



Universidad Nacional del Sur

TESIS DE DOCTOR EN MATEMÁTICA

CONTRIBUCIONES AL ANÁLISIS
ESTADÍSTICO DE IMÁGENES

Jorge Alberto Martinez

Bahía Blanca - Argentina

2020

Prefacio

Esta Tesis se presenta como parte de los requisitos para optar al grado Académico de Doctor en Matemática, de la Universidad Nacional del Sur. La misma no ha sido presentada previamente para la obtención de otro título en esta Universidad u otra. La misma contiene los resultados obtenidos en investigaciones realizadas en el ámbito del Departamento de Matemática, bajo la dirección de la Dra. Ana Georgina Flesia, Profesora Asociada de la Facultad de Matemática, Astronomía y Física (FaMAF) de la Universidad Nacional de Córdoba e Investigadora Independiente CIEM-CONICET y la supervisión local de la Dra. María Cristina Maciel, Profesora Titular del Departamento de Matemática de la Universidad Nacional del Sur.

Jorge Alberto Martinez



UNIVERSIDAD NACIONAL DEL SUR
Secretaría de Posgrado y Educación Continua

La presente tesis ha sido aprobada el día .../.../....., mereciendo la calificación de
...(.....)

A Silvana

Agradecimientos

Quiero expresar mi más sincero agradecimiento a mi directora, la Dra. Ana Georgina Flesia por haberme dirigido en esta tesis, por su hospitalidad y paciencia, como así también por sus sugerencias y recomendaciones brindadas en cada viaje que realicé a Córdoba.

A la Dra. María Cristina Maciel, mi supervisora local, por poder integrar su grupo de investigación, por su apoyo y acompañamiento en esta tarea.

Al Dr. Oscar H. Bustos por su ayuda desinteresada, su buena predisposición y calidez para conmigo.

A Silvina, por su infinita contención y ayuda, ¡Gracias por estar siempre a mi lado!.

Finalmente, a toda mi familia y amigos quienes me alentaron en todo momento.

Resumen

Este trabajo se compone de dos partes que se enmarcan en el área del procesamiento estadístico de imágenes. La primera, corresponde al modelado estadístico de una imagen y la segunda a la medición de la calidad de una imagen en el contexto de fusión. Cada una de las partes de este trabajo es autocontenida.

En el Capítulo 1 de la primer parte, se describe el marco teórico concerniente al modelado estocástico de una imagen de textura mediante los Campos Aleatorios de Gibbs-Markov. Además, se establecen las condiciones suficientes para obtener la descomposición de un potencial de Gibbs en configuraciones basadas en el esquema de codificación de Besag. En el Capítulo 2, se propone un nuevo método de estimación basado en el esquema de codificación de Besag, denominado estimador de Mínimo Cuadrado Condicional Coding (MCCC), implementado mediante un algoritmo paralelo. Para evaluar su performance, se llevó a cabo un estudio de simulación Monte Carlo en situaciones específicas de un modelo de textura Gibbs-Markov, el Modelo Autobinomial. El estudio consistió en un análisis descriptivo de su comportamiento, el testeo de su sensibilidad ante presencia de ruido (distintos grados de contaminación de diferentes patrones homogéneos de textura), la evaluación de su algoritmo paralelo y su capacidad en la clasificación supervisada de diferentes texturas presentes en una imagen satelital concreta.

La segunda parte aborda la problemática de la medida de la calidad de una imagen en el contexto de fusión. En el Capítulo 4 se presentan las características de las medidas de calidad basadas en el índice de similitud estructural, como así también algunas propiedades de interés de este índice. Luego, en el Capítulo 5 se propone una nueva medida de calidad de fusión de imágenes que tiene como objetivo ser ciega (sin imagen de referencia), de cálculo significativamente simple y de fácil implementación. Para validar esta propuesta se desarrolla tres experimentos que evalúan: su propósito general, la estabilidad y capacidad discriminativa, como así también, la correlación con la opinión subjetiva. Por último, se presenta una aplicación de la medida propuesta en la obtención de imágenes de pinturas de alta calidad, a través de un proceso de exclusión que selecciona con éxito el mejor subconjunto de imágenes de entrada en términos de la imagen fusionada.

Algunos de estos resultados fueron publicados en [47, 48, 12, 58, 49, 50]. Estos son:

- J. Martinez, S. Pistonesi y A.G. Flesia Inference strategies for texture parameters. In A. Pardo and J. Kittler, editors, *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, 460 – 467, Cham, 2015. Springer International Publishing.
- J. Martinez, S. Pistonesi, M.C. Maciel y A.G. Flesia. Image Fusion Quality Measure Based on a Multi-scale Approach. In G. Bebis et al., editors, *Advances in Visual Computing. ISVC 2016*, 836 – 845, Cham, 2016. Springer International Publishing.
- O.H Bustos, V. Rulloni, y J. Martinez. Distribuciones de Gibbs: Volumen 1. Trabajos de Matemática, Serie B 65/17, Facultad de Matemática, Astronomía y Física. Universidad Nacional de Córdoba, 2017.
- S. Pistonesi, J. Martinez, S.M. Ojeda y R. Vallejos. Structural Similarity Metrics for Quality Image Fusion Assessment: Algorithms. *Image Processing On Line*, 8 : 345 – 368, 2018
- J. Martinez, S. Pistonesi, M.C. Maciel y A.G. Flesia. Parameter estimation in a Gibbs-Markov field texture model based on a coding approach. *In 2018 IEEE Statistical Signal Processing Workshop (SSP)*, 105-09, Junio 2018.
- J. Martinez, S. Pistonesi, M.C. Maciel y A.G. Flesia. Multi-scale fidelity measure for image fusion quality assessment. *Information Fusion*, 50 : 197 – 211, 2019.

Además, se encuentra en progreso el siguiente trabajo

- J. Martinez and S. Pistonesi. *Gibbs-potential decomposition based on coding sets*.

Abstract

This work consists of two parts, framed in the area of statistical image processing. The first corresponds to the statistical modeling of an image and the second to the measurement of the quality of an image in the fusion context.

In Chapter 1, the theoretical framework concerning the stochastic modeling of a texture image using the Gibbs-Markov Random Fields is described. Furthermore, sufficient conditions are established to obtain the decomposition of a Gibbs potential in configurations based on the Besag coding scheme.

In Chapter 2, we propose a new estimation method based on the Besag coding scheme, called the Coding Conditional Least Square estimator (CCLS), implemented by a parallel algorithm. In order to evaluate its performance, we have conducted a Monte Carlo simulation study in specific situations of a Gibbs-Markov texture model, the Auto-binomial Model. The study consisted of a descriptive analysis of its behavior, the sensitivity testing before the presence of noise (different contamination degrees in different homogeneous texture patterns), the evaluation of its parallel algorithm, and its capacity in the supervised classification of different textures on a specific satellite image.

The second part addresses the problem of measuring the quality of an image in the fusion context. Chapter 4 presents the quality measures characteristics of based on the structural similarity index, as well as some interest properties of this index. subsequently, in Chapter 5, a new image fusion quality measure is proposed. This measure aims to be blind (without a reference image) and significantly simple to calculate. It is also, easily to implemented. In order to validate this proposal, we have developed three experiments to asses the following aspects: its general purpose, stability and discriminative capacity, as well as the correlation with subjective opinion. Finally, we present an application of the proposed measure in obtaining images of high quality paintings, by means of an exclusion process (leave one out) that successfully selects the best subset of input images in terms of the fused image.

Some of these results were published in [47, 48, 12, 58, 49, 50]. These are:

- J. Martinez, S. Pistonesi and A.G. Flesia Inference strategies for texture parameters. In A.

Pardo and J. Kittler, editors, *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, 460 – 467, Cham, 2015. Springer International Publishing.

- J. Martinez, S. Pistonesi, M.C. Maciel and A.G. Flesia. Image Fusion Quality Measure Based on a Multi-scale Approach. In G. Bebis et al., editors, *Advances in Visual Computing. ISVC 2016*, 836 – 845, Cham, 2016. Springer International Publishing.
- O.H Bustos, V. Rulloni, and J. Martinez. Distribuciones de Gibbs: Volumen 1. Trabajos de Matemática, Serie B 65/17, Facultad de Matemática, Astronomía y Física. Universidad Nacional de Córdoba, 2017.
- S. Pistonesi, J. Martinez, S.M. Ojeda and R. Vallejos. Structural Similarity Metrics for Quality Image Fusion Assessment: Algorithms. *Image Processing On Line*, 8 : 345 – 368, 2018
- J. Martinez, S. Pistonesi, M.C. Maciel and A.G. Flesia. Parameter estimation in a Gibbs-Markov field texture model based on a coding approach. *In 2018 IEEE Statistical Signal Processing Workshop (SSP)*, 105-09, June 2018.
- J. Martinez, S. Pistonesi, M.C. Maciel and A.G. Flesia. Multi-scale fidelity measure for image fusion quality assessment. *Information Fusion*, 50 : 197 – 211, 2019.

Also, the following work is in progress

- J. Martinez and S. Pistonesi. *Gibbs-potential decomposition based on coding sets.*

Índice general

Prefacio	I
Resumen	VII
Abstract	IX
I Modelización de una imagen	1
1. Modelado de imágenes de textura	5
1.1. Texturas	5
1.2. Imágen de textura	6
1.3. Modelado probabilístico	9
1.4. Distribuciones de Gibbs	10
1.5. Campos Markovianos	13
1.6. Ejemplo: Modelo de Ising 2-D	15
1.7. Descomposición de un potencial de Gibbs en conjuntos Coding	16
2. Estimación de parámetros texturales	23
2.1. Modelo de Textura: Autobinomial	23
2.2. Métodos de estimación	26
2.2.1. Mínimo Cuadrado Condicional	27
2.2.2. Mínimo Cuadrado Condicional Coding	28
2.3. Análisis de los métodos de estimación	28
2.3.1. Medidas de desempeño	28
2.3.2. Sensibilidad	29
2.3.3. Evaluación de los algoritmos	34
2.3.4. Clasificación supervisada de una imagen real	35
2.4. Resultados experimentales	36
2.4.1. Análisis descriptivo	37
2.4.2. Resultados de sensibilidad	42

2.4.3.	Evaluación de los algoritmos	48
2.4.4.	Algunos resultados de los estimadores para un MAB de cuarto y quinto orden	50
3.	Conclusión y estudios futuros	59
3.1.	Conclusión	59
3.2.	Estudios futuros	59
II	Calidad de una imagen	63
4.	Medidas de calidad	67
4.1.	Medidas de similitud estructural de la imagen	67
4.1.1.	Índice de calidad de imagen universal (Q)	67
4.1.2.	Índice de similitud estructural (SSIM)	68
4.1.3.	Una clase de métricas basadas en SSIM	70
4.2.	Medidas de calidad en fusión de imágenes	77
4.2.1.	Medidas de Piella (Q_S , Q_W y Q_E)	77
4.2.2.	Medida de Cvejic (Q_C)	78
4.2.3.	Medida de Yang (Q_Y)	78
4.2.4.	Versión extendida	80
5.	Evaluación de la calidad de similitud estructural a escala múltiple	83
5.1.	Índice de calidad de similitud estructural multiescala (MSSIM)	83
5.2.	Medida de calidad propuesta ($MS-Q_W$)	84
5.3.	Validación	87
5.3.1.	Primer experimento: evaluación de propósito general	87
5.3.2.	Segundo experimento: estabilidad y capacidad discriminativa de la medida	90
5.3.3.	Tercer experimento: evaluación de acuerdo con la opinión subjetiva	91
5.4.	Resultados experimentales	92
5.4.1.	Resultados del primer experimento: evaluación de propósito general	92
5.4.2.	Resultados del segundo experimento: estabilidad y capacidad discriminativa de la medida	93
5.4.3.	Resultados del tercer experimento: evaluación de acuerdo con la opinión subjetiva	95
5.5.	Otros resultados	97
5.6.	Aplicación	99

5.6.1. Obtención de fotografías de alta calidad de pinturas por fusión de imágenes	99
6. Conclusión y estudios futuros	109
6.1. Conclusión	109
6.2. Estudios futuros	110
A. Algoritmos de simulación	113
A.1. Cadenas de Markov (de imágenes)	113
A.2. Gibbs sampler	114

Parte I

Modelización de una imagen

Introducción

El modelado de imágenes implica la construcción de modelos o procedimientos para la especificación de imágenes. Estos modelos pueden describir imágenes que se observan y también permiten generar imágenes sintéticas a partir de los parámetros del modelo.

Los campos aleatorios de Gibbs Markov (Gibbs Markov random field - GMRF) constituyen una herramienta poderosa para la extracción de información espacial contenida en una imagen, dado que proporcionan una forma conveniente y consistente para modelar la dependencia del contexto de los píxeles de la imagen, dando un aspecto visual característico que se entiende como textura.

La estimación de los parámetros asociados con un modelo de textura ha sido objeto de interés para los investigadores, debido a que este es el elemento clave para caracterizar el contenido de una imagen y generar texturas sintéticas. Besag [4], en su artículo fundacional, introdujo un procedimiento de codificación para la estimación de parámetros en GMRF, definido sobre una malla rectangular en el contexto de datos binarios. Este proceso de codificación consiste en etiquetar los sitios de la malla con diferentes códigos alternativamente. De esta manera, la malla se divide en conjuntos disjuntos llamados codificaciones, en inglés *codings*. Luego, por la propiedad Markoviana, las variables asociadas con los sitios de un conjunto de codificación sobre la malla son independientes, dados los valores observados en los sitios restantes. En consecuencia, la probabilidad condicional para cada conjunto de codificación es el producto de las probabilidades locales sólo considerando aquellos sitios que constituyen el conjunto de codificación. Luego, se obtiene una estimación del vector de parámetros maximizando la probabilidad condicional; y finalmente, los resultados de cada conjunto de codificación se combinan adecuadamente, para obtener la estimación final del vector de parámetros.

Este esquema de codificación se ha empleado en diferentes trabajos aplicados al procesamiento de imágenes [17, 34, 57] para imágenes binarias y en nivel de gris. En 2012, Kaiser et al. [35] extendieron el enfoque de codificación de Besag a configuraciones más amplias que incluyen grafos y mallas, así como otras mallas irregulares, y renombrando a estos conjuntos de codificación *concliqués*. Recientemente, Kaplan et al. 2019 [36] propusieron un nuevo algoritmo Gibbs

Sampler (GS), basado en el marco de Kaiser, denominado Concliqué-Gibbs Sampler (CGS).

En esta primera parte de la tesis, se presenta un teorema de descomposición de un potencial de Gibbs en configuraciones basadas en el esquema de codificación de Besag y posteriormente, se propone un enfoque novedoso del estimador de Mínimo Cuadrado Condicional (MCC) [40] basado en el esquema de codificación de Besag, que permite estimar el vector de parámetros de un modelo de textura de Gibbs-Markov, el modelo Auto-Binomial (MAB). Este procedimiento de estimación, denominado estimador de Mínimo Cuadrado Condicional Coding (MCCC), se implementa en un algoritmo paralelo. Con motivo de evaluar su performance, a partir de un estudio Monte-Carlo, se realiza un análisis descriptivo, se testea su sensibilidad ante presencia de ruido, se evalúan los algoritmos y por último, se presenta una aplicación real en el contexto de la clasificación supervisada de imágenes satelitales.

Capítulo 1

Modelado de imágenes de textura

En este capítulo se presenta el marco teórico concerniente al modelado estocástico de una imagen de textura, desarrollado en el libro de Georgii (1988)[24], las notas de Bustos et al. (2017)[12], como así también en Bustos et al. (1998)[10], Bustos y Guerrero (2011)[9] y Rulloni (2014)[62]. Al final, se demuestra un teorema de descomposición de un potencial de Gibbs basado en los conjuntos coding de Besag (Martinez y Pistonesi (2020)[46]).

1.1. Texturas

En una imagen, la textura es una de las características visuales que permite identificar un sector como perteneciente a una determinada clase. Reconocemos muchas partes de una imagen por la textura en lugar de su forma, por ejemplo: el cabello, la piel, diferentes tipo de tejido o una zona cultivada de la que no, como se ilustra en la Figura 1.1.



Figura 1.1: Imágenes de texturas: Julián y Viñedo.

Si la textura pertenece a una clase que tiene una interpretación física particular, como es la corteza de un árbol, agua, arena o cabello, entonces puede considerarse como una textura “natural”. Por otro lado, una textura puede pertenecer a una clase identificada por las características visuales artificiales, que tienen una interpretación matemática concisa. Ver Figura 1.2. No existe

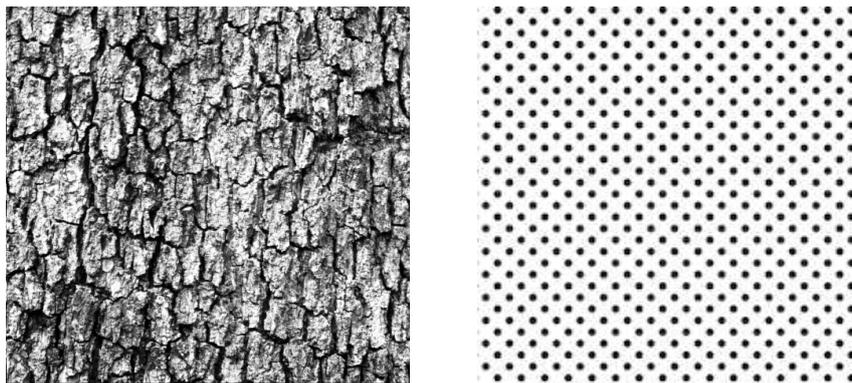


Figura 1.2: Imágenes de texturas: Corteza de un árbol (“Bark”) y lunares negros, respectivamente.

una definición universal de textura, ya que no existe un modelo matemático de textura que lo abarque todo. No obstante, desde una perspectiva humana, podemos conjeturar que la textura es una cualidad que distingue la regularidad en la apariencia visual de los detalles locales [17].

Dentro de los modelos matemáticos de texturas existentes, se encuentran los modelos estocásticos que persiguen tanto la descripción de texturas observadas como la síntesis de texturas artificiales. Algunos de estos modelos son válidos para caracterizar la textura en vecindades pequeñas en una única escala, es decir, sólo pueden utilizarse de manera eficiente para el análisis de microtexturas y otros tipos, llamadas macrotexturas, son capaces de caracterizar de manera eficiente texturas a diferentes escalas, combinando características espaciales y frecuenciales de la imagen de textura.

Los modelos de Campos Aleatorios de Markov (CAM) permiten generar y representar texturas de una imagen. En estos modelos, los parámetros caracterizan dos aspectos importantes: especifican las propiedades locales que determinan los niveles de grises de una determinada área y la relación existente entre los niveles de grises de las áreas que forman la imagen.

1.2. Imágen de textura

Una imagen (discreta) se puede definir como la representación de información organizada espacialmente en una grilla o malla bidimensional. De esta forma, el concepto de imagen se muestra equivalente al arreglo bidimensional de datos. Las celdas de esta grilla son conocidas como píxeles uniformemente espaciados y contienen datos ordenados según la ubicación de procedencia. Estos datos pueden provenir de la digitalización de algún proceso óptico o bien de la recolección de información espacial. Se define el espacio de estados como el conjunto de valores posibles que tiene cada pixel y se lo denota E . En la práctica, E es un conjunto finito. En una imagen binaria cada dato tiene sólo dos valores posibles (dos estados), como por ejemplo 0 y 1, pudiendo representar la ausencia y presencia local de alguna característica de interés, mientras

que las imágenes a escala de grises cada dato puede tomar un valor entero de 0 (negro) a 255 (blanco).

De ahora en más, se considera a E un conjunto finito, la métrica discreta (en E) y \mathcal{E} su σ -álgebra de Borel, con lo cual $\mathcal{E} = \mathcal{P}(E)$. Se define el soporte de una imagen como la grilla de pixeles en donde está definida la imagen, y se denota $S \subseteq \mathbb{Z}^2$ ($S \neq \emptyset$ no necesariamente finito). Una observación de una imagen no se puede lograr sobre S , pero si sobre un subconjunto finito $\Lambda \subset S$. Desde el punto de vista formal, una imagen es una matriz

$$x = [x_{ij}] = \begin{bmatrix} x_{11} & \cdots & x_{1N} \\ \vdots & \ddots & \\ x_{M1} & & x_{MN} \end{bmatrix}$$

en el caso de soporte regular finito (i.e. $\Lambda = ([1, M] \times [1, N] \cap \mathbb{Z}^2)$) y un arreglo $x = \{x_s\}_{s \in S}$, en el caso general.

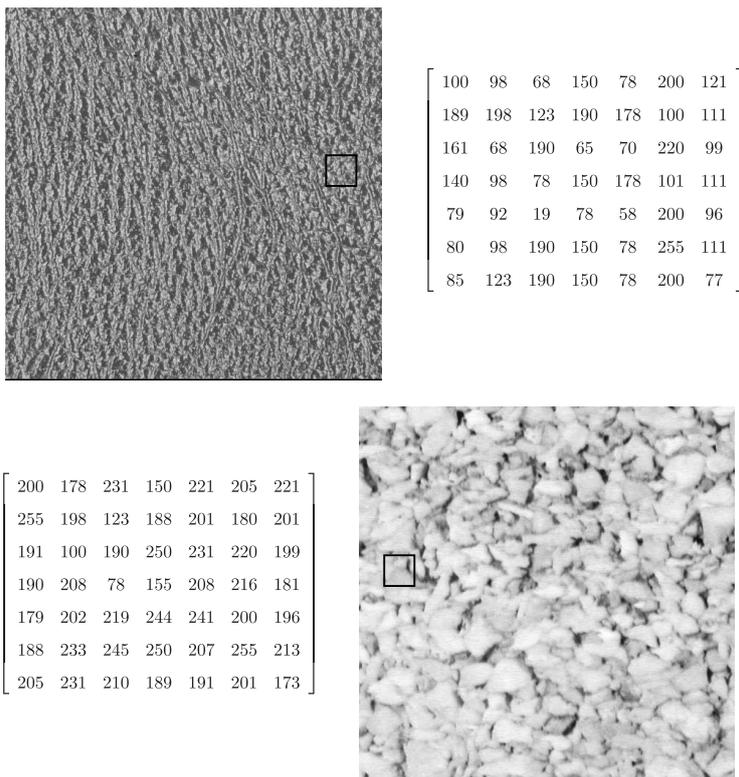


Figura 1.3: Imágenes de textura: “Pressed calf leather” “Gravell”, perteneciente a la base de datos USC-SIPI. Para el sector seleccionado se indica los valores de niveles de grises correspondientes.

También una imagen puede considerarse como una función, donde $x(s) = x_s \in E$, para $s \in S$. Se denota E^S al conjunto de imágenes con soporte S . Caso particular:

$$E^V := \{x/x : V \rightarrow E\},$$

conjunto de imágenes con soporte V y espacio de estados E ($\emptyset \neq V \subseteq S$). De ahora en más se llamará imagen a la función, a la matriz o a la representación visual de manera indistinta, entendiéndose cada alusión en el contexto que se realice.

Dada una imagen $x \in E^S$:

- Su restricción a V (o proyección E^V , $\emptyset \neq V \subseteq S$) se define y denota: $x_V \in E^V$ tal que $x_V(s) \forall s \in V$; o bien $x = \{x_s\}_{s \in V}$.
- Su traslación por $t \in S$ se define y denota $x^t \in E^S$ tal que $x^t(s) = x(s-t)$.

Definición 1.2.1 Sean $\emptyset \neq V \subseteq \Delta \subseteq \mathbb{Z}^2$, se define la proyección $\sigma_{\Delta, V} : E^\Delta \rightarrow E^V$:

$$\sigma_{\Delta, V}(x) = x_V, x \in E^\Delta$$

Notación 1.2.1 σ_V en lugar de $\sigma_{S, V}, \sigma_{\Delta, s}$, en lugar de $\sigma_{\Delta, \{s\}}$, para $s \in \Delta$, y σ_s en lugar de $\sigma_{S, s}$.

Definición 1.2.2 Sean $z \in E^{V_1}$ y $w \in E^{V_2}$, para $\emptyset \neq V_1 \subseteq S$, $\emptyset \neq V_2 \subseteq S$ y $V_1 \cap V_2 = \emptyset$, se define la imagen $zw \in E^{V_1 \cup V_2}$ ($zw : V_1 \cup V_2 \rightarrow E$) como la función que cumple

$$zw(t) = \begin{cases} z(t) & t \in V_1 \\ w(t) & t \in V_2 \end{cases}$$

Sea \mathcal{E}^V la σ -álgebra producto de E^V . Para simplificar la notación:

$$\mathcal{F} := \mathcal{E}^S$$

Para cada $B \in \mathcal{E}^V$, se define el conjunto $\sigma_V^{-1}(B) := \{w \in E^S / \sigma_V(w) \in B\}$ y para cada $x_V \in E^V$, el conjunto $(\sigma_V = x_V) = \{w \in E^S / \sigma_V(w) = x_V\}$

Sea ahora \mathcal{F}_V la σ -álgebra de E^S definida por:

$$\mathcal{F}_V := \{\sigma_V^{-1}(B) / B \in \mathcal{E}^V\}$$

Se denota:

$$\mathcal{I}_V := \mathcal{F}_{S \setminus V}$$

Dado que S puede ser infinito, es conveniente indentificar el conjunto de subsoportes finitos:

$$\mathcal{S} := \{\Lambda \subset S / 1 \leq \#(\Lambda) < \infty\}.$$

1.3. Modelado probabilístico

El modelado estadístico asume que una imagen $x \in E^S$ es una realización de un campo aleatorio X .

Definición 1.3.1 Sea (Ω, \mathcal{D}) un espacio medible, $X : \Omega \rightarrow E^S$, se dice que X es campo aleatorio (sobre E^S) si X es \mathcal{D} -medible (es decir $X^{-1}(A) \in \mathcal{D}, \forall A \in \mathcal{F}$).

Definición 1.3.2 Sea (Ω, \mathcal{D}) un espacio medible, $X : \Omega \rightarrow E^S$ campo aleatorio y sea P una probabilidad sobre (Ω, \mathcal{D}) , se denota y define P_X , la probabilidad sobre (E^S, \mathcal{F}) que cumple:

$$P_X(A) := P(X \in A), \forall A \in \mathcal{F}.$$

Para cada $V \subset S$ y $s \in S$ sean los campos aleatorios (sobre E^V y E , respectivamente):

- $X_V : \Omega \rightarrow E^V$,
- $X_s : \Omega \rightarrow E$, en lugar de $X_s : \Omega \rightarrow E^{\{s\}}$

definidos por:

- $X_V := \sigma_V \circ X$, es decir $X_V(\omega) = x_V$, si $X(\omega) = x$
- $X_s(\omega) = x_s$, si $X(\omega) = x$ para todo $\omega \in \Omega$.

Es necesario definir (o construir) una medida μ en E^S que describa la probabilidad de que $A \in \mathcal{F}$ sea un conjunto de “realizaciones” de X , es decir $\mu(A) = P(X \in A)$ (o bien $\mu(\{x\}) = P_X(\{x\}) = P(X = x)$ cuando S es finito). Esta medida μ debe modelar la posible interacción de cada pixel con su entorno. En la práctica, la relación de cada pixel con su entorno se torna independiente a medida que uno se aleja de él, circunscribiéndose sólo a un vecindario. Los campos aleatorios que cumplen con esta propiedad son llamados campos markovianos (Markov random fields (MRFs) en inglés).

Cuando S es finito, el teorema de equivalencia de Hammersley-Clifford (1971) [27] nos dice que cualquier campo markoviano posee una medida de Gibbs y viceversa: cualquier campo que posea una medida de Gibbs es un campo markoviano. Entonces es válido derivar métodos y algoritmos para campos aleatorios de Markov usando la teoría de campos aleatorios de Gibbs.

Una medida de Gibbs cuando S es finito tiene la forma:

$$\mu(\{x\}) = P(X = x) = \frac{\exp(-H(x))}{Z}$$

en donde Z es la constante de normalización, conocida también como función de partición, y H la función de energía que debe contener la información “relevante” de cada x . La información “relevante” de interés es la otorgada por la interacción de cada pixel de la imagen con su entorno y se describe con unas funciones “locales” $(\Phi_\Lambda, \Lambda \in \mathcal{S})$ llamadas potenciales, tal que:

$$H(x) = \sum_{\Lambda \in \mathcal{S}} \Phi_\Lambda(x)$$

Para S posiblemente infinito la construcción de una medida de Gibbs-Markov no es trivial sino que se la construye de manera local a través de funciones llamadas especificaciones. Se procura la existencia y unicidad de una probabilidad global tal que las especificaciones se comporten como las condicionales de esta probabilidad y localmente como medidas de Gibbs-Markov con soporte finito. Se busca un conjunto de propiedades que la caractericen y ayuden a definirla.

En las siguientes secciones y subsecciones se presentan definiciones y notaciones específicas de cada tema que titula la sección.

1.4. Distribuciones de Gibbs

Sea S no necesariamente finito (general) y se consideran, de ahora en más, todas las definiciones, supuestos y notaciones de la sección anterior.

Definición 1.4.1 *Se dice que la colección $\Phi := (\Phi_\Lambda)_{\Lambda \in \mathcal{S}}$ es un potencial si para todo $\Lambda \in \mathcal{S}$:*

- $\Phi_\Lambda : E^S \rightarrow \mathbb{R}$ es \mathcal{F}_Λ -medible (esto es $\Phi_\Lambda(x) = \phi_\Lambda(x_\Lambda)$ con ϕ_Λ función \mathcal{E}^Λ -medible)
- Existe $H_\Lambda(x) := \sum_{\Delta \in \mathcal{S} \cap \Lambda} \Phi_\Delta(x) < \infty$ para cada $x \in E^S$. H_Λ se define como la función de energía local para Φ , siendo $\mathcal{S} \cap \Lambda := \{\Delta \in \mathcal{S} / \Delta \cap \Lambda \neq \emptyset\}$

Se definen estas funciones con dependencias locales $\{\Phi_\Delta : \Delta \in \mathcal{S} \cap \Lambda\}$ llamadas funciones potenciales con el objetivo de poder separar las diferentes interacciones locales presentes globalmente en la función de energía. Se tiene una función potencial Φ_Λ para cada subconjunto $\Lambda \in \mathcal{S}$ (Si S es finito el número de funciones potenciales es igual al número de subconjunto de S , que es $2^{\#S}$).

Definición 1.4.2 *Un potencial Φ es λ -admisibles, $\lambda \in \mathcal{M}(E, \mathcal{E})$, si*

$0 < \int_{E^\Lambda} \exp(-H_\Lambda(\xi_{S \setminus \Lambda})) \lambda^\Lambda(d\xi) < \infty, \forall x \in E^S, \Lambda \in \mathcal{S}$. donde λ^Λ es la única medida sobre $(E^\Lambda, \mathcal{E}^\Lambda)$ tal que si $\Lambda = \{s_1, \dots, s_n\}$ entonces: $\lambda^\Lambda \left(\sigma_{\Lambda, \{s_1\}}^{-1}(B_1) \cap \dots \cap \sigma_{\Lambda, \{s_n\}}^{-1}(B_n) \right) = \lambda(B_1) \dots \lambda(B_n)$ cualesquiera sean B_1, \dots, B_n en \mathcal{E} .

Definición 1.4.3 Sea Φ un potencial λ -admisibles, se define la función de partición local de Φ sobre Λ y se denota por Z_Λ (para cada $\Lambda \in \mathcal{S}$) a la que cumple:

$$Z_\Lambda(x) := \int_{E^\Lambda} \exp(-H_\Lambda(\xi x_{S \setminus \Lambda})) \lambda^\Lambda(d\xi), \text{ para todo } x \in E^S \text{ y } \Lambda \in \mathcal{S}. \quad (1.1)$$

Notación 1.4.1 Para cada $s \in S$ se pondrá: H_s en lugar de $H_{\{s\}}$ y Z_s en lugar de $Z_{\{s\}}$

Definición 1.4.4 Sea Φ un potencial λ -admisibles, se define entonces $\rho_\Lambda : E^S \rightarrow [0, +\infty)$ (para cada $\Lambda \in \mathcal{S}$), tal que:

$$\rho_\Lambda(x) := \frac{\exp(-H_\Lambda(x))}{Z_\Lambda(x)}, \forall x \in E^S. \quad (1.2)$$

Definición 1.4.5 Para cada $s \in S$ se denota y define la característica local $\rho_s := \rho_{\{s\}}$, es decir:

$$\rho_s(x) = \frac{\exp(-H_s(x))}{Z_s(x)}, \forall x \in E^S. \quad (1.3)$$

Definición 1.4.6 Sea Φ un potencial λ -admisibles, para cada $\Lambda \in \mathcal{S}$ sea $\gamma_{\Phi, \Lambda} : \mathcal{F} \times E^S \rightarrow [0, 1]$ definida por

$$\gamma_{\Phi, \Lambda}(A|x) := \int_{E^\Lambda} 1_A(\xi x_{S \setminus \Lambda}) \rho_{\Phi, \Lambda}(\xi x_{S \setminus \Lambda}) \lambda^\Lambda(d\xi), \forall A \in \mathcal{F} \text{ y } x \in E^S,$$

tal que verifica:

- Para $x \in E^S$, la función $\gamma_{\Phi, \Lambda}(\cdot|x)$ es una probabilidad sobre (E^S, \mathcal{F}) ,
- Para $A \in \mathcal{F}$, la función $\gamma_{\Phi, \Lambda}(A|\cdot)$ es una función \mathcal{I}_Λ -medible,
- $B \in \mathcal{I}_\Lambda$, $x \in E^S$ entonces $\gamma_{\Phi, \Lambda}(B|x) = 1_B(x)$,
- $\Delta \subseteq \Lambda$, $A \in \mathcal{F}$ entonces $\gamma_{\Phi, \Lambda}(A|\cdot) = \gamma_{\Phi, \Lambda} \gamma_{\Phi, \Delta}(A|\cdot) := \int_{E^S} \gamma_{\Phi, \Delta}(A|w) \gamma_{\Phi, \Lambda}(dw|\cdot)$.

Se dice que $\gamma_{\Phi, \Lambda} := (\gamma_{\Phi, \Lambda})_{\Lambda \in \mathcal{S}}$ es una especificación inducida por el potencial Φ .

Si E es finito, λ una medida de conteo y $A = \sigma_\Lambda^{-1}(x_\Lambda)$, de acuerdo a Georgii (1988) (Observación 1.28 (5))[24], se obtiene la siguiente propiedad:

$$\rho_\Lambda(x) = \gamma_{\Phi, \Lambda}(\sigma_\Lambda^{-1}(x_\Lambda)|x) \quad \forall x \in E^S \quad \forall \Lambda \in \mathcal{S} \quad (1.4)$$

El conjunto de medidas globales de Gibbs para Φ se define:

$$\mathcal{G}(\Phi) := \left\{ \nu \in \mathcal{P}(E^S, \mathcal{F}) / \int_B \gamma_{\Phi, \Lambda}(A|x) \nu(dx) = \nu(B \cap A), \forall A \in \mathcal{F}, B \in \mathcal{I}_\Lambda \right\}.$$

Si $\mu \in \mathcal{G}(\Phi)$ se dirá que es una medida global de Gibbs para Φ . Se cumple que $\gamma_\Lambda(\cdot|\cdot)$ es una probabilidad condicional regular dada \mathcal{J}_Λ para μ (en símbolos $\gamma(A|\cdot) \in \mu(A|\mathcal{J}_\Lambda)$), es decir que $\gamma_\Lambda(A|x)$ es la probabilidad de A con respecto a μ condicionada a $x_{S \setminus \Lambda}$. Si $\mu \in \mathcal{G}(\Phi)$, μ es la medida que se quería construir.

Definición 1.4.7 Sea (Ω, \mathcal{D}, P) espacio de probabilidad, Φ potencial y $X : \Omega \rightarrow E^S$ campo aleatorio. Se dice que X es un campo con distribución de Gibbs para Φ si $P_X \in \mathcal{G}(\Phi)$ (esto es, P_X medida de Gibbs para Φ).

Nota 1.4.1 $\#\mathcal{G}(\Phi)$ puede ser vacío, o tener un solo elemento o tener infinitos elementos (transición de fase).

A continuación veremos una clase particular de potencial para que $\mathcal{G}(\Phi) \neq \emptyset$ (existencia).

Definición 1.4.8 Un potencial Φ se dice acotado si $\Phi_\Lambda \in \mathcal{L}^\infty$, para todo Λ^1

Definición 1.4.9 Un potencial Φ acotado, se dice sumable si $\sum_{\Lambda \in \mathcal{S}_{\{s\}}} \|\Phi_\Lambda\|_{\mathcal{L}^\infty} < \infty$. Siendo $K = \{\Phi/\Phi \text{ es sumable}\}$, el conjunto de los potenciales sumables.

Según Georgii (1988) (Teorema 4.23) [24], si (E, \mathcal{E}) es un espacio de Borel estándar y $\lambda \in \mathcal{M}(E, \mathcal{E})$ es finito, entonces para todo $\Phi \in K$, $\mathcal{G}(\Phi)$ es no vacío y compacto. A partir de este resultado se deduce que

$$\text{si } \Phi \text{ es sumable} \Rightarrow \mathcal{G}(\Phi) \neq \emptyset.$$

Puesto que en este trabajo se considera a E con la métrica discreta, un espacio métrico separable y completo ; $\mathcal{E} = \mathcal{P}(E)$ su σ -álgebra de Borel; λ , la medida de conteo definida en (E, \mathcal{E}) σ -finita y γ , la especificación inducida por Φ y λ , en consecuencia, la existencia de una medida global de Gibbs está asegurada cuando el potencial es sumable.

Definición 1.4.10 Para $\Lambda \in \mathcal{S}$ se define el núcleo de probabilidad (de $(E^S, \mathcal{J}_\Lambda)$ a $(E^\Lambda, \mathcal{E}^\Lambda)$) γ_Λ^0 tal que:

$$\gamma_\Lambda^0(B|x) := \gamma_\Lambda(\sigma_\Lambda \in B|x) = \gamma_\Lambda(\sigma_\Lambda^{-1}(B)|x), x \in E^S, B \in \mathcal{E}^\Lambda$$

Notación 1.4.2 γ_s^0 en lugar de $\gamma_{\{s\}}^0$, para $s \in S$.

Definición 1.4.11 Sea $V \subset S$ y $t \in S$ tal que $t + V = V$ se define $\tau_t : E^V \rightarrow E^V$ la función traslación (por t) evaluada en $x \in E^V$ se denota:

$$\tau_t(x) := x^t \in E^V$$

¹ $\mathcal{L}^\infty := \mathcal{L}^\infty(E^S, \mathcal{F}, \mathbb{R}) := \{f : E^S \rightarrow \mathbb{R} / \|f\|_{\mathcal{L}^\infty} < \infty\}$

y cumple $\tau_t(x)(s) = x^t = x(s-t), \forall s \in V$. Es decir, cuando $\mathbf{0} := (0,0) \in V$, el valor de x^t en el pixel t es el mismo que x en el pixel $\mathbf{0}$ (cero).

Se asume que $t + S = S$ para todo $t \in S$

Definición 1.4.12 Se dice que un potencial Φ es invariante a traslaciones si siempre se cumple: $\Phi_\Lambda(x) = \Phi_{t+\Lambda}(x^t)$.

Definición 1.4.13 Sea $k \in \mathbb{N}, \Theta \subset \mathbb{R}^k$ un conjunto de parámetros compacto con interior no vacío ($\text{int}(\Theta) \neq \emptyset$) llamado espacio paramétrico y sea $\Phi := (\Phi_\Lambda)_{\Lambda \in \mathcal{S}}$ tal que $\Phi_\Lambda : \Theta \times \mathcal{E}^S \rightarrow \mathbb{R}$. Si $\Phi^\underline{\theta} := (\Phi_\Lambda(\underline{\theta}, \cdot))_{\Lambda \in \mathcal{S}}$ es un potencial para cada $\underline{\theta} \in \Theta$, se dice que $(\Phi^\underline{\theta})_{\underline{\theta} \in \Theta}$ es una familia de potenciales parametrizada por Θ y que el potencial $\Phi^\underline{\theta}$ está parametrizado por el vector de parámetro $\underline{\theta}$. Θ se dice identificable si

$$\underline{\theta}_1 \in \Theta, \underline{\theta}_2 \in \Theta, \underline{\theta}_1 \neq \underline{\theta}_2 \Rightarrow \mathcal{G}(\Phi^{\underline{\theta}_1}) \cap \mathcal{G}(\Phi^{\underline{\theta}_2}) = \emptyset$$

La identificación de parámetros resulta indispensable en la discriminación y caracterización de texturas en imágenes.

1.5. Campos Markovianos

A continuación se define el entorno de influencia de cada pixel, para ello, se considera:

Definición 1.5.1 Para cada $s \in S$, sea $\partial s \in \mathcal{S}$. Se dice que $\partial = \{\partial s / s \in S\}$ es un sistema de entornos finitos de S si:

- $s \notin \partial s$ y
- $s \in \partial t \Leftrightarrow t \in \partial s$

∂s se denomina entorno de s y $\partial \Delta := \bigcup_{s \in \Delta} \partial s \setminus \Delta$ entorno de Δ y $\Delta^* := \Delta \cup \partial \Delta$.

Al par $G := (S, \partial)$, lo llamaremos grafo (no dirigido) sobre S con sistema de entorno ∂ .

Se asume que el sistema de entornos es homogéneo (invariante a traslaciones). Es decir, conociendo ∂s se puede conocer ∂t desplazando el vecindario de s : $\partial t = (t-s) + \partial s$, se asume $0 \in S$, luego $\partial s = s + \partial 0, \forall s \in S$.

Entre los elementos de $\Lambda \in \mathcal{S}$, puede definirse un orden, mediante el cual el sistema de entornos ∂ puede determinarse más explícitamente.

Definición 1.5.2 Sea $\delta \in D$, $D := \{i^2 + j^2 : i \geq 0, j \geq 0, (i, j) \neq (0, 0), i, j \in \mathbb{N}\}^2$ Sea $s \in S$, se define el orden δ del sistema de entorno, si verifica $\partial^\delta s = \{t \in S : 0 < \|s - t\|^2 \leq \delta\}$.

Definición 1.5.3 Se denomina cliqué, a un conjunto de píxeles en el que todo par de píxeles son mutuamente vecinos o está constituido por un solo píxel. El número y el tipo de cliques está vinculado a la estructura de vecinos considerada. Entonces, sea ∂ un sistema de entornos, se define y denota el conjunto de los ∂ -cliques como: $C := C(\partial) := \{(\Lambda \in \mathcal{S} : \#(\Lambda) = 1) \text{ o } (s \in \Lambda, t \in \Lambda, s \neq t \Rightarrow s \in \partial t)\}$

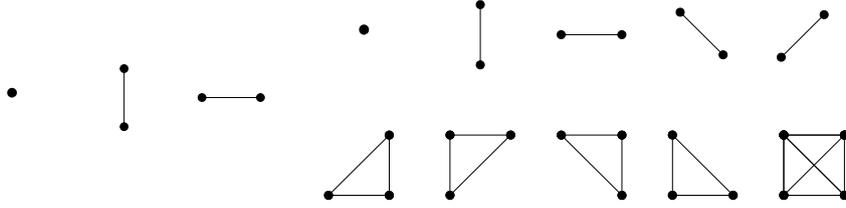


Figura 1.4: Tipos de cliques para el vecindario de orden 1 y 2 en el modelo de textura, respectivamente.

Definición 1.5.4 Se dice que Φ es un ∂ -potencial, potencial de vecindario, si $\Phi_\Lambda = 0$ para todo Λ que no pertenece al conjunto de cliques.

Notar que, si Φ es un ∂ -potencial, como $\partial 0$ es finito, Φ es sumable y por lo tanto $\mathcal{G}(\Phi) \neq \emptyset$.

Un conjunto importante de aplicaciones la constituyen las distribuciones de Gibbs con ∂ -potencial nulos fuera de cliques con tres ó más elementos, llamados los potenciales de a pares.

Definición 1.5.5 Sea $x \in E^S$ y $B \in \mathcal{E}^\Lambda$, dado ∂ un sistema de entornos, se dice que la especificación γ es ∂ -markoviana si la función $\gamma_\Lambda^0(B|x)$ depende de $x_{\partial\Lambda}$.

Definición 1.5.6 Sea $X : \Omega \rightarrow E^S$ campo aleatorio, se dice que es un campo de Gibbs-Markov si $P_X \in \mathcal{G}(\Phi)$ y además Φ es ∂ -potencial.

Por último, sea el espacio de estados E (finito), (Ω, \mathcal{D}, P) espacio de probabilidad, $X : \Omega \rightarrow E^S$ campo aleatorio, se conoce P_X a través de sus condicionales, las especificaciones $\gamma_\Lambda(\cdot|\cdot)$ ($\Lambda \in \mathcal{S}$). Sea $\Lambda \in \mathcal{S}$, dado que Φ es un ∂ -potencial, $\gamma_\Lambda(\sigma_\Lambda \in B|w)$ depende sólo de $w_{\partial\Lambda}$, es decir, $\gamma_\Lambda(\sigma_\Lambda \in B|\cdot)$ es $\mathcal{F}_{\partial\Lambda}$ -medible como función de $w \in E^S$.

Para cada $w_{\partial\Lambda} \in E^{\partial\Lambda}$, sea $P_{w_{\partial\Lambda}}$ la probabilidad sobre E^S tal que para cada $A \in \mathcal{F}$, $P_{w_{\partial\Lambda}}(A)$ (función evaluada en w , pero que sólo depende de $w_{\partial\Lambda}$) es una versión de $P(X \in A|\mathcal{F}_{\partial\Lambda})$. Dado que $\gamma_\Lambda(C|\cdot)$ es una versión de la probabilidad condicional $P(X \in C|\mathcal{F}_{S \setminus \Lambda})$ y $\gamma_\Lambda(\sigma_\Lambda \in B|\cdot)$ es $\mathcal{F}_{\partial\Lambda}$ -medible, $\gamma_\Lambda(\sigma_\Lambda \in B|\cdot)$ es una versión de la probabilidad condicional $P(X_\Lambda \in B|\mathcal{F}_{\partial\Lambda})$.

² $D := \{1, 2, 4, 5, 8, 9, \dots\}$

Luego, para cada $x_\Lambda \in E^\Lambda$, si $B = \{x_\Lambda\}$, $P_{w_{\partial\Lambda}}(X_\Lambda = x_\Lambda)$ y $\gamma_\Lambda(\sigma_\Lambda = x_\Lambda|w)$ coinciden (*P* c.s., pues son versiones de la probabilidad condicional $P(X_\Lambda = x_\Lambda|\mathcal{F}_{\partial\Lambda})$). De esta forma, se exhibe la relación entre las especificación y las probabilidades condiciones. Considerando la ecuación (1.4) se tiene que

$$P_{w_{\partial\Lambda}}(X_\Lambda = x_\Lambda) = \gamma_\Lambda(\sigma_\Lambda = x_\Lambda|w) = \rho_\Lambda(x) \quad \forall x \in E^S \quad \forall \Lambda \in \mathcal{F}.$$

1.6. Ejemplo: Modelo de Ising 2-D

En 1925, Ernest Ising, en su tesis doctoral [33], desarrolla un nuevo modelo para explicar teóricamente uno de los fenómenos más notables en la física del estado sólido, el ferromagnetismo. El modelo bidimensional de Ising consiste en un sistema de partículas fijas que forman una red periódica bidimensional, asociada a cada partícula, hay una variable spin que puede tomar únicamente dos estados: +1 (spin up) o -1 (spin down).

Sea ∂ el sistema de entornos de primer orden (Ver Figura 1.6) y $E = \{-1, 1\}$ el espacio de estados.

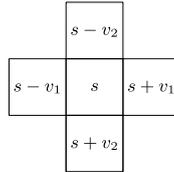


Figura 1.5: Sistemas de vecindades del pixel s primer orden, siendo $v_1 = (0, 1)$ y $v_2 = (1, 0)$.

El modelo de Ising con parámetro β (sin campo externo) es el definido por el ∂ -potencial (invariante por traslaciones)

$$\Phi_\Lambda(x) = \begin{cases} \beta x_t x_{t+v_i} & \Lambda = \{t, t+v_i\}, t \in S, 1 \leq i \leq 2 \\ 0 & \text{caso contrario} \end{cases} \quad \forall x \in E^S.$$

Su función de energía local en s es:

$$\begin{aligned} H_s(x) &= \beta \sum_{i=1}^2 (x_s x_{s-v_i} + x_s x_{s+v_i}), \\ &= \beta (\#\{t \in \partial s / x_t = x_s\} - \#\{t \in \partial s / x_t \neq x_s\}), \\ &= \beta (2\#\{t \in \partial s / x_t = x_s\} - 4). \end{aligned}$$

y

$$\rho_s(x) = \gamma_s^0(\{x_s\} | x) = \frac{\exp[-\beta (2\#\{t \in \partial s / x_t = x_s\} - 4)]}{\exp[-\beta (2\#\{t \in \partial s / x_t = x_s\} - 4)] + \exp[\beta (2\#\{t \in \partial s / x_t = x_s\} - 4)]}.$$

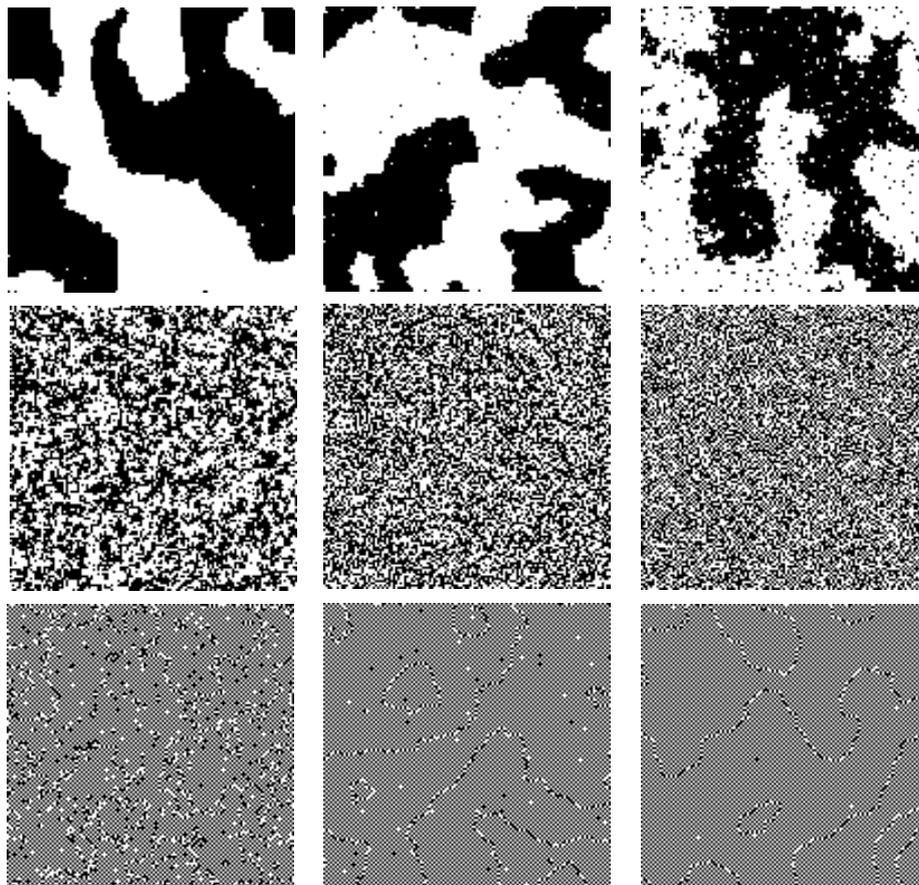


Figura 1.6: De izquierda a derecha y de arriba hacia abajo: nueve ejemplos de texturas generadas según el modelo de Ising 2D de primer orden mediante el algoritmo Gibbs Sampler considerando el parámetro $\beta = -1, -0.75, -0.5, -0.25, 0, 0.25, 0.5, 0.75$ y 1 , respectivamente.

Para los físicos, resulta de interés, un modelo de Ising más general, respecto al conjunto de partículas, al sistema de vecinos, pero con el mismo espacio de estados, donde la función de energía está dada por la siguiente expresión [79]:

$$H(x) = -\frac{1}{kT} \left[J \sum_{t \in \partial s} x_s x_t - mB \sum_s x_s \right],$$

donde la constante J pondera la interacción entre los vecinos cercanos, la constante B pondera la interacción de un campo externo, m es una constante positiva que depende del material considerado, k es el factor Boltzmann y T representa la temperatura absoluta.

1.7. Descomposición de un potencial de Gibbs en conjuntos Coding

A continuación se presenta un teorema que permite que la función de energía H se pueda descomponer en las contribuciones de las configuraciones en subconjuntos particulares del soporte S , llamados conjuntos Coding. Estos conjuntos determinan una partición de S . En el caso de

grillas regulares con entornos definidos, los conjuntos de codificación corresponden a los llamados conjuntos de codificación de Besag (1974) [4]. De esta manera, se obtiene la independencia condicional bajo el modelo GMRF entre los elementos de estos conjuntos de codificaciones, permitiendo la computación paralela.

Definición 1.7.1 Sea $G = (S, \partial)$ un grafo sobre S con sistema de entorno dado por ∂ . Entonces:

- Sea $A \in \mathcal{S}$ se llama ∂ -completo si: $\#(A) = 1$ o $(s \in A, t \in A, s \neq t) \Rightarrow s \in \partial t$.
- Sea $C \subset S$ con $\#(C) \geq 2$ se llama ∂ -coding, si $(s \in C, t \in C, s \neq t) \Rightarrow s \notin \partial t$,
- Una familia no vacía \mathcal{C} de subconjuntos ∂ -coding de S es una familia ∂ -coding, si $(C_1 \in \mathcal{C}, C_2 \in \mathcal{C}, C_1 \neq C_2) \Rightarrow C_1 \cap C_2 = \emptyset$ y $(\partial C_1 \cup \partial C_2) \subset S \setminus (C_1 \cup C_2)$.

1	2	1	2	1	2	1	2
2	1	2	1	2	1	2	1
1	2	1	2	1	2	1	2
2	1	2	1	2	1	2	1
1	2	1	2	1	2	1	2
2	1	2	1	2	1	2	1
1	2	1	2	1	2	1	2
2	1	2	1	2	1	2	1

1	2	1	2	1	2	1	2
3	4	3	4	3	4	3	4
1	2	1	2	1	2	1	2
3	4	3	4	3	4	3	4
1	2	1	2	1	2	1	2
3	4	3	4	3	4	3	4
1	2	1	2	1	2	1	2
3	4	3	4	3	4	3	4

Figura 1.7: Patrón de codificación según sistema de vecinos de primer (2 codings) y segundo (4 codings) orden, respectivamente.

Teorema 1.7.1 Sea $G = (S, \partial)$ un grafo, $\Phi = (\Phi_\Lambda)_{\Lambda \in \mathcal{S}}$ un potencial sobre (E^S, \mathcal{F}) , \mathcal{C} una familia no vacía de subconjuntos de S , que es ∂ -coding, $H_\Phi = (H_{\Phi, \Lambda})_{\Lambda \in \mathcal{S}}$ la familia de funciones de energía asociada a Φ , esto es:

$$H_{\Phi, \Lambda}(x) := \sum_{\Delta \in \mathcal{S} \cap \Lambda} \Phi_\Delta(x), \quad x \in E^S, \quad \Lambda \in \mathcal{S}.$$

Si:

- (i) Φ es un ∂ -potencial.
- (ii) Sean $C_1 \in \mathcal{C}, C_2 \in \mathcal{C}$, conjuntos no vacíos tal que $C_1 \neq C_2$.

Entonces

- (a) Para todo $\Lambda \in \mathcal{S}$, $H_{\Phi, \Lambda \cap (C_1 \cup C_2)} = H_{\Phi, \Lambda \cap C_1} + H_{\Phi, \Lambda \cap C_2}$

(b) Sea $h_\Phi = (h_{\Phi,\Lambda})_{\Lambda \in \mathcal{S}}$ la familia de funciones de Boltzmann asociadas a Φ , esto es, $h_{\Phi,\Lambda}(x) := \exp(-H_{\Phi,\Lambda}(x))$, $x \in E^S$, $\Lambda \in \mathcal{S}$. Entonces

$$h_{\Phi,\Lambda \cap (C_1 \cap C_2)} = h_{\Phi,\Lambda \cap C_1} \cdot h_{\Phi,\Lambda \cap C_2}, \Lambda \in \mathcal{S}.$$

(c) Sea $\lambda \in \mathcal{M}(E, \mathcal{E})$ y supongamos que Φ es λ -admisibles. Sean $\Lambda \in \mathcal{S}$, $x \in E^S$ y $\rho_\Lambda : E^S \rightarrow [0, +\infty)$ tal que

$$\rho_{\Phi,\lambda,\Lambda}(x) := \frac{h_{\Phi,\Lambda}(x)}{Z_{\Phi,\lambda,\Lambda}(x)}$$

donde

$$Z_{\Phi,\lambda,\Lambda}(x) := \int_{E^\Lambda} h_{\Phi,\Lambda}(\xi x_{S \setminus \Lambda}) \lambda^\Lambda(d\xi)$$

y para la cual se puede definir la especificación $\gamma_{\Phi,\Lambda} : \mathcal{F} \times E^S \rightarrow [0, 1]$ tal que

$$\gamma_{\Phi,\lambda,\Lambda}(A|x) := \int_{E^\Lambda} 1_A(\xi x_{S \setminus \Lambda}) \rho_{\Phi,\lambda,\Lambda}(\xi x_{S \setminus \Lambda}) \lambda^\Lambda(d\xi)$$

se cumplen (bajo (i) y (ii)),

$$(c.1) \quad Z_{\Phi,\lambda,\Lambda \cap (C_1 \cup C_2)}(x) = Z_{\Phi,\lambda,\Lambda \cap C_1}(x) \cdot Z_{\Phi,\lambda,\Lambda \cap C_2}(x), \forall x \in E^S, \forall \Lambda \in \mathcal{S}.$$

$$(c.2) \quad \rho_{\Phi,\lambda,\Lambda \cap (C_1 \cup C_2)}(x) = \rho_{\Phi,\lambda,\Lambda \cap C_1}(x) \cdot \rho_{\Phi,\lambda,\Lambda \cap C_2}(x), \forall x \in E^S, \forall \Lambda \in \mathcal{S}.$$

Con motivo de demostrar el Teorema 1.7.1, se establecen los siguientes tres lemas.

Para simplificar la notación, consideramos:

$$\Lambda_0 := \Lambda \cap (C_1 \cup C_2), \Lambda_1 := \Lambda \cap C_1 \text{ y } \Lambda_2 := \Lambda \cap C_2.$$

Lema 1.7.1 Sea $\Delta \in \mathcal{S}$ tal que $\Phi_\Delta \neq 0$, entonces no es posible que $\Delta \cap \Lambda_1 \neq \emptyset$ y $\Delta \cap \Lambda_2 \neq \emptyset$.

Demostración:

Sea $\Delta \in \mathcal{S}$ tal que $\Phi_\Delta \neq 0$ y supongamos que $\Delta \cap \Lambda_1 \neq \emptyset$ y $\Delta \cap \Lambda_2 \neq \emptyset$. Sea $s \in \Delta \cap \Lambda_1$ y $t \in \Delta \cap \Lambda_2$. Dado que $C_1 \cap C_2 = \emptyset$ (por (ii) Teorema 1.7.1) entonces $s \neq t$. Como $\Phi_\Delta \neq 0$, Δ es ∂ -completo, entonces como $s \in \Delta$, $t \in \Delta$ y $s \neq t$, $s \in \partial t$. Más aún, $t \in C_2 \subset S \setminus C_1$ entonces $t \in \partial C_1$. Pero $\partial C_1 \subset (S \setminus C_1) \cap (S \setminus C_2)$, entonces $t \notin C_2$. Una contradicción. \square

Lema 1.7.2 Sean $\Delta, \Lambda \in \mathcal{S}$, $\Lambda \subseteq \Delta$ y sea $f : E^S \rightarrow \mathbb{R}^m$, $m \in \mathbb{N}$, entonces:

(1) f es \mathcal{F}_Λ -medible $\Leftrightarrow f(x) = f(y) \forall x, y \in E^S$ tal que $\sigma_\Lambda(x) = \sigma_\Lambda(y)$,

(2) Si f es \mathcal{F}_Δ -medible, entonces la función $x \mapsto \int_{E^\Lambda} f(\xi x_{S \setminus \Lambda}) \lambda^\Lambda(d\xi)$ es $\mathcal{F}_{\Delta \setminus \Lambda}$ -medible.

Demostración:

Parte (1). Consideremos $x, y \in E^S$ tal que $\sigma_\Lambda(x) = \sigma_\Lambda(y)$. Si $f(x) \neq f(y)$, entonces existen dos conjuntos abiertos G_1 y G_2 in \mathbb{R}^m , tal que $f(x) \in G_1$ y $f(y) \in G_2$. Dado que f es \mathcal{F}_Λ -medible, se verifica que $f^{-1}(G_1), f^{-1}(G_2) \in \mathcal{F}_\Lambda$ y además $f^{-1}(G_1) \cap f^{-1}(G_2) = \emptyset$. Entonces, existen $B_1, B_2 \in \mathcal{E}^\Lambda, B_1 \cap B_2 = \emptyset$, tal que $f^{-1}(G_1) = \sigma_\Lambda^{-1}(B_1)$ e $f^{-1}(G_2) = \sigma_\Lambda^{-1}(B_2)$. Sin embargo, dado que $x \in f^{-1}(G_1)$ y $y \in f^{-1}(G_2)$, tenemos $\sigma_\Lambda(x) = \sigma_\Lambda(y) \in B_1 \cap B_2 = \emptyset$, lo cual es una contradicción.

Recíprocamente, sea $x \in E^S$, fijado $w \in E^S$ y dado un abierto G in \mathbb{R}^m , probaremos que f es \mathcal{F}_Λ -medible, consideremos $f(x) \in G \iff f(x_\Lambda w_{S \setminus \Lambda}) \in G \iff x_\Lambda \in \{y \in E^\Lambda : f(x_\Lambda w_{S \setminus \Lambda}) \in G\} \iff x \in \sigma_\Lambda^{-1}(\{y \in E^\Lambda : f(x_\Lambda w_{S \setminus \Lambda}) \in G\})$.

Entonces,

$$\begin{aligned} f^{-1}(G) &= \{x \in E^S : f(x) \in G\} \\ &= \sigma_\Lambda^{-1}(\{y \in E^\Lambda : f(x_\Lambda w_{S \setminus \Lambda}) \in G\}) \in \mathcal{F}_\Lambda \end{aligned}$$

Esto completa la demostración de la parte (1).

Parte (2). Sea $\Lambda \subseteq \Delta$, y $x, y \in E^S$ tal que $\sigma_{\Delta \setminus \Lambda}(x) = \sigma_{\Delta \setminus \Lambda}(y)$.

Por lo tanto, para cada $\xi \in E^\Lambda$, tenemos $f(\xi x_{S \setminus \Lambda}) = f(\xi y_{\Delta \setminus \Lambda} x_{S \setminus \Lambda \cup \Delta}) = f(\xi y_{S \setminus \Lambda})$ donde la última igualdad sigue a partir de la parte (1) y porque f es \mathcal{F}_Λ -medible. Entonces

$\int_{E^\Lambda} f(\xi x_{S \setminus \Lambda}) \lambda^\Lambda(d\xi) = \int_{E^\Lambda} f(\xi y_{S \setminus \Lambda}) \lambda^\Lambda(d\xi)$, tal que, por la parte (1), la proposición es probada. \square

Lema 1.7.3 Para cada $\Lambda \in \mathcal{S}$, entonces

(1) $H_{\Phi, \Lambda}(x)$ es \mathcal{F}_{Λ^*} -medible, $\forall \Lambda \in \mathcal{S}$ y con $\Lambda^* := \Lambda \cup \partial\Lambda$.

(2) Z_Λ is $\mathcal{F}_{\partial\Lambda}$ -medible.

(3) $\rho_{\Phi, \Lambda}$ is \mathcal{F}_{Λ^*} -medible.

Demostración

Parte (1). La demostración sigue del Lema 1.2.2 en Bustos et al.(2017) [12].

Sean x e $y \in E^S$ tales que $\sigma_{\Lambda^*}(x) = \sigma_{\Lambda^*}(y)$ y sea $\Delta \in \mathcal{S} \cap \Lambda$. Veamos que $\Phi_\Delta(x) = \Phi_\Delta(y) \forall \Delta \in \mathcal{S} \cap \Lambda$

Puesto que Φ es un ∂ -potencial (Ver definición 1.5.4), si $\Delta \notin C$ (conjunto cliqué) entonces $\Phi_\Delta(x) = 0 = \Phi_\Delta(y)$, pero si $\Delta \in C$, veamos que $\Delta \subset \Lambda^*$ ya que si esto vale, entonces por la parte (1) del Lema 1.7.2, se tiene también que $\Phi_\Delta(x) = \Phi_\Delta(y)$. Veamos que $\Delta \subset \Lambda^*$ entonces:

sea $t \in \Delta$. Si $t \in \Lambda$ entonces $t \in \Lambda^*$, pero si no, como $\Delta \in \mathcal{S} \cap \Lambda$, existe $s \in \Lambda \cap \Delta$. Luego $s \neq t$ y como $\Delta \in \mathcal{C}$, se tiene que $t \in \partial s \subseteq \Lambda^*$, luego $t \in \Lambda^*$.

Entonces $\Phi_\Delta(x) = \Phi_\Delta(y)$, y por la definición de H_Λ y por la parte (1) del Lema 1.7.2, esto completa la demostración de la parte (1).

Parte (2). Sea $x \in E^S$, $\Lambda \in \mathcal{S}$.

Por la Parte (1), $h_{\Phi, \Lambda^*}(x) = \exp(-H_{\Phi, \Lambda^*}(x))$ es \mathcal{F}_{Λ^*} -medible.

De esta manera, la función $x \mapsto \int_{E^\Lambda} h_{\Phi, \Lambda}(\xi x_{S \setminus \Lambda}) \lambda^\Lambda(d\xi)$ es $\mathcal{F}_{\Lambda^* \setminus \Lambda}$ -medible, por consiguiente, Z_Λ es $\mathcal{F}_{\partial\Lambda}$ -medible, donde $\partial\Lambda = \Lambda^* \setminus \Lambda$.

Esto completa la demostración de la parte (2).

Parte (3). La demostración sigue a partir de la definición de $\rho_{\Phi, \Lambda}$, y de la partes (1) y (2) de estos lemas. Esto completa la demostración. \square

Ahora probaremos el Teorema 1.7.1.

Demostración (del Teorema 1.7.1).

Parte (a).

$$\begin{aligned}
 H_{\Phi, \Lambda_0} &= \sum_{\Delta \in \mathcal{S} \cap \Lambda_0} \Phi_\Delta \\
 &= \sum_{\Delta \in \mathcal{S} \cap (\Lambda_1 \cup \Lambda_2)} \Phi_\Delta \\
 &= \sum_{\Delta \in \mathcal{S} \cap \Lambda_1} \Phi_\Delta + \sum_{\Delta \in \mathcal{S} \cap \Lambda_2} \Phi_\Delta \\
 &= H_{\Phi, \Lambda_1} + H_{\Phi, \Lambda_2}
 \end{aligned}$$

donde la tercera igualdad es válida por el Lema 1.7.1. Esto prueba la parte (a) del Teorema.

Parte (b). Sigue inmediatamente de la parte (a) del Teorema 1.7.1. En efecto, sea $\Lambda \in \mathcal{S}$ y $x \in E^S$,

$$\begin{aligned}
 h_{\Phi, \Lambda_0}(x) &= \exp(-H_{\Phi, \Lambda_0}(x)) \\
 &= \exp(-[H_{\Phi, \Lambda_1}(x) + H_{\Phi, \Lambda_2}(x)]) \\
 &= \exp(-H_{\Phi, \Lambda_1}(x)) \cdot \exp(-H_{\Phi, \Lambda_2}(x)) \\
 &= h_{\Phi, \Lambda_1}(x) \cdot h_{\Phi, \Lambda_2}(x).
 \end{aligned}$$

donde la segunda igualdad sigue a partir de la parte (a)

Parte (c). Ahora se prueba la parte (c.1). Sea $x \in E^S$, entonces

$$\begin{aligned}
 Z_{\Phi, \lambda, \Lambda_0}(x) &= \int_{E^{\Lambda_0}} h_{\Phi, \Lambda_0}(\xi x_{S \setminus \Lambda_0}) \lambda^{\Lambda_0}(d\xi) \\
 &= \int_{E^{\Lambda_1}} \int_{E^{\Lambda_2}} h_{\Phi, \Lambda_0}(\xi_1 \xi_2 x_{S \setminus \Lambda_0}) \lambda^{\Lambda_2}(d\xi_2) \lambda^{\Lambda_1}(d\xi_1) \\
 &= \int_{E^{\Lambda_1}} \int_{E^{\Lambda_2}} h_{\Phi, \Lambda_1}(\xi_1 \xi_2 x_{S \setminus \Lambda_0}) h_{\Phi, \Lambda_2}(\xi_1 \xi_2 x_{S \setminus \Lambda_0}) \lambda^{\Lambda_2}(d\xi_2) \lambda^{\Lambda_1}(d\xi_1) \\
 &= \int_{E^{\Lambda_1}} \int_{E^{\Lambda_2}} h_{\Phi, \Lambda_1}(\xi_1 x_{S \setminus \Lambda_1}) h_{\Phi, \Lambda_2}(\xi_2 x_{S \setminus \Lambda_2}) \lambda^{\Lambda_2}(d\xi_2) \lambda^{\Lambda_1}(d\xi_1) \\
 &= \int_{E^{\Lambda_1}} h_{\Phi, \Lambda_1}(\xi_1 x_{S \setminus \Lambda_1}) \lambda^{\Lambda_1}(d\xi_1) \cdot \int_{E^{\Lambda_2}} h_{\Phi, \Lambda_2}(\xi_2 x_{S \setminus \Lambda_2}) \lambda^{\Lambda_2}(d\xi_2) \\
 &= Z_{\Phi, \lambda, \Lambda_1}(x) \cdot Z_{\Phi, \lambda, \Lambda_2}(x).
 \end{aligned}$$

donde la tercera igualdad sigue a partir de la parte (b) del Teorema 1.7.1 y la cuarta igualdad vale, puesto que h_{Φ, Λ_1} is $\mathcal{F}_{\Lambda_1 \cup \partial \Lambda_1}$ -medible (ver Lema 1.7.3 (1)), y como $\partial C_1 \cup \partial C_2 \subset (S \setminus C_1) \cap (S \setminus C_2)$ y $C_1 \cap C_2 = \emptyset$, no es posible que $\xi_2 \in \Lambda_1 \cup \partial \Lambda_1$.

Luego, $h_{\Phi, \Lambda_1}(\xi_1 \xi_2 x_{S \setminus \Lambda_0}) = h_{\Phi, \Lambda_1}(\xi_1 x_{S \setminus \Lambda_1})$.

Análogamente, $h_{\Phi, \Lambda_2}(\xi_1 \xi_2 x_{S \setminus \Lambda_0}) = h_{\Phi, \Lambda_2}(\xi_2 x_{S \setminus \Lambda_2})$.

Parte (c.2). Es consecuencia directa de la parte (b) y parte (c.1) del Teorema 1.7.1. En efecto, sea $x \in E^S$ y $\Lambda \in \mathcal{S}$

$$\begin{aligned}
 \rho_{\Phi, \lambda, \Lambda_0}(x) &= \frac{h_{\Phi, \Lambda_0}(x)}{Z_{\Phi, \lambda, \Lambda_0}(x)} \\
 &= \frac{h_{\Phi, \Lambda_1}(x) \cdot h_{\Phi, \Lambda_2}(x)}{Z_{\Phi, \lambda, \Lambda_1}(x) \cdot Z_{\Phi, \lambda, \Lambda_2}(x)} \\
 &= \frac{h_{\Phi, \Lambda_1}(x)}{Z_{\Phi, \lambda, \Lambda_1}(x)} \cdot \frac{h_{\Phi, \Lambda_2}(x)}{Z_{\Phi, \lambda, \Lambda_2}(x)} \\
 &= \rho_{\Phi, \lambda, \Lambda_1}(x) \cdot \rho_{\Phi, \lambda, \Lambda_2}(x).
 \end{aligned}$$

donde la segunda igualdad sigue a partir de las partes (b) y (c.1)

De esta manera la prueba del Teorema 1.7.1 está completa. ■

Teniendo en cuenta este resultado, se propone un nuevo método de estimación del vector de parámetros de un modelo GMRF, Auto-binomial, que se desarrolla en el siguiente Capítulo 2.

Capítulo 2

Estimación de parámetros texturales

En este capítulo se presenta un modelo basado en los CAM, que puede capturar una amplia gama de texturas. Además, se desarrolla y evalúa un método de estimación de los parámetros del modelo de textura considerado, basado en los conjuntos coding de Besag (Ver Martinez et al.(2018) [49]).

2.1. Modelo de Textura: Autobinomial

Para caracterizar el contenido de una imagen compuesta por microtexturas o generar texturas sintéticas, el modelo más apropiado es el modelo Autobinomial (MAB), este es un modelo definido por ∂ -potenciales parametrizados e invariantes a traslaciones.

Consideremos el MAB, según Schröder et al. (1997) [64]

$$\Phi_{\underline{\theta}, \Lambda}(x) = \begin{cases} \beta_i \frac{x_t x_{t+v_i}}{M} & \Lambda = \{t, t + v_i\}, t \in S, 1 \leq i \leq r \\ -\log \binom{M}{x_t} + \beta_0 x_t & \Lambda = \{t\}, \quad \forall x \in E^S. \\ 0 & \text{caso contrario} \end{cases}$$

donde $t \in S$, el máximo valor del nivel de gris $M = 255$, $i = 1, \dots, r$ ($r = 2$ y 4 , corresponde al primer y segundo orden, con 4 y 8 vecinos, respectivamente), $\partial s = \{s \pm v_i\}_{i=1}^r$, $v_1 = (0, 1)$, $v_2 = (1, 0)$, $v_3 = (-1, 1)$, $v_4 = (1, 1)$, y además $\underline{\theta} = (\beta_0, \beta_1, \beta_2)$ y $\underline{\theta} = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4)$ es el vector de parámetro para el MAB de primer y segundo orden.

Dado que Φ es invariante por traslación [24], $\underline{\theta}$ no depende de $s \in S$.

Para $s \in S$, consideremos la siguiente definición

$$\begin{aligned} \gamma_s^0(B|x) &:= \gamma_s(B \times E^{S \setminus s} | x) \\ &= \frac{\sum_{\xi \in B} \exp(-H_s(\xi x_{S \setminus s}))}{\sum_{\xi \in E} \exp(-H_s(\xi x_{S \setminus s}))} \\ &= \frac{\sum_{\xi \in B} \exp(-\sum_{\Lambda \in s \cap \mathcal{S}} \Phi_\Lambda(\xi x_{S \setminus s}))}{\sum_{\xi \in E} \exp(-\sum_{\Lambda \in s \cap \mathcal{S}} \Phi_\Lambda(\xi x_{S \setminus s}))}, B \in \mathcal{E}^\Lambda. \end{aligned}$$

Por lo tanto, la característica local es

$$\begin{aligned} \rho_s(x) &= \gamma_s^0(\{x_s\} | x) \\ &= \frac{\binom{M}{x_s} \exp\left[-x_s \left(\beta_0 + \sum_{i=1}^r \beta_i \frac{x_{s-v_i} + x_{s+v_i}}{M}\right)\right]}{\left(1 + \exp\left[-\left(\beta_0 + \sum_{i=1}^r \beta_i \frac{x_{s-v_i} + x_{s+v_i}}{M}\right)\right]\right)^M} \\ &= \binom{M}{x_s} \omega_s^{x_s} (1 - \omega_s)^{M-x_s} \quad \forall x_s = 0, 1, \dots, M, \end{aligned} \quad (2.1)$$

con

$$\omega_s = \frac{1}{1 + \exp(-\eta_s)} \text{ y } \eta_s = \beta_0 + \sum_{i=1}^r \beta_i \frac{x_{s-v_i} + x_{s+v_i}}{M}, 1 \leq i \leq r. \quad (2.2)$$

Claramente, $\gamma_s^0(\cdot | x)$ depende de ∂s .

La magnitud η_s refleja la influencia conjunta de los vecinos ponderada por el vector de parámetros θ .

Para el modelo de primer orden (ver Figura 2.1),

$$\eta_s = \beta_0 + \beta_1 \frac{x_{s-v_1} + x_{s+v_1}}{M} + \beta_2 \frac{x_{s-v_2} + x_{s+v_2}}{M}. \quad (2.3)$$

En un modelo de segundo orden (ver Figura 2.1), η_s , se define de la siguiente manera

$$\begin{aligned} \eta_s &= \beta_0 + \beta_1 \frac{x_{s-v_1} + x_{s+v_1}}{M} + \beta_2 \frac{x_{s-v_2} + x_{s+v_2}}{M} \\ &+ \beta_3 \frac{x_{s-v_3} + x_{s+v_3}}{M} + \beta_4 \frac{x_{s-v_4} + x_{s+v_4}}{M}, \end{aligned} \quad (2.4)$$

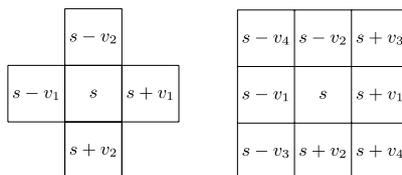


Figura 2.1: Sistemas de vecindades del píxel s , primer orden y segundo orden, respectivamente, siendo $v_1 = (0, 1)$, $v_2 = (1, 0)$, $v_3 = (-1, 1)$ y $v_4 = (1, 1)$

Por lo tanto, la variable aleatoria X_s modelando la intensidad de la imagen en el píxel s tiene una distribución binomial condicional con parámetros M y ω_s . El valor esperado y la varianza de X_s son respectivamente:

$$E[X_s] = M\omega_s = \frac{M}{1 + e^{-\eta_s}} \quad V[X_s] = M\omega_s(1 - \omega_s) = M \frac{e^{-\eta_s}}{(1 + e^{-\eta_s})^2}. \quad (2.5)$$

En las Tablas 2.1 y 2.2 se presentan los valores de los vectores de los parámetros de los casos bajo estudio del MAB de primer y segundo orden, respectivamente que serán considerados en las posteriores secciones.

Tabla 2.1: Valores considerados en el vector de parámetros para el MAB primer orden.

Casos	Parámetros		
	β_0	β_1	β_2
(i)	-0.5	1.5	-1
(ii)	1.5	-2	0.5
(iii)	-2	1	1

Tabla 2.2: Valores considerados en el vector de parámetros para el MAB segundo orden.

Casos	Parámetros				
	β_0	β_1	β_2	β_3	β_4
(i)	-2	1	1	-0.6	-0.6
(ii)	-2	1	-0.5	2	-0.5
(iii)	0	1	1	-1	-1

En las siguientes Figuras 2.2 y 2.3, se muestran seis imágenes de 128×128 bajo el MAB de primer y segundo orden con $E = \{1, \dots, 255\}$, generadas por los valores de los vectores de parámetros especificados en las Tablas 2.1 y 2.2, mediante el algoritmo Gibbs Sampler (GS) con 200 iteraciones [34] (suficiente para que las imágenes de textura se estabilicen. Ver Anéxo A).

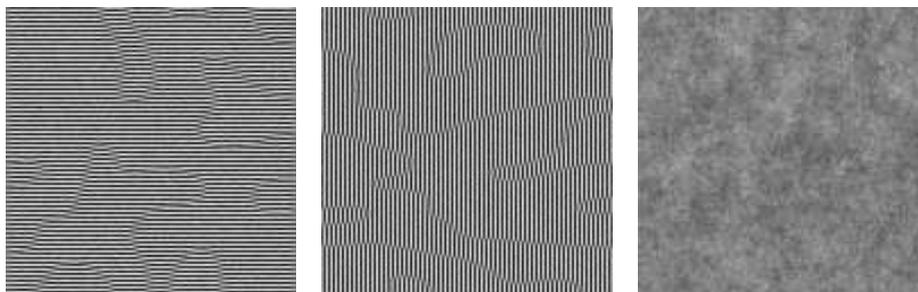


Figura 2.2: Tres ejemplos de realizaciones de MAB de primer orden usando los vectores de parámetros especificados correspondientes a los casos (i), (ii) y (iii), respectivamente.

En ellas se pueden visualizar diferentes patrones o texturas, puesto que los parámetros del MAB consideran las interacciones de los píxeles. El parámetro β_1 controla las direcciones horizontales pues multiplica cada producto $x_s x_{s+v_1}$; por ejemplo, en el primer panel de la Figura

2.2 se puede apreciar un patrón horizontal pues el valor del parámetro β_1 es 1.5 y los valores del resto de los parámetros son menores. De esta manera, los parámetros β_1 y β_2 consideran las interacciones horizontales y verticales, mientras que β_3 y β_4 las diagonales izquierda y derecha, respectivamente en los modelos de segundo orden [17].

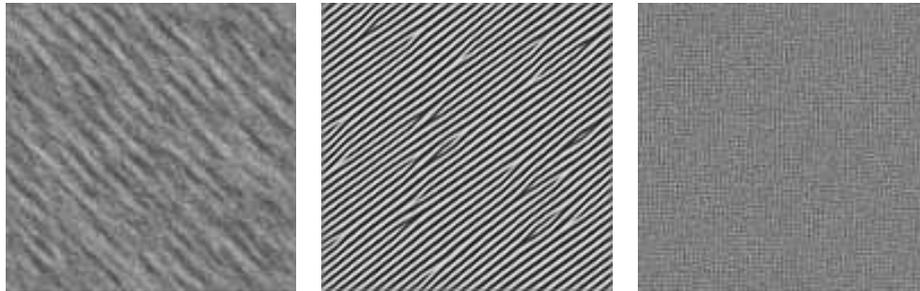


Figura 2.3: Tres ejemplos de texturas generadas según un MAB de segundo orden utilizando los vectores de parámetros especificados correspondientes a los casos (i), (ii) y (iii), respectivamente.

Si bien el modelo de primer orden es el más simple, restringe la cantidad de texturas posibles que se pueden generar (patrones horizontales y verticales). El modelo de segundo orden es más complejo, pero ofrece una variedad de texturas más amplia a los MAB de orden superior [17, 64].

La elección de los valores de los parámetros del MAB en este estudio, se basan en el hecho de que permiten generar texturas que presentan patrones típicos que identifican a los MAB de primer y segundo orden, respectivamente.

Las texturas generadas a través del MAB de primer orden, se caracterizan principalmente por presentar patrones horizontales y verticales, como exhiben los casos (i) y (ii), respectivamente. También, este modelo, permite generar texturas similares al caso (iii), donde no se observa un patrón de orientación definido. A través del MAB de segundo orden, se generan texturas que revelan patrones diagonales a izquierda y a derecha, como los que manifiestan los casos (i) y (ii), respectivamente. Además, puede generar texturas tipo clústers y texturas con patrones sin una orientación concreta, como se observa en el caso (iii).

2.2. Métodos de estimación

En esta sección se presentan dos métodos de estimación del vector de parámetros del MAB que son especialmente adecuados debido a las propiedades especiales de este modelo [63]. El primero de ellos es el Mínimo Cuadrado Condicional (MCC) propuesto por Lele y Ord (1986) [40] y el segundo es el Mínimo Cuadrado Condicional Coding (MCCC) propuesto por Martínez et al. (2018) [49].

Para la presentación de estos dos métodos de estimación, se considerará la característica local definida en (2.1).

2.2.1. Mínimo Cuadrado Condicional

El estimador Mínimo Cuadrado Condicional (MCC) se define como el argumento que minimiza la ecuación:

$$\hat{\theta} = \arg \min_{\theta} \sum_{s \in \Lambda_{CS}} (x_s - E[X_s])^2. \quad (2.6)$$

De la ecuación (2.5), la ecuación (2.6) puede reescribirse como:

$$\hat{\theta} = \arg \min_{\theta} \sum_{s \in \Lambda_{CS}} \left(x_s - \frac{M}{1 + e^{-\eta_s}} \right)^2. \quad (2.7)$$

De esta forma, el vector que minimiza la función de costo cuadrático de la ecuación (2.7) es el estimador condicional de mínimos cuadrados. Este enfoque intenta minimizar el error entre la imagen observada y la imagen esperada. Schröder et al. (1998) [63] propuso intercambiar esta ecuación con una expansión de Taylor reducida, que condujo a una estimación de un modelo lineal (ML), mediante el uso de la aproximación del campo medio, η_s puede ser aproximado por:

$$\eta_s \approx -\log \left(\frac{M}{x_s} - 1 \right) + \tilde{\epsilon}_s$$

donde $\tilde{\epsilon}_s$ es un término Gaussiano de ruido de media cero y varianza δ . Junto con la definición de η_s , ecuación (2.4), el proceso de estimación de parámetros se reduce a resolver un sistema de ecuaciones lineales:

$$G\theta = d + n, \quad (2.8)$$

donde θ es el vector de parámetros desconocidos, d el vector de valores transformados de los píxeles x_s

$$d_s = -\log \left(\frac{M}{x_s} - 1 \right), \quad (2.9)$$

y n un vector Gaussiano de ruido, G es una matriz de los valores de los píxeles vecinos,

$$G_{s,n} = \begin{cases} 1 & n = 0 \\ \frac{x_{s-v_i}^{(n)} + x_{s+v_i}^{(n)}}{M} & \text{caso contrario} \end{cases} \quad (2.10)$$

Los valores x_{s-v_i} y x_{s+v_i} denotan los dos vecinos del píxel x_s descriptos en el cliqué de pares (Ver Figura 1.5). La s -ésima fila especifica los niveles de grises de todos los vecinos del píxel s y la n -ésima columna contiene al vecino n de todos los píxeles de la imagen. El parámetro β_0

no pondera ningún vecino, por ende vale 1.

Considerando que el término de ruido n posee distribución normal, el estimador de máxima verosimilitud de θ esta dado por:

$$\hat{\theta}_{MCC} = (G^t G)^{-1} G^t d, \quad (2.11)$$

Una ventaja importante de este estimador es su cálculo que resulta significativamente rápido y sencillo, por este motivo, es ampliamente usado.

2.2.2. Mínimo Cuadrado Condicional Coding

Sea \mathcal{C} un familia no vacía de subconjuntos de S , la cual es ∂ -coding. Sea $C_1, C_2, \dots, C_n \in \mathcal{C}$, $n \geq 2$. El estimador Mínimo Cuadrado Condicional Coding (MCCC) se define como el argumento que minimiza la ecuación:

$$\hat{\theta}_i = \arg \min_{\theta} \sum_{s \in \Lambda \cap C_i \subset S} (x_s - E[X_s])^2. \quad (2.12)$$

Para cada C_i conjunto coding y considerando las ecuaciones (2.7) - (2.10) se obtiene

$$\hat{\theta}_i = (G_i^t G_i)^{-1} G_i^t d_i \quad (2.13)$$

Por último, $\hat{\theta}_{MCCC} = \frac{\sum_i^n \hat{\theta}_i}{n}$.

2.3. Análisis de los métodos de estimación

Con el objetivo de evaluar el rendimiento del MCCC propuesto y compararlo con el MCC, se abordará un estudio de simulación que considera cuatro aspectos: características descriptivas y texturales, sensibilidad de los estimadores en presencia y ausencia de ruido blanco gaussiano aditivo (AGWN), rendimiento de los algoritmos de cálculo paralelo y la capacidad de los estimadores en la clasificación supervisada en una imagen real. A continuación se detallan los diferentes aspectos:

2.3.1. Medidas de desempeño

En esta subsección, se presentan dos estadísticas que persiguen objetivos diferentes, la medida de Borges y el Error Cuadrático Medio empírico, para evaluar el desempeño del método de estimación propuesto.

Medida de Borges

Según Borges (1999) [5], lo que esencialmente impulsa el proceso de formación de textura son las probabilidades condicionales, y estas están relacionadas con los parámetros de textura mediante una transformación no lineal invertible pero compleja. La métrica Borges (ς) mide qué tan bien los parámetros de textura generados por los diferentes procedimientos de estimación reproducen las probabilidades condicionales correctas. Esta métrica se calcula como la distancia al cuadrado de las probabilidades condicionales verdaderas a las dadas por los parámetros de textura estimados. Para un MAB se define como,

$$\varsigma(\underline{\theta}, \hat{\underline{\theta}}) = \sum_{s \in \Lambda_{CS}} \left[\rho_{\Phi_{\underline{\theta}, \{x_s\}}}(x) - \rho_{\Phi_{\hat{\underline{\theta}}, \{x_s\}}}(x) \right]^2 \quad (2.14)$$

ECM empírico

El error cuadrático medio empírico (ECM) [69], Se define como:

$$\text{ECM}(\hat{\beta}_i) = \frac{\sum_{k=1}^N (\hat{\beta}_{i_k} - \beta_i)^2}{N},$$

donde $\hat{\beta}_i$, $i = 1, \dots, 4$, es el valor estimado del parámetro β_i y N representa el tamaño de una muestra de imágenes de textura.

2.3.2. Sensibilidad

Durante la adquisición, transmisión, almacenamiento y procesos de recuperación, la imagen se contamina con ruido. Esta distorsión hace que la interpretación y análisis de la escena sea muy difícil, por lo tanto, es esencial conocer el comportamiento de los métodos de estimación del vector de parámetros del MAB bajo diferentes patrones de contaminación en texturas [57].

Contaminación de texturas

Para estudiar la sensibilidad de los estimadores MCC y MCCC bajo contaminación aditiva, se consideró un proceso de contaminación [57]. En este caso el proceso X no es perfectamente observable debido a una pequeña fracción γ de observaciones generadas por el proceso de valores atípicos $\{\delta_{ij}, v_{ij}\}$, donde δ_{ij} es tal que $P(\delta_{ij} = 1) = \gamma$ y $P(\delta_{ij} = 0) = 1 - \gamma$, y las variables v_{ij} que tienen una distribución arbitraria R . El modelo observado es $Y_{ij} = X_{ij} + \delta_{ij}v_{ij}$. Por lo tanto, con probabilidad $(1 - \gamma)$ se observa el proceso X , y con probabilidad γ , las observaciones X_{ij} están corrompidas por un error de distribución R . En el estudio se consideran diferentes valores de γ : 0.05, 0.1, 0.15, 0.20, 0.25, 0.30, 0.35 y 0.40 y se asume que las v_{ij} están distribuidas

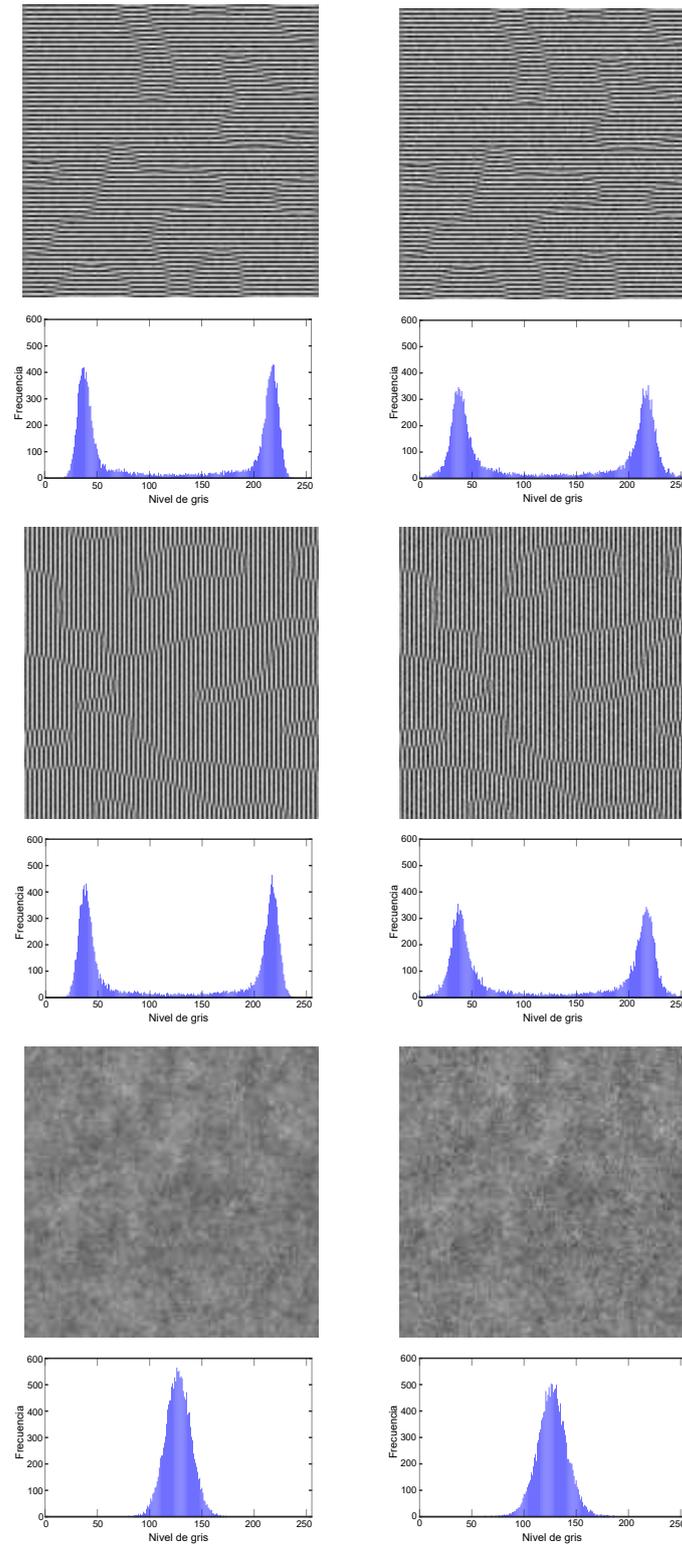


Figura 2.4: Textura original (sin contaminar) y contaminada con AWGN al 40% de los casos (i), (ii) y (iii) y su histograma, correspondiente al MAB de primer orden.

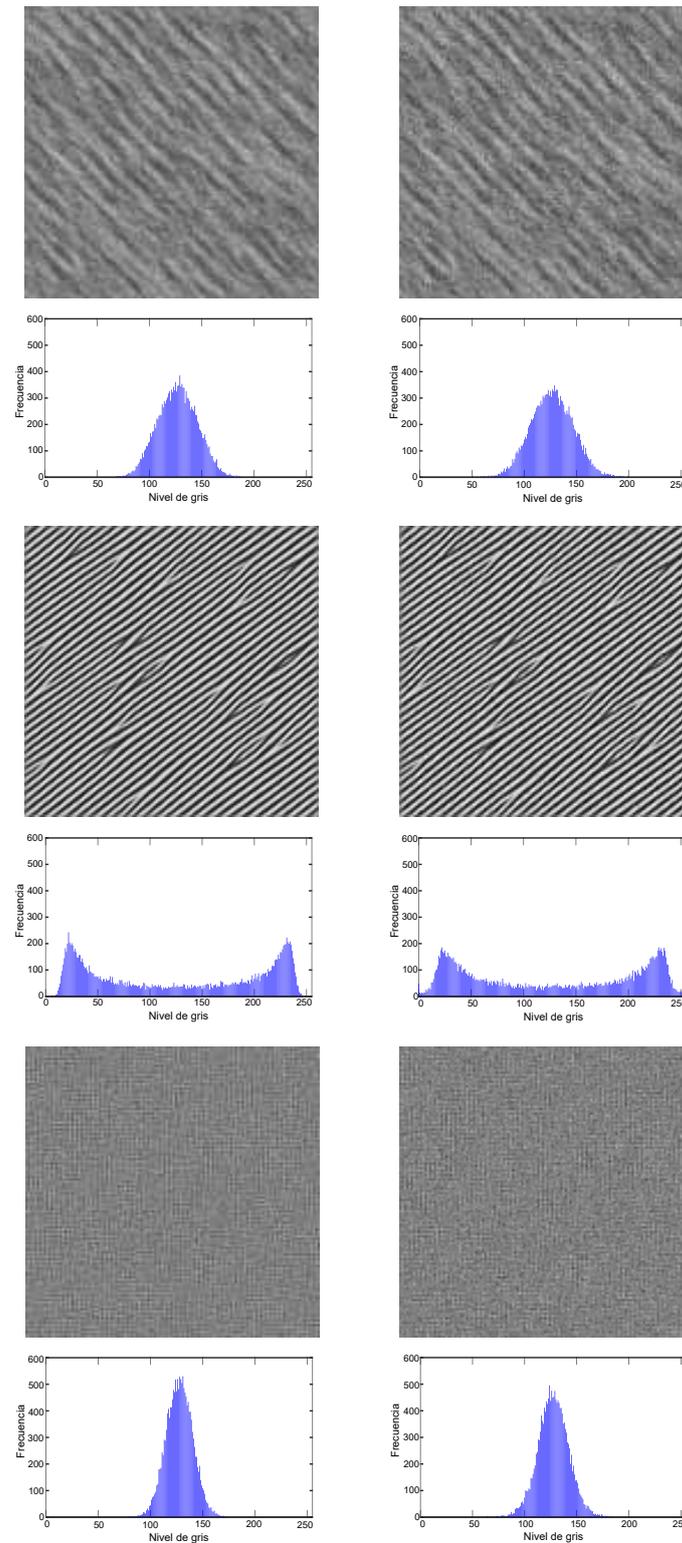


Figura 2.5: Textura original (sin contaminar) y contaminada con AWGN al 40% para los casos (i), (ii) y (iii) y su histograma, correspondiente al MAB de segundo orden.

normalmente con media 0 y varianza igual a 128, llamado ruido blanco gaussiano aditivo (en inglés, additive white gaussian noise - AWGN).

La elección del tipo de ruido y los niveles de contaminación considerados para este estudio, permiten simular distorsiones en las imágenes que se producen al realizar operaciones de adquisición, transmisión y procesamiento de las mismas, conservando el patrón de la textura. En las Figuras 2.4 y 2.5, se pueden visualizar la imagen original, sin contaminación ($\gamma = 0$) y la imagen con el máximo nivel de contaminación ($\gamma = 0.40$), junto con sus respectivos histogramas para cada una de las texturas correspondientes al MAB de primer y segundo orden, respectivamente. En todos los casos, se puede observar que el patrón de textura se mantiene, no así la distribución de frecuencia de los niveles de grises de los píxeles en las mismas.

Análisis de varianza de medidas repetidas (ANOVA MR)

Para cada MAB bajo estudio, se realizó una comparación de los estimadores utilizando el Análisis de Varianza de Medidas Repetidas (ANOVA MR) [39], considerando la norma euclídea como variable de respuesta, puesto que es empleada en diferentes aplicaciones reales del MAB [51, 52].

Para este análisis, se establecieron dos esquemas de trabajo:

- (a) sin contaminación, y
- (b) con contaminación (con diferentes niveles de AWGN).

El estudio de ANOVA MR permite responder las siguientes preguntas de investigación:

- Para el esquema (a): sin contaminación, ¿existe un método de estimación: MCC y MCCC, que tenga un rendimiento diferente en norma euclídea respecto del otro?.
- Para el esquema (b): con contaminación, ¿hay al menos un nivel de AWGN donde un método de estimación particular se comporta de manera diferente al otro?.

Se evaluaron los principales efectos sobre la norma Euclídea promedio, inducida por dos factores (intra e inter sujetos).

Para el esquema (a), se realizó ANOVA de dos vías con medidas repetidas en un factor:

Factor intra-sujetos:

- I) Métodos de estimación, con 2 niveles: MCC y MCCC.

Factor inter-sujetos:

- II) Casos de estudio, con 3 niveles: (i), (ii), y (iii).

Por lo tanto, el diseño del modelo multivariado utilizado es el siguiente:

$$y_{ijk} = \mu + \tau_j + \delta_{i(j)} + \gamma_k + (\tau\gamma)_{jk} + \epsilon_{ijk} \quad (2.15)$$

con $i = 1, \dots, 500$, $j = 1, 2$, y $k = 1, 2, 3$. donde:

- y_{ijk} denota el valor de la norma euclídea entre los parámetros originales con los que la imagen i generada, considerando el vector de parámetros del caso j , y el vector de parámetros estimado en la imagen i , utilizando el método k .
- μ es la media general.
- τ_j es el efecto principal del caso j , sujeto a $\sum_j \tau_j = 0$.
- $\delta_{i(j)}$ es el efecto aleatorio introducido por la imagen i dentro del caso j .
- γ_k es el efecto principal del método de estimación k , sujeto a $\sum_k \gamma_k = 0$.
- $(\tau\gamma)_{jk}$ es el efecto agregado de la interacción entre el caso j y método de estimación k , sujeto a $\sum_j (\tau\gamma)_{jk} = 0$ para todo k , y $\sum_k (\tau\gamma)_{jk} = 0$ para todo j .
- para todo i, j, k , con $i = 1, \dots, 500$; $j = 1, \dots, 5$, y $k = 1, 2$, ϵ_{ijk} y $\delta_{i(j)}$ son variables aleatorias independientes, con $\epsilon_{ijk} \sim N(0, \sigma_\epsilon^2)$ i.i.d. y $\delta_{i(j)} \sim N(0, \sigma_\delta^2)$ i.i.d..

Para el esquema **(b)**, se utiliza ANOVA de tres vías con medidas repetidas en dos factores,

Factor intra-sujetos:

- I) Métodos de estimación, con 2 niveles : MCC y MCCC.
- II) Contaminación (AWGN), con 9 niveles: 0, 0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35 y 0.40.

Factor inter-sujetos:

- III) Casos de estudio, con 3 niveles: (i), (ii) y (iii).

El diseño del modelo multivariante utilizado en este caso es:

$$y_{ijkt} = \mu + \tau_j + \delta_{i(j)} + \gamma_k + \alpha_t + (\tau\gamma)_{jk} + (\tau\alpha)_{jt} + (\gamma\alpha)_{kt} + (\tau\gamma\alpha)_{jkt} + \epsilon_{ijkt} \quad (2.16)$$

donde:

- y_{ijkt} denota el valor de la norma euclídea entre el vector de parámetros original de la imagen i del caso j , y el vector de parámetros estimado usando el método k , y la imagen original i contaminada con ruido aditivo en el nivel t
- α_t es el efecto principal del nivel t de contaminación, sujeto a $\sum_t \alpha_t = 0$.
- $(\tau\alpha)_{jt}$ es el efecto agregado de la interacción entre el caso j y el nivel de contaminación t , sujeto a $\sum_j (\tau\alpha)_{jt} = 0$ para todo t y $\sum_t (\tau\alpha)_{jt} = 0$ para todo j .
- $(\gamma\alpha)_{kt}$ es el efecto agregado de la interacción entre el método de estimación k y el nivel de contaminación t , sujeto a $\sum_k (\gamma\alpha)_{kt} = 0$ para todo t y $\sum_t (\gamma\alpha)_{kt} = 0$ para todo k ,
- $(\tau\gamma\alpha)_{jkt}$ es el efecto agregado de la interacción entre caso j , el método de estimación k y nivel de contaminación t , sujeto a $\sum_k (\tau\gamma\alpha)_{jkt} = 0$ para todo t , $\sum_t (\tau\gamma\alpha)_{jkt} = 0$ para todo k , y $\sum_t (\tau\gamma\alpha)_{jkt} = 0$ para todo k , y
- ϵ_{ijkt} el error experimental independiente. Se ha utilizado restricciones estándar de suma a cero en los parámetros.

Se asume que ϵ_{ijkt} y $\delta_{i(j)}$ son variables aleatorias independientes con $\epsilon_{ijkt} \sim N(0, \sigma_\epsilon^2)$ i.i.d. y $\delta_{i(j)} \sim N(0, \sigma_\delta^2)$ i.i.d., $i = 1, \dots, 500$; $j = 1, \dots, 5$; $k = 1, 2$ y $t = 1, \dots, 5$. Se ha omitido la definición de los parámetros definidos en el modelo anterior.

En los modelos multivariados 2.15 y 2.16, el efecto de imagen (denotado por $\delta_{i(j)}$) está anidado dentro del factor de caso. No se considera la interacción entre imágenes y tratamientos porque no resulta de interés en este estudio.

2.3.3. Evaluación de los algoritmos

Con respecto a la programación paralela, se determina la aceleración (en inglés Speed-up) de los algoritmos MCCC [38]. El Speed-up es el cociente entre el tiempo de ejecución para un procesador, preferiblemente utilizando el mejor algoritmo secuencial, y el tiempo de ejecución para p procesadores. En este estudio, se considera $p = 2$ (2 codings) y 4 (4 codings) para el MAB de primer y segundo orden, respectivamente y $p = 9$ (9 codings) para el tercer y cuarto orden. Además, se comparan los algoritmos del MCC y MCCC, mediante el cálculo del cociente entre los tiempos de ejecución del algoritmo MCC y MCCC. Para reducir el sesgo entre las diferentes ejecuciones, cada versión paralela se ejecuta diez veces y se reporta la media aritmética \pm desvío estándar en ambas situaciones consideradas.

2.3.4. Clasificación supervisada de una imagen real

Se analiza el comportamiento de los estimadores MCC y MCCC en la clasificación supervisada de texturas en el contexto de las imágenes de teledetección. Para ello, se supone que el MAB es un modelo válido. El proceso consiste en la estimación del vector de parámetros θ en un área de entrenamiento seleccionada para cada clase, y el vector resultante es empleado como un vector de caracterización de textura en un espacio de tres (primer orden) y cinco dimensiones (segundo orden). El mapa de clasificación se obtuvo etiquetando un píxel estimando el vector de parámetros en una ventana centrada en el píxel de la imagen. Luego, al etiquetar el parche con el modelo con los parámetros más cercanos ($\Delta\theta$ hasta el umbral T), utilizando la norma euclídea. Para esta aplicación se consideraron un tamaño de ventana: 17×17 píxeles, y umbral $T = k/10$, con $k = 1, \dots, 40$, según Martínez et. al (2015) [47]. El rendimiento de la clasificación de imágenes se evaluó con la Fiabilidad Global (FG) y las medidas estadísticas de Kappa. Para el estudio, utilizamos la imagen Pléiades-1A (base de datos de Calanques), fechada el 17 de abril de 2013 [20]. La escena seleccionada en el experimento fue un viñedo ($43^\circ 15' 45,9''$ N y $5^\circ 37' 55,4''$ E), con 1000×1000 píxeles. Ver Figura 2.6.

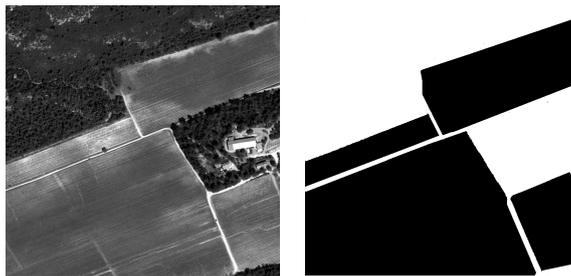


Figura 2.6: Imagen satelital Pléiades-1A de alta resolución: Viñedo (1000×1000 píxeles) y su mapa de clases donde negro es viñedo y blanco es no viñedo.

Medidas de evaluación de la clasificación

Para evaluar la calidad del proceso de clasificación de una imagen se utilizan medidas que se definen a partir de la matriz de confusión. Esta matriz identifica las diferencias entre el mapa de clases original M_O y el reconstruido, M_R . Se trata de una matriz cuadrada de orden k (cantidad de clases en consideración). Esta sección presenta la definición de dos medidas ampliamente adoptadas: la fiabilidad global y el coeficiente Kappa.

Sea M la matriz de confusión estándar generada por dos mapas de clases: M_O y el M_R . Usualmente, las filas de M , son ocupadas por la categorización de los píxeles deducida por el clasificador, mientras que las columnas corresponden a las clases de referencia establecidas por observaciones del terreno. El valor n_{ij} correspondiente a la matriz de entrada, $M(i, j)$, representa el número de píxeles de la verdadera clase j que es asignado a la clase i en la imagen

reconstruida. La diagonal de la matriz M expresa el número de píxeles clasificados donde se produce un total acuerdo entre el mapa de referencia (original) y el reconstruido, mientras que los marginales suponen errores de asignación. Los totales columna y fila se denotarán por (n_{+j}) y (n_{i+}) , respectivamente, n será el total de píxeles considerados y $p_{ij} = n_{ij} = n$; $p_{i+} = n_{i+} = n$; $p_{+i} = n_{+j} = n$ las proporciones de píxeles mal clasificados.

A partir de una matriz de confusión pueden deducirse índices relativos a la exactitud de la clasificación.

Fiabilidad Global

Este coeficiente introducido por Anderson et al. (1976) [3] calcula la proporción de coincidencias entre la imagen de referencia y la reconstruida (píxeles correctamente clasificados).

$$FG = \frac{\sum_{i=1}^k n_{ii}}{n} \cdot 100 \%, \quad (2.17)$$

Habitualmente, se expresa en porcentaje. Un esquema de clasificación se considera bueno si su fiabilidad global esta por encima del 85 %.

Coficiente Kappa

El estimador Kappa (κ) fue desarrollado originalmente por Cohen (1960) [15] y mide el grado de coincidencia en un esquema de clasificación, ayuda a determinar si la salida del esquema es mejor que la asignación aleatoria.

$$\kappa = \frac{p_o - p_c}{1 - p_c}.$$

Este coeficiente se basa en la coincidencia real $p_o = \sum_{i=1}^k p_{ii}$ y la posible coincidencia esperada por puro azar $p_c = \sum_{i=1}^k p_{i+}p_{i+}$. Resume la bondad de la clasificación en un sólo valor. Si el κ es igual a 1 esto indica un acuerdo pleno, mientras que un valor de 0 sugiere que el acuerdo observado es puramente debido al azar. En caso en que el acuerdo observado sea inferior al esperado por azar, el estadístico Kappa es negativo. Para más detalles ver Agresti (1996) [1].

2.4. Resultados experimentales

Mediante el algoritmo Gibbs Sampler (GS) (Anéxo A), se generaron 500 imágenes de textura de 128×128 píxeles de tamaño, con 256 niveles de gris, para los MAB en estudio, considerando tres casos con diferentes patrones de textura que fueron presentados en la Sección 2.1 en las Figuras 2.2 y 2.3.

2.4.1. Análisis descriptivo

Según el estudio descriptivo, en todos los casos de los dos órdenes MAB bajo estudio, el comportamiento de ambos estimadores es comparable de acuerdo a los resultados presentados en las Tablas 2.3 y 2.4 y como se puede apreciar en las Figuras 2.7 y 2.9.

Además se determinaron las estimaciones normalizadas para analizar la magnitud relativa de los sesgos respecto del valor verdadero del parámetro. En las figuras 2.8 y 2.10 se muestran las distribuciones de las estimaciones normalizadas (EN) donde se puede apreciar que las diferentes dispersiones que tienen las estimaciones para los distintos valores de los parámetros son semejantes.

Estos resultados también se corrobora cuando analizamos el desempeño de los métodos MCC y MCCC de acuerdo a el ECM. Ver Tablas 2.3 y 2.4 y gráfico de barras de la Figura 2.11.

Por otro lado, la medida de Borges, confirma la capacidad de MCC y MCCC para estimar los parámetros de textura para reproducir las probabilidades condicionales correctas, de los MAB de primer y segundo orden. Ver gráfico de barras de error de la Figura 2.11.

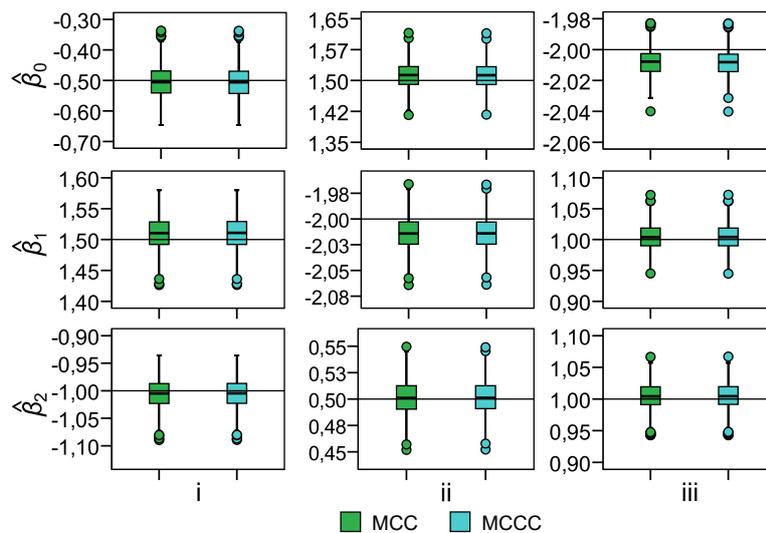


Figura 2.7: Los boxplots de cada columna representan la distribución de las estimaciones de los parámetros de los MAB de primer orden, para los casos (i), (ii) y (iii) respectivamente. La línea horizontal indica el valor del parámetro especificado.

Tabla 2.3: Resultados experimentales: media, desviación estándar (DE), coeficiente de variación (CV%) y ECM de las estimaciones MCC, MCCC en MAB de primer orden.

Parámetros	Caso (i)			
	β_0	β_1	β_2	
Valor verdadero	-0.5	1.5	-1	
Método de Estimación				
MCC	Media	-0.50418	1.51097	-1.00555
	DE	0.05249	0.02618	0.02630
	CV %	10.41133	1.73407	2.61602
	ECM	0.00276	0.00077	0.00072
MCCC	Media	-0.50485	1.51011	-1.0052
	DE	0.05252	0.02619	0.02632
	CV %	10.40368	1.73472	2.61837
	ECM	0.00277	0.00078	0.00071
Caso (ii)				
Valor verdadero	1.5	-2	0.5	
MCC	Media	1.51179	-2.01357	0.50176
	DE	0.03012	0.01484	0.01515
	CV %	1.99247	0.73704	3.02026
	ECM	0.00104	0.00040	0.00023
MCCC	Media	1.51172	-2.01353	0.50180
	DE	0.03008	0.01482	0.01513
	CV %	1.99031	0.73625	3.01670
	ECM	0.00104	0.00040	0.00023
Caso (iii)				
Valor verdadero	-2	1	1	
MCC	Media	-2.00808	1.00418	1.00390
	DE	0.00859	0.02143	0.02139
	CV %	0.42817	2.13433	2.13092
	ECM	0.00013	0.00047	0.00047
MCCC	Media	-2.00834	1.00430	1.00403
	DE	0.00859	0.02144	0.02140
	CV %	0.42813	2.13511	2.13155
	ECM	0.00014	0.00047	0.00047

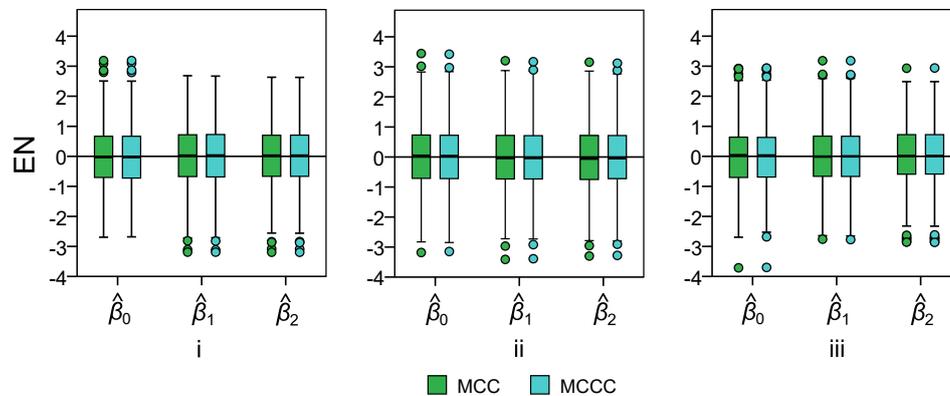


Figura 2.8: Los boxplots de cada columna representan la distribución de las estimaciones normalizadas (EN) de los parámetros de los MAB de primer orden, para los casos (i), (ii) y (iii), respectivamente.

Tabla 2.4: Resultados experimentales: media, desviación estándar (DE), coeficiente de variación (CV%) y ECM de las estimaciones MCC, MCCC en MAB de segundo orden.

Parametro		Caso (i)				
		β_0	β_1	β_2	β_3	β_4
Valor verdadero		-2	1	1	-0.6	-0.6
Método de estimación						
MCC	Media	-2.00677	1.00393	1.00426	-0.60265	0.60123
	DE	0.00734	0.02619	0.02784	0.01883	0.02667
	CV %	0.36585	2.60915	2.77235	3.12460	4.43720
	ECM	0.00009	0.00070	0.00079	0.00036	0.00071
MCCC	Media	-2.00683	1.00397	1.00441	-0.60276	0.60121
	DE	0.00738	0.02622	0.02784	0.01888	0.02670
	CV %	0.36796	2.61258	2.7726	3.13326	4.44172
	ECM	0.00010	0.00070	0.00079	0.00036	0.00071
Valor verdadero		Caso (ii)				
		-2	1	-0.5	2	-0.5
MCC	Media	-2.01952	1.02500	-0.50826	2.00794	-0.50515
	DE	0.02281	0.01364	0.02253	0.01584	0.01664
	CV %	1.12948	1.33143	4.43373	0.78900	3.29427
	ECM	0.00090	0.00081	0.00057	0.00031	0.00030
MCCC	Media	-2.01988	1.02497	-0.50835	2.00818	-0.50492
	DE	0.02278	0.01366	0.02252	0.01582	0.01662
	CV %	1.12803	1.33362	4.43004	0.78786	3.29201
	ECM	0.00091	0.00081	0.00057	0.00031	0.00030
Valor verdadero		Caso (iii)				
		0	1	1	-1	-1
MCC	Media	0.00233	1.00241	1.00288	-1.00407	-1.00360
	DE	0.04807	0.02475	0.02504	0.01930	0.01892
	CV %	2059.08778	2.46918	2.49738	1.92270	1.88523
	ECM	0.00231	0.00061	0.00063	0.00038	0.00037
MCCC	Media	0.00153	1.00293	1.00339	-1.00416	-1.00373
	DE	0.04810	0.02479	0.02508	0.01934	0.01892
	CV %	3142.26182	2.4719	2.50049	1.92693	1.88506
	ECM	0.00231	0.00062	0.00063	0.00039	0.00037

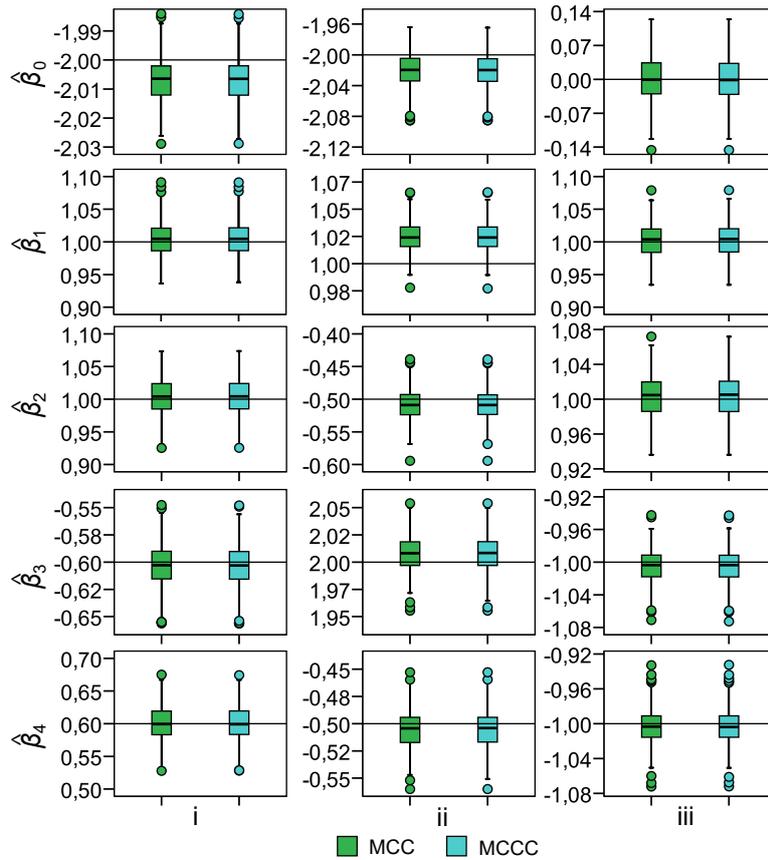


Figura 2.9: Los boxplots de cada columna representan las estimaciones de los parámetros de los modelos de segundo orden, para los casos (i) – (iii) respectivamente. La línea horizontal indica el valor del parámetro especificado.

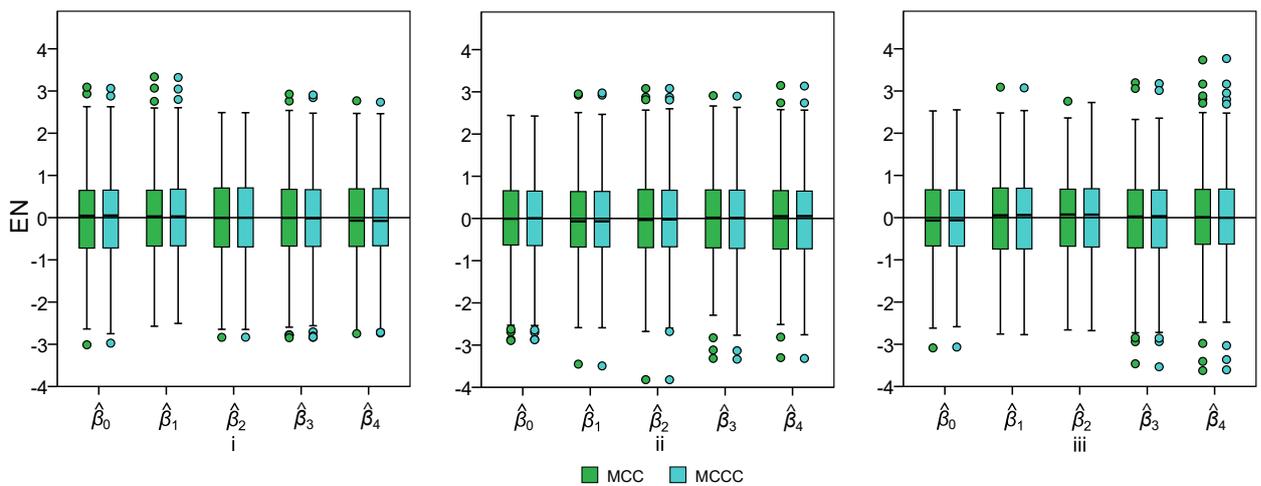


Figura 2.10: Los boxplots de cada columna representan la distribución de las estimaciones normalizadas (EN) de los parámetros de los MAB de segundo orden, para los casos (i), (ii) y (iii), respectivamente.

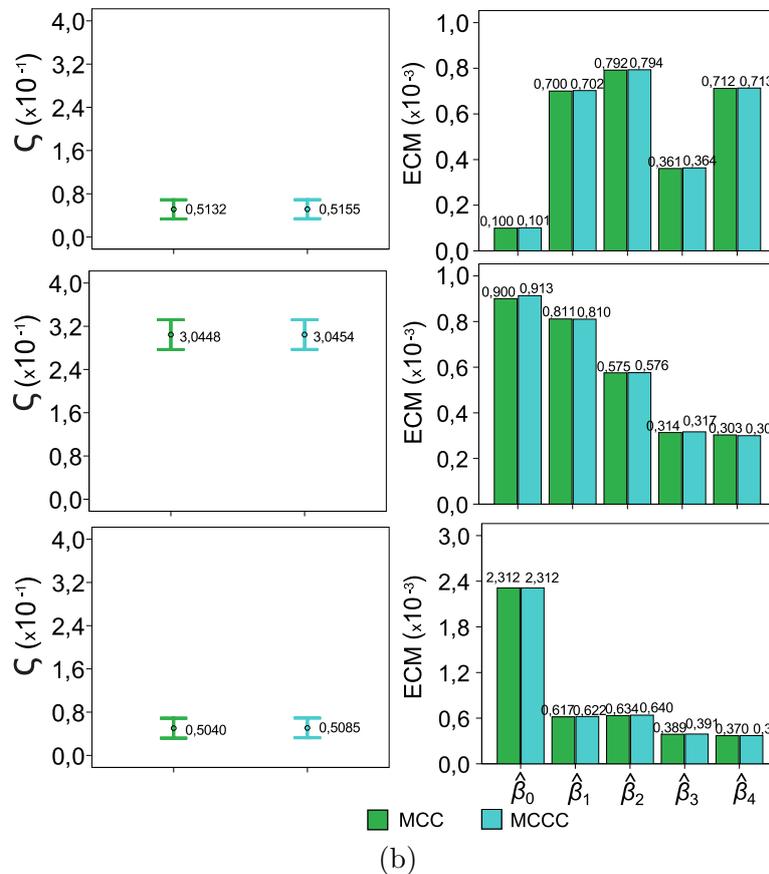
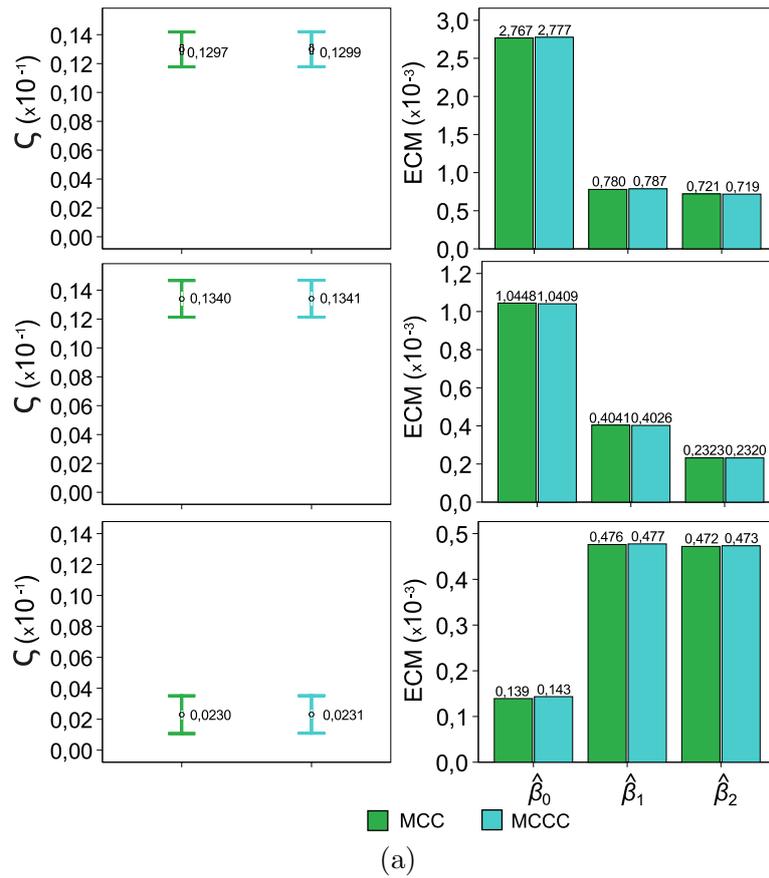


Figura 2.11: Primera columna, las barras de error (Media \pm DE) de la medida de Borges y la segunda columna, la gráfica de barras del rendimiento de ECM de ambos métodos de estimación, para los casos (i), (ii) y (iii) considerados en el (a) MAB de primer orden y (b) MAB de segundo orden, respectivamente.

2.4.2. Resultados de sensibilidad

A continuación se presentan las respuestas obtenidas mediante el estudio de ANOVA MR a los interrogantes planteados al finalizar la Subsección 2.3.2, que son:

- Para el esquema (a): sin contaminación, ¿existe un método de estimación: MCC y MCCC, que tenga un rendimiento diferente en norma euclídea respecto del otro?
- Para el esquema (b): con contaminación, ¿hay al menos un nivel de AWGN donde un método de estimación particular se comporta de manera diferente al otro?

Esquema (a) ANOVA MR en texturas sin contaminar

Las tablas 2.5 y 2.6 resumen los resultados de ANOVA MR, donde la variable respuesta es la norma euclídea, considerando un MAB de primer y segundo orden, respectivamente. Las mismas no incluyen el efecto del factor inter-sujetos Casos de estudio, debido a que en ellas se exhiben únicamente los estadísticos univariados que permiten contrastar las hipótesis nulas referidas a los efectos en los que se encuentra involucrado el factor intra-sujetos, en este caso el factor Método de Est. y la interacción Método de Est. * Casos.

Tabla 2.5: MAB de primer orden: ANOVA dos vías con MR en un factor: Test de Efectos Intra-sujetos.

F. V.	S.C.	g.l.	C.M.	F	p
Método Est.	3.72E-06	1	3.72E-06	11.299	0.001
Método Est. *Caso	5.23E-06	2	2.62E-06	7.956	<0.0001
Error(Método Est.)	0.00E-07	1497	3.29E-07		

Tabla 2.6: MAB de segundo orden: ANOVA de dos vías con MR en un factor: Test de Efectos Intra-sujetos.

F. V.	S.C.	g.l.	C.M.	F	p
Método Est.	1.21E-05	1	1.21E-05	13.074	<0.0001
Método Est. *Caso	1.20E-06	2	5.99E-07	0.649	0.523
Error(Método Est.)	0.001	1497	9.22E-07		

El enfoque univariado se utilizó en ambos MAB, el supuesto de esfericidad no se puede verificar debido a que el factor intra-sujetos posee dos niveles. El efecto principal (métodos de estimación) fue estadísticamente significativo ($p < 0.0001$) en ambos modelos. La interacción entre el método de estimación y los casos bajo estudio resultó significativa únicamente en el MAB de primer orden. Por tal motivo, se realizó un estudio por caso para dicho modelo, presentado en la siguiente subsección.

Pruebas individuales de ANOVA MR para el MAB de primer orden

Se efectuaron ANOVA de MR en un factor para cada uno de los casos considerados. El objetivo es analizar si los métodos de estimación presentan un comportamiento en norma euclídea promedio diferente en cada uno de los casos del MAB de primer orden.

La tabla 2.7 muestra el resultado de la prueba de ANOVA MR para la norma euclídea como variable de respuesta, para estimadores MCC y MCCC, considerando un MAB de primer para los casos (i), (ii) y (iii), respectivamente. Se utilizó la corrección de Greenhouse-Geisser [26] ($p < 0.0001$), porque no se cumplió la condición esfericidad. Además, la corrección de Huynh-Feldt [32] (aquí omitida) también se calculó y se obtuvo la misma conclusión. Luego, comparando el método MCC y MCCC, los puntajes de la norma euclídea entre los parámetros originales y los estimados en cada uno de los casos de textura (i), (ii) o (iii), podemos afirmar que son estadísticamente diferentes para los casos (i) y (ii) con $p = 0.022$ y $p < 0.01$ respectivamente, mientras que para el caso (iii) no se encontraron evidencias suficientes que aseguren que los estimadores son diferentes en términos de la norma euclídea promedio ($p = 0.123$).

Tabla 2.7: MAB de primer orden: ANOVA de una vía con MR en un factor: Test de Efectos Intra-sujetos (corrección de Greenhouse-Geisser).

Caso (i)						
F. V.	S.C.	g.l.	C.M.	F	<i>p</i>	
Método Est.	3.76E-06	1	3.76E-06	5.263	0.022	
Error(Método Est.)	0.000	499	7.15E-07			
Caso (ii)						
F. V.	S.C.	g.l.	C.M.	F	<i>p</i>	
Método Est.	4.61E-06	1	4.61E-06	134.015	0.000	
Error(Método Est.)	1.72E-05	499	3.44E-08			
Caso (iii)						
F. V.	S.C.	g.l.	C.M.	F	<i>p</i>	
Método Est.	5.63E-07	1	5.63E-07	2.383	0.123	
Error(Método Est.)	0.000	499	2.36E-07			

A partir de los resultados obtenidos, los estimadores MCC y MCCC, manifiestan un comportamiento distinto en norma euclídea en los diferentes tipos de texturas sin contaminar consideradas, a excepción de las texturas del ABM de primer orden que no poseen un patrón definido (caso (iii)).

Esquema (b) ANOVA MR en texturas contaminadas

Las tablas 2.8 y 2.9 muestran la salida de las pruebas ANOVA de tres vías MR en dos factores, asumiendo la norma euclídea como variable de respuesta. La condición de esfericidad no fue satisfecha; en su lugar se utilizó la corrección Greenhouse-Geiser [26]. Además, la corrección

Huynh - Feldt [32] (aquí omitida) también se calculó y se obtuvo la misma conclusión. La interacción entre los métodos de estimación y la contaminación AWGN fue estadísticamente significativa. El efecto del factor caso también fue significativo; a saber, sin tener en cuenta los métodos de estimación y la contaminación AWGN, los puntajes de la norma euclídea varían entre los casos (i), (ii) y (iii).

Tabla 2.8: MAB de primer orden: ANOVA de tres vías MR en dos factores. Test de Efectos Intra-sujetos (corrección de Greenhouse-Geisser)

F. V.	S.C.	g.l.	C.M.	F	<i>p</i>
Método Est.	0.001	1	0.001	210.473	< 0.0001
Método Est.*Caso	0.001	2	0.001	188.103	< 0.0001
Error(Método Est.)	0.005	1497	3.24E-06		
Contaminación	405.181	2.862	141.582	18617.980	< 0.0001
Contaminación*Caso	56.916	5.724	9.944	1307.644	< 0.0001
Error(Contaminación)	32.579	4284.130	0.008		
Método Est.*Contaminación	0.000	3.327	4.68E-05	71.089	< 0.0001
Método Est.*Contaminación*Caso	0.000	6.654	4.46E-05	67.797	< 0.0001
Error(Método Est.*Contaminación)	0.003	4980.161	6.58E-07		

Exploramos la interacción entre el método de estimación y el nivel de contaminación por AWGN, utilizando las gráficas de perfil de la norma euclídea promedio. Las gráficas de los dos métodos de estimación son paralelas, excepto en texturas sin contaminación de los tres casos bajo estudio cuando se considera el MAB de primer orden (ver Figura 2.12 - primera columna).

Tabla 2.9: MAB de segundo orden: ANOVA de tres vías MR en dos factores. Test de Efectos Intra-sujetos (corrección de Greenhouse-Geisser).

F. V.	S.C.	g.l.	C.M.	F	<i>p</i>
Método Est.	0.005	1	0.005	433.811	< 0,0001
Método Est.*Caso	0.004	2	0.002	168.717	< 0,0001
Error(Método Est.)	0.018	1497	1.23E-05		
Contaminación	1059.230	3.244	326.472	64624.566	< 0,0001
Contaminación*Caso	154.850	6.489	23.864	4723.782	< 0,0001
Error(Contaminación)	24.537	4856.985	0.005		
Método Est.*Contaminación	0.001	3.196	3.13E-04	143.458	< 0,0001
Método Est.*Contaminación*Caso	0.001	6.392	1.56E-04	52.473	< 0,0001
Error(Método Est.*Contaminación)	0.011	4784.576	2.19E-06		

Es importante tener en cuenta que los valores más pequeños de la norma euclídea indican un mejor ajuste del vector de parámetros estimado al vector de parámetros real. En todos los casos de estudio, el valor medio de la norma euclídea se incrementa, mientras que los niveles de contaminación AWGN aumentan en las texturas. Este patrón es común a los dos métodos de estimación.

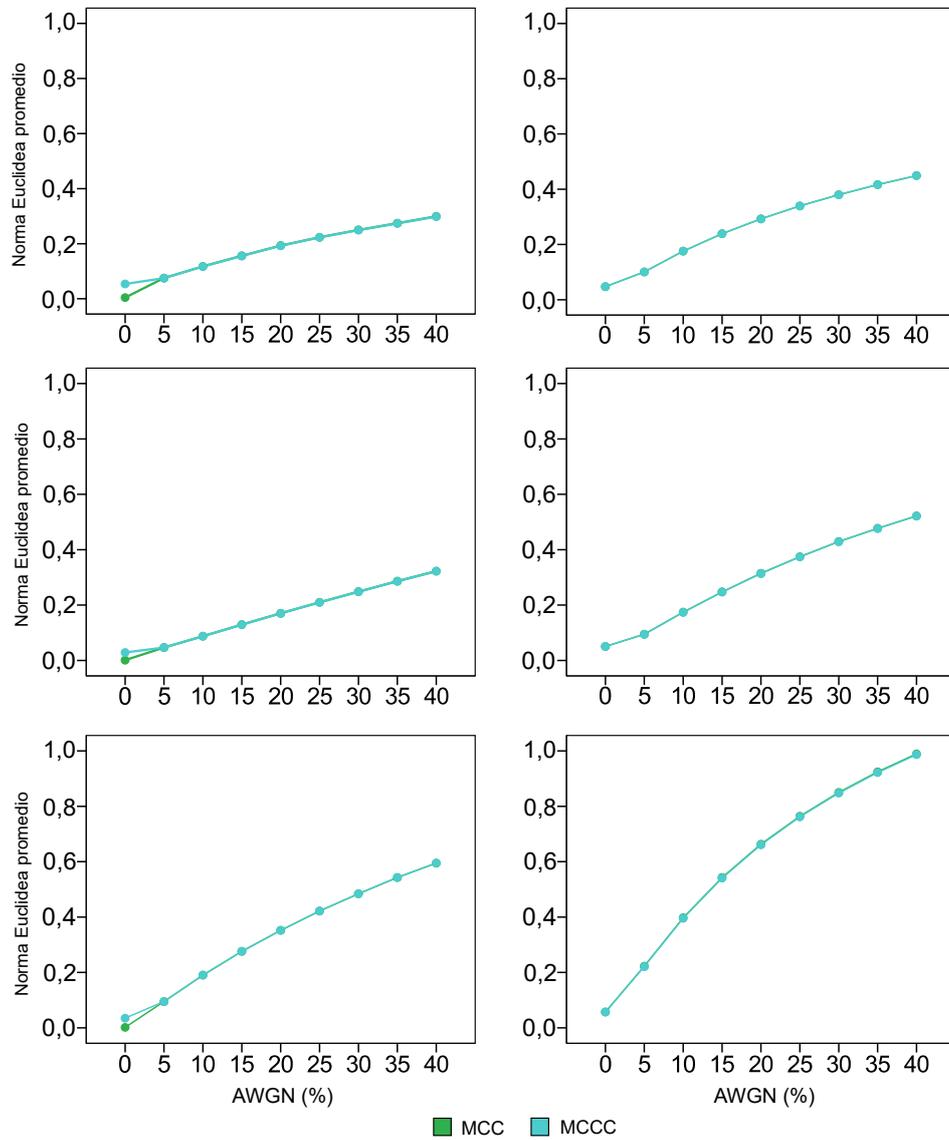


Figura 2.12: Desempeño de la norma Euclídea calculada entre los parámetros originales y los estimados por los métodos: MCC y MCCC, a diferentes niveles de contaminación AWGN para las texturas de los casos: (i), (ii) y (iii), consideradas en el MAB de primer orden (primera columna) y de segundo orden (segunda columna), respectivamente.

Debido a que la interacción de ambos factores podría enmascarar sus efectos sobre los valores de la norma euclídea entonces, para cada método de estimación, se realizó un estudio del desempeño de la norma euclídea. Particularmente, se investigó cuál de los métodos está más afectado por la contaminación AWGN.

Pruebas individuales de ANOVA MR

Para cada uno de los métodos de estimación, y caso de estudio, se realizaron pruebas de ANOVA MR individuales. El objetivo es analizar el efecto de la contaminación sobre la norma euclídea promedio en cada situación.

Las tablas 2.10 y 2.12 muestran el resultado de la prueba de ANOVA MR para la norma euclídea como variable de respuesta, para estimadores MCC y MCCC, considerando un MAB de primer y segundo orden, respectivamente. Para las dos pruebas, se utilizó la corrección de Greenhouse-Geisser [26] ($p < 0.0001$), porque no se cumplió la condición esfericidad. Además, la corrección de Huynh-Feldt [32](aquí omitida) también se calculó y se obtuvo la misma conclusión. Luego, usando el método MCC o MCCC, los puntajes de la norma euclídea entre los parámetros originales y los estimados que provienen de los distintos casos de textura (*i*), (*ii*) o (*iii*), son significativamente diferentes, a diferentes niveles de contaminación AWGN. Además, las interacciones fueron estadísticamente significativas, es decir, los puntajes de la norma euclídea para diferentes niveles de contaminación aditiva varían entre los casos de estudio.

Los resultados de las comparaciones múltiples con el ajuste de Bonferroni se pueden ver en las Tablas 2.11 y 2.13, para ambos MAB. En todos los casos bajo estudio, puede observarse que la norma euclídea es sensible a los cambios en los niveles de contaminación aditiva sobre las texturas, sin importar el método de estimación utilizado. La norma euclídea de cada método de estimación experimenta un deterioro progresivo a medida que aumenta el nivel de AWGN. Este comportamiento se manifiesta en todos los casos de los MAB de primer y segundo orden. Solo cabe destacar que para ambos MAB, en las texturas correspondientes al tercer caso, los scores de la norma euclídea de cada uno de los métodos de estimación, exhiben un incremento mayor ante los niveles de contaminación aditiva comparado con los otros dos tipos de texturas.

Tabla 2.10: MAB de primer orden: ANOVA de dos vías MR en un sólo factor: Test de Efectos Intra-sujetos (corrección de Greenhouse-Geisser).

MCC					
F. V.	S.C.	g.l.	C.M.	F	<i>p</i>
Contaminación	202.801	2.862	70.867	18646.450	< 0,0001
Contaminación*Caso	28.342	5.723	4.952	1302.927	< 0,0001
Error(Contaminación)	16.282	4283.969	0.004		
MCCC					
F. V.	S.C.	g.l.	C.M.	F	<i>p</i>
Contaminación	202.381	2.862	70.713	18585.816	< 0,0001
Contaminación*Caso	28.575	5.724	4.992	1312.106	< 0,0001
Error(Contaminación)	16.301	4284.435	0.004		

Tabla 2.11: MAB de primer orden: Comparaciones múltiples. Diferentes letras en cada caso indican diferencias entre los puntajes de la norma euclídea con respecto a la contaminación por AWGN (prueba de ajuste de Bonferroni $p < 0,05$). Para todos los casos, $N = 500$.

Método de Estimación	AWGN (%)	Casos		
		(i)	(ii)	(iii)
MCC	0	0.05367 ^a	0.02863 ^a	0.03497 ^a
	5	0.07517 ^b	0.04685 ^b	0.09477 ^b
	10	0.11758 ^c	0.08735 ^c	0.19025 ^c
	15	0.15612 ^d	0.12965 ^d	0.27614 ^d
	20	0.19333 ^e	0.17038 ^e	0.35198 ^e
	25	0.22352 ^f	0.21010 ^f	0.42161 ^f
	30	0.25023 ^g	0.24855 ^g	0.48408 ^g
	35	0.27457 ^h	0.28627 ^h	0.54266 ^h
	40	0.29931 ⁱ	0.32269 ⁱ	0.59456 ⁱ
MCCC	0	0.05379 ^a	0.02876 ^a	0.03492 ^a
	5	0.07468 ^b	0.04657 ^b	0.09489 ^b
	10	0.11683 ^c	0.08735 ^c	0.19044 ^c
	15	0.15527 ^d	0.12923 ^d	0.27637 ^d
	20	0.19241 ^e	0.16993 ^e	0.35219 ^e
	25	0.22255 ^f	0.20962 ^f	0.42192 ^f
	30	0.24918 ^g	0.24801 ^g	0.48444 ^g
	35	0.27346 ^h	0.28570 ^h	0.54304 ^h
	40	0.29821 ⁱ	0.32207 ⁱ	0.59495 ⁱ

Tabla 2.12: MAB de segundo orden: ANOVA de dos vías MR en un sólo factor: Test de Efectos Intra-sujetos .

MCC					
F. V.	S.C.	g.l.	C.M.	F	<i>p</i>
Contaminación	530.493	3.241	163.678	64718.835	< 0,0001
Contaminación*Caso	77.731	6.482	11.992	4741.519	< 0,0001
Error(Contaminación)	12.271	4851.899	0.003		
MCCC					
F. V.	S.C.	g.l.	C.M.	F	<i>p</i>
Contaminación	528.738	3.248	162.798	64475.186	< 0,0001
Contaminación*Caso	77.120	6.496	11.873	4702.056	< 0,0001
Error(Contaminación)	12.276	4861.980	0.003		

El análisis descriptivo y de sensibilidad de los estimadores presentado en las secciones 2.4.1

Tabla 2.13: MAB de segundo orden: Comparaciones múltiples. Diferentes letras en cada caso indican diferencias entre los puntajes de la norma euclídea con respecto a la contaminación por AWGN (prueba de ajuste de Bonferroni $p < 0,05$). Para todos los casos, $N = 500$.

Método de Estimación	AWGN (%)	Casos		
		(i)	(ii)	(iii)
MCC	0	0.04719 ^a	0.05053 ^a	0.05717 ^a
	5	0.10075 ^b	0.09479 ^b	0.22271 ^b
	10	0.17584 ^c	0.17452 ^c	0.39754 ^c
	15	0.23999 ^d	0.24773 ^d	0.54280 ^d
	20	0.29263 ^e	0.31477 ^e	0.66306 ^e
	25	0.33949 ^f	0.37476 ^f	0.76398 ^f
	30	0.38014 ^g	0.42957 ^g	0.85005 ^g
	35	0.41657 ^h	0.47720 ^h	0.92427 ^h
	40	0.44928 ⁱ	0.52192 ⁱ	0.98910 ⁱ
MCCC	0	0.04730 ^a	0.05062 ^a	0.05735 ^a
	5	0.10064 ^b	0.09431 ^b	0.22115 ^b
	10	0.17574 ^c	0.17388 ^c	0.39658 ^c
	15	0.23890 ^d	0.24695 ^d	0.54086 ^d
	20	0.29256 ^e	0.31386 ^e	0.66090 ^e
	25	0.33946 ^f	0.37390 ^f	0.76164 ^f
	30	0.38005 ^g	0.42872 ^g	0.84760 ^g
	35	0.41653 ^h	0.47638 ^h	0.92167 ^h
	40	0.44925 ⁱ	0.52109 ⁱ	0.98639 ⁱ

y 2.4.2, respectivamente, se llevó a cabo mediante la utilización del software estadístico SPSS 15.0 versión 15.0.1 .

2.4.3. Evaluación de los algoritmos

Dada una imagen de entrada de tamaño 128×128 píxeles, el algoritmo del MCC, según la ecuación 2.10, computa una única matriz G de dimensión 15876×3 y 15876×5 , para el MAB de primer y segundo orden, respectivamente. Luego, mediante la ecuación 2.11 determina el valor del estimