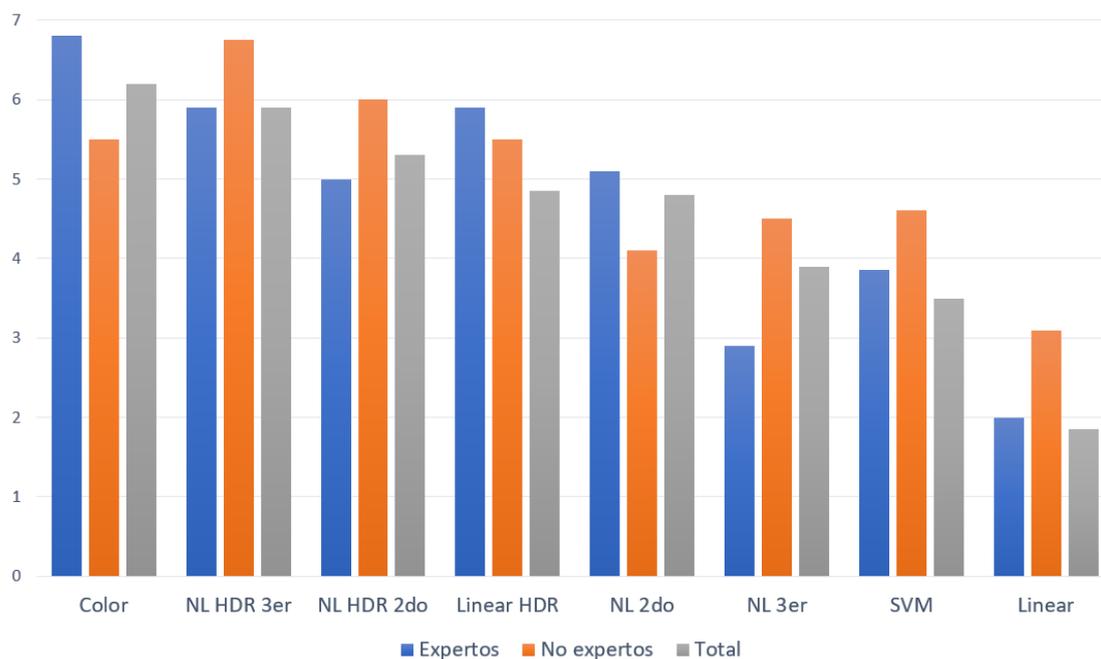


FIGURA 4.8: Gráfica de barras de los puntajes promedio de cada uno de los métodos presentados en la primera prueba de percepción visual. En azul los resultados para las personas con experiencia en procesamiento de imágenes o fotografía. En naranja los que no tienen experiencia. En gris el puntaje tomando en cuenta el total de participantes: los dos anteriores grupos más los que no indicaron ninguna categoría.

reales y nítidos porque fue tomada con una cámara especializada en color, sin embargo es importante destacar la preferencia y altos puntajes obtenidos por los métodos de regresión. En color naranja, se muestra la gráfica de barras para las personas sin conocimiento técnico. Esta gráfica muestra que la regresión no lineal de tercer orden y no lineal de segundo orden son las más destacadas, considerando que no muestran mucha diferencia en puntaje, siendo la primera de 6.67 y la segunda de 6.02. Este resultado se puede deber a que este grupo de personas no considera la tonalidad en la percepción de una imagen de buena calidad y por ello no eligieron a la imagen a color como la mejor. Por último en color gris, se muestran los resultados generales con todas las respuestas. El primer lugar se lo lleva la imagen a color con un puntaje de 6.25, seguido de regresión lineal de tercer orden HDR con un puntaje de 5.92. Así mismo el último lugar se lo llevó la regresión

lineal con un puntaje de 1.73. Es importante resaltar en esta primera prueba que los métodos con mayores puntajes fueron los mismos en obtener los mejores resultados numéricos en SSIM, pero no los mejores en MSE (SSIM y k -NN). Esto se debe a que la estructura de la imagen es un factor que consideran importante a la hora de calificar una imagen como buena o mala.

4.2.3 Prueba de percepción de imágenes de botellas

Las imágenes que se muestran en la Figura 4.9 son las que se presentaron en la segunda prueba de percepción visual. A diferencia de todas las pruebas anteriores, en esta prueba se presentan imágenes de botellas metálicas en donde se muestra una azul y otra roja. Esta prueba es importante debido a que se está reconstruyendo con una imagen diferente a la que se utilizó para entrenar el modelo de color. Estas imágenes probarán la fidelidad de los métodos ante nuevas muestras de colores. El análisis de la data obtenida es similar a la primera prueba y se detalla a continuación.

En la Figura 4.10, en color azul, se muestra la gráfica de barras de los puntajes promedio para las personas que consideran tener conocimientos en fotografía o procesamiento de imágenes. El mejor resultado lo obtiene la regresión no lineal de tercer orden con HDR con un puntaje de 7.01, seguido de regresión no lineal de segundo orden casi empatado con la imagen a color con un puntaje de 5.24 y 5.11 respectivamente. A diferencia de la primera prueba, la imagen a color no fue la favorita y el método no lineal lo supera en un 27 por ciento aproximadamente. En color naranja, se muestra la siguiente gráfica de barras que muestra los resultados de las personas que no consideran poseer cualidades técnicas. En esta, la relación de puntajes se mantiene casi igual, con la diferencia que la brecha entre no lineal de tercer y segundo orden decrece, siendo el puntaje de la primera de 6.08 y 5.67 de la segunda. La imagen a color queda desplazada al último lugar con un puntaje de 3.33, esto se puede explicar debido a que de manera similar con la primera prueba, el segundo grupo de personas, no considera la intensidad de color en su



FIGURA 4.9: Imágenes de botellas usados en la segunda prueba de percepción visual ordenadas de mejor a peor (a-h) según los resultados de la prueba de percepción. (a) No lineal de tercer orden HDR tipo 3. (b) No lineal de segundo orden HDR tipo 3. (c) Lineal HDR tipo 3. (d) Lineal. (e) Imagen a color. (f) No lineal de segundo orden. (g) No lineal de tercer orden. (h) SVM.

percepción de calidad de una imagen. Por último, en color gris, se muestran los resultados generales de esta prueba. En líneas generales, los puntajes siguen la tendencia de las dos anteriores, siendo el puntaje más alto 6.41 para el método de regresión no lineal de tercer orden HDR, seguido de los otros dos métodos de regresión con HDR con puntajes de 5.33 y 5.08. En tercer, cuarto y quinto lugar se encuentran la regresión lineal, la imagen a color y la regresión no lineal de segundo orden con puntajes de 4.11, 4.07 y 3.97. En último lugar se ubica SVM con un puntaje de 3.38. Este último lugar se explica por

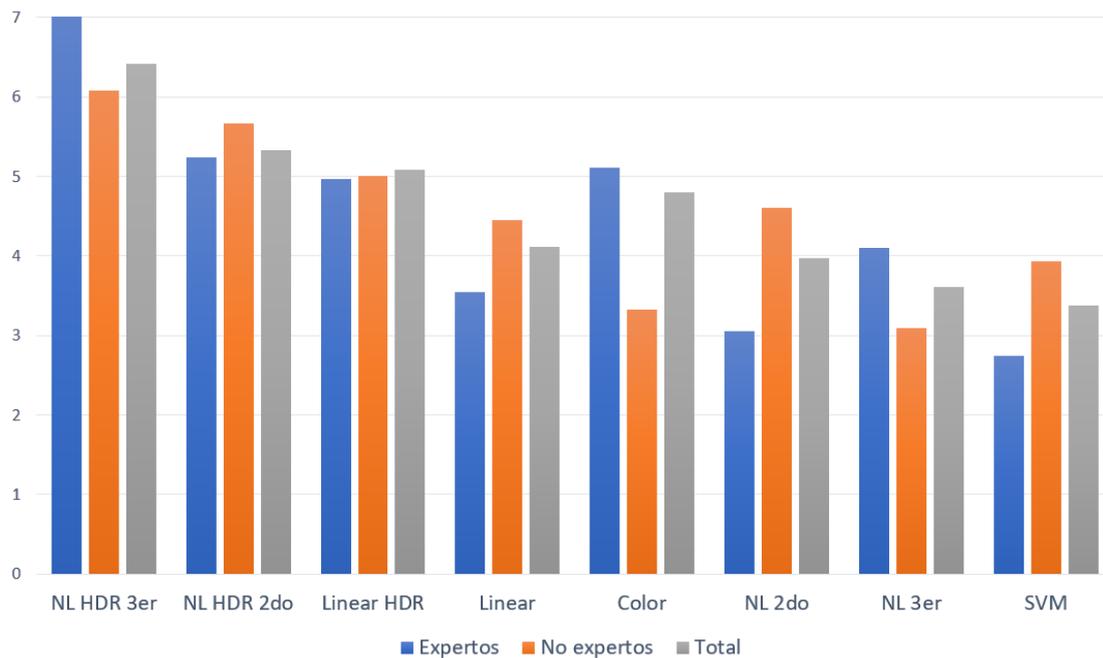


FIGURA 4.10: Gráfica de barras de los puntajes promedio de cada uno de los métodos presentados en la segunda prueba de percepción visual. En azul los resultados para las personas con experiencia en procesamiento de imágenes o fotografía. En naranja los que no tienen experiencia. En gris el puntaje tomando en cuenta el total de participantes: los dos anteriores grupos más los que no indicaron ninguna categoría.

la forma en que SVM reconstruye el color, es decir, solo selecciona entre uno de los 24 colores disponibles en la cartilla de Macbeth y clasifica cada píxel de acuerdo al modelo. Este procedimiento genera que la imagen no se vea orgánica y los cambios de tonalidad sean abruptos. En esta segunda prueba, al igual que en la primera, siguieron primando los métodos de regresión utilizando HDR y SVM quedó desplazada a los últimos lugares.

4.3 Prueba de ampliación de rango dinámico

Adicionalmente a los cambios realizados en la reconstrucción de color al introducir imágenes HDR en el procedimiento, se realizó una variación utilizando modelos ya entrenados para mejorar los resultados. Esta variación consiste en entrenar el modelo solo

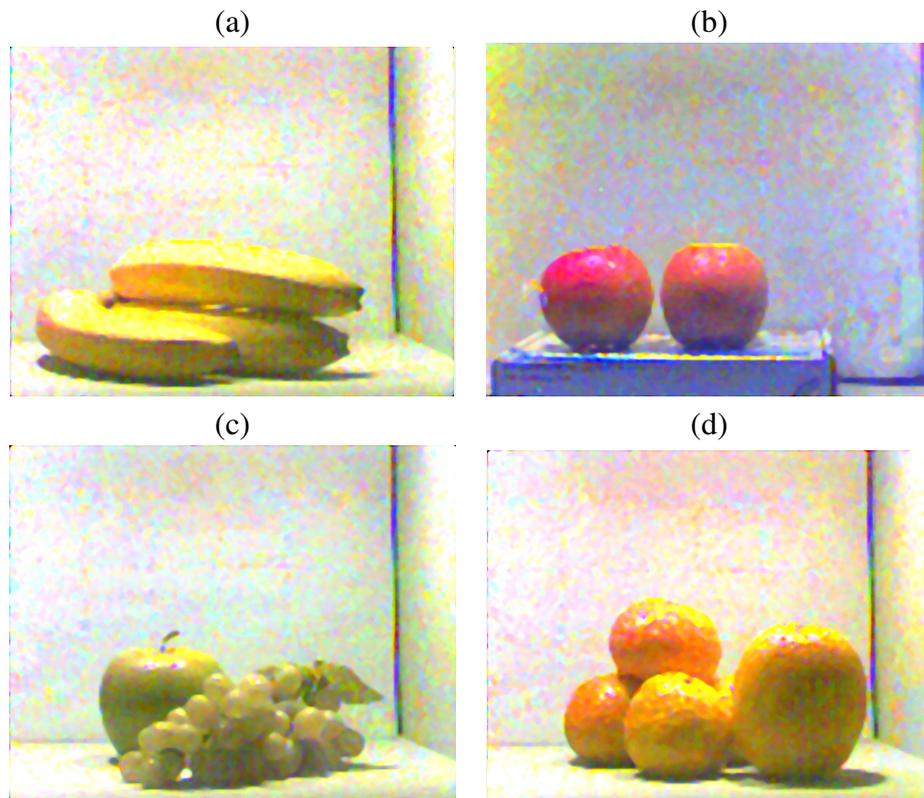


FIGURA 4.11: Imágenes de frutas reconstruidas en color usando regresión no lineal de 3er orden con HDR de tercer tipo y ampliando el rango dinámico. (a) Plátanos. (b) Manzanas. (c) Uvas y una manzana verde. (d) Mandarinas y una naranja.

con dos imágenes de tiempos de exposición diferente y utilizarlo para obtener una imagen a color con otro grupo de imágenes de diferente tiempo de exposición. Por ejemplo, se podría entrenar el modelo utilizando dos imágenes de tiempos de exposición $1/10$ s y $1/20$ s y se obtendría un modelo. Este modelo se usaría para reconstruir el color utilizando imágenes de tiempos de exposición $1/5$ s, $1/10$ s, $1/20$ s, $1/40$ s. Esta variación permite que imágenes sobresaturadas o que fueron tomadas con mucha iluminación puedan ser reconstruidas en color correctamente. Debido a que las imágenes de referencia (cartilla de Macbeth) con las que se entrena el modelo no presentan un gran rango dinámico, las nuevas muestras que serán reconstruidas en color también deben tener un rango dinámico similar. Para lidiar con el problema que se ocasiona al reconstruir color de muestras con

alto rango dinámico comparado con la referencia, se realizó el ajuste descrito en el párrafo anterior. Esta última prueba se realizó utilizando imágenes de cuatro tipos de frutas: plátanos, manzanas, uvas y mandarinas. Así mismo, se usó el mejor método seleccionado según las pruebas cuantitativas y cualitativas. El método para obtener el modelo de color elegido fue regresión no lineal de tercer orden, ya que fue el mejor en la percepción visual y el mejor en las pruebas cuantitativas sin contar SVM y k -NN ya que estos métodos fueron descartados debido a las pruebas de percepción visual. El tipo de separación de muestras fue el segundo es decir la interpolación a 360×450 píxeles debido a que es el que presenta mejores resultados en promedio para la regresión no lineal y aumenta la resolución de la imagen. Por último la forma de aplicar HDR elegida fue la tercera por ser superior a las anteriores por más del diez por ciento. Los resultados se muestran en la Figura 4.11. El fondo blanco de todas las imágenes presenta un error de saturación, ya que algunos valores RGB de los píxeles no llegan a 255 para ser totalmente blancos, sin embargo, los colores de las frutas fueron reconstruidos correctamente y las tonalidades saturadas o brillantes se distinguen claramente. Esto comprueba que ampliando el rango dinámico de las muestras comparado con el modelo entrenado, se puede reconstruir correctamente el color de imágenes con diferente exposición de luz a las imágenes de referencia con las que se entrena el modelo.

4.4 Trabajos de investigación publicados

A lo largo del proceso de elaboración del presente trabajo, resultados parciales y parte de la metodología descrita en este documento fueron enviados a conferencias con revisión por pares para su publicación en repositorios científicos reconocidos internacionalmente. En total se publicó tres *papers* [20] [21] [22]. De los cuales dos fueron aceptados en IEEE y uno en Springer. Finalmente, es importante mencionar que los resultados de este trabajo están por presentarse en un *journal* internacional junto con la universidad de Harvard [23].

CONCLUSIONES

El objetivo principal de este trabajo era diseñar un algoritmo que obtenga una imagen a color a partir de nueve bandas multispectrales captadas por un prototipo basado en nanowires. Según los resultados obtenidos, utilizando el algoritmo propuesto se logró disminuir los errores de construcción del prototipo tales como: manchas en la grilla de filtros, error de orientación de los filtros, aberración cromática y saturación de píxeles. En las pruebas cualitativas y cuantitativas, el mejor método para generar un modelo de reconstrucción de color es la regresión no lineal de tercer orden, ya que a pesar de que los métodos k -NN y SVM muestran mejores resultados en MSE, necesitan una base de datos más amplia para la correcta reconstrucción de colores fuera de los de la cartilla de Macbeth y por ello no fueron elegidos como los mejores en la prueba de percepción visual. Por otro lado, en las pruebas cuantitativas, se demostró que el método de separación de muestras por interpolación a 360×450 píxeles es la mejor para la reconstrucción de color, ya que aunque el MSE no es el mejor, la imagen se ve más nítida. Así mismo, incluir HDR en la reconstrucción de color mostró mejorar los resultados en promedio 9.1 por ciento, en la primera forma de utilizar imágenes de alto rango dinámico y 29.3 por ciento, en la tercera. Esto fue contrastado con las pruebas cualitativas donde las imágenes donde se usó HDR obtuvieron los mejores puntajes. Por último, ampliando el rango dinámico de las muestras sin cambiar el modelo, es posible obtener una reconstrucción de color en imágenes con diferente exposición de luz comparada con las imágenes de referencia.

RECOMENDACIONES

Con los resultados y el análisis realizado, se recomienda utilizar la separación de muestras de tamaño 360×450 píxeles, regresión no lineal de tercer orden para obtener el modelo de color y utilizar la tercera forma de aplicar HDR para reconstruir el color en imágenes multiespectrales de nueve bandas capturadas por el prototipo propuesto para obtener la mejor calidad de imagen.

Por otro lado, los métodos de SVM y k -NN no mostraron buenos resultados debido a la poca cantidad de colores en la base de datos. Por ello, para un trabajo futuro, de utilizar algún método de aprendizaje supervisado similar, se recomienda aumentar la base de datos de colores, ya que este trabajo demostró que 24 colores no fueron suficientes para caracterizar un modelo correcto de color. Así mismo, se recomienda ampliar la base de datos con imágenes de referencia con diferente exposición de luz podría generar modelos más precisos para imágenes en condiciones no ideales.

ANEXOS

Los algoritmos desarrollados y utilizados en este trabajo se encuentran en el siguiente link: <https://cutt.ly/KniNPJk>.

En el repositorio se encuentran dos carpetas: en la primera, «Base de datos», se encuentran las capturas del prototipo utilizadas, y en la segunda, «Core Folder», se encuentran los archivos de MATLAB que contienen los algoritmos utilizados. Dentro de esta última carpeta existen 3 tipos de datos: los coeficientes de cada modelo entrenado, datos necesarios para el funcionamiento de los algoritmos como posiciones de los colores en las cartillas de Macbeth y valores RGB, y por último se encuentran los algoritmos propiamente, de los cuales los principales son:

- HDR_type1.m, HDR_type2.m, y HDR_type3.m
- KNN.m y SVM.m
- LinealRegression.m
- NonLinearMultipleRegressionCorrected.m

Cada archivo tiene el nombre de los algoritmos descritos en el capítulo 3.

REFERENCIAS BIBLIOGRÁFICAS

- [1] R. A. Española, “Diccionario de la lengua española (22. aed.),” *Consultado en* <http://www.rae.es/rae.html>, vol. 126, 2001.
- [2] H. R. Morris, C. C. Hoyt, and P. J. Treado, “Imaging spectrometers for fluorescence and raman microscopy: Acousto-optic and liquid crystal tunable filters,” *Applied Spectroscopy*, vol. 48, no. 7, pp. 857–866, 1994. [Online]. Available: <https://doi.org/10.1366/0003702944029820>
- [3] A. F. Goetz, “Three decades of hyperspectral remote sensing of the earth: A personal view,” *Remote Sensing of Environment*, vol. 113, pp. S5 – S16, 2009, imaging Spectroscopy Special Issue. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S003442570900073X>
- [4] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, “Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 2, pp. 354–379, April 2012.
- [5] B. E. Bayer, “Color imaging array,” Jul. 20 1976, uS Patent 3,971,065.
- [6] T. Sakamoto, C. Nakanishi, and T. Hase, “Software pixel interpolation for digital still cameras suitable for a 32-bit mcu,” *IEEE Transactions on Consumer Electronics*, vol. 44, no. 4, pp. 1342–1352, Nov 1998.

- [7] P.-J. Lapray, X. Wang, J.-B. Thomas, and P. Gouton, "Multispectral filter arrays: Recent advances and practical implementation," *Sensors*, vol. 14, no. 11, pp. 21 626–21 659, 2014. [Online]. Available: <http://www.mdpi.com/1424-8220/14/11/21626>
- [8] H. Park and K. B. Crozier, "Multispectral imaging with vertical silicon nanowires," *Scientific Reports*, vol. 3, p. 2460, Aug. 2013. [Online]. Available: <https://doi.org/10.1038/srep02460>
- [9] R. Ramanath, W. E. Snyder, G. L. Bilbro, and W. A. Sander, "Robust multispectral imaging sensors for autonomous robots," Technical Report, Tech. Rep., 2001.
- [10] P. E. Buchsbaum and M. J. Morris, "Method for making monolithic patterned dichroic filter detector arrays for spectroscopic imaging," Oct. 28 2003, uS Patent 6,638,668.
- [11] H. Qi and R. Ramanath, "The design and evaluation of a generic method for generating mosaicked multispectral filter arrays," *IEEE Transactions on Image Processing*, vol. 15, no. 9, pp. 2780–2791, Sep. 2006.
- [12] L. Miao, H. Qi, and R. Ramanath, "A generic binary tree-based progressive demosaicking method for multispectral filter array," in *2006 International Conference on Image Processing*, Oct 2006, pp. 3221–3224.
- [13] J. Brauers and T. Aach, "A color filter array based multispectral camera," in *12. Workshop Farbbildverarbeitung*. Ilmenau, 2006.
- [14] A. Majumdar, "Multi-spectral demosaicing technique for single-sensor imaging," in *2013 Fourth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG)*, Dec 2013, pp. 1–4.
- [15] H. Park, Y. Dan, K. Seo, Y. J. Yu, P. K. Duane, M. Wober, and K. B. Crozier, "Filter-free image sensor pixels comprising silicon nanowires with selective color

- absorption,” *Nano Lett.*, vol. 14, no. 4, pp. 1804–1809, Apr. 2014. [Online]. Available: <https://doi.org/10.1021/nl404379w>
- [16] K. Seo, M. Wober, P. Steinvurzel, E. Schonbrun, Y. Dan, T. Ellenbogen, and K. B. Crozier, “Multicolored vertical silicon nanowires,” *Nano Letters*, vol. 11, no. 4, pp. 1851–1856, 2011, pMID: 21413684. [Online]. Available: <https://doi.org/10.1021/nl200201b>
- [17] H. Park, K. Seo, and K. B. Crozier, “Adding colors to polydimethylsiloxane by embedding vertical silicon nanowires,” *Applied Physics Letters*, vol. 101, no. 19, p. 193107, 2012. [Online]. Available: <https://doi.org/10.1063/1.4766944>
- [18] E. Reinhard, W. Heidrich, P. Debevec, S. Pattanaik, G. Ward, and K. Myszkowski, *High dynamic range imaging: acquisition, display, and image-based lighting*. Morgan Kaufmann, 2010.
- [19] Z. Wang, E. P. Simoncelli, and A. C. Bovik, “Multiscale structural similarity for image quality assessment,” in *The Thirty-Seventh Asilomar Conference on Signals, Systems Computers, 2003*, vol. 2, 2003, pp. 1398–1402 Vol.2.
- [20] J. C. Barbaran, V. Murray, M. Wober, and F. Habbal, “High-dynamic-range imaging using nanowires,” in *Brazilian Technology Symposium*. Springer, 2019, pp. 133–142.
- [21] ———, “Color estimation using nanowires sensors,” in *2019 IEEE XXVI International Conference on Electronics, Electrical Engineering and Computing (INTERCON)*. IEEE, 2019, pp. 1–4.
- [22] ———, “A first approach to multispectral high-dynamic-range imaging using nanowires sensors,” in *2019 IEEE XXVI International Conference on Electronics, Electrical Engineering and Computing (INTERCON)*. IEEE, 2019, pp. 1–4.

- [23] V. Murray, J. C. Barbaran, M. Wober, and F. Habbal, “Color estimation using nanowires and machine learning,” in *to be submitted to Nature Electronics*. Nature, 2021, pp. 1–10.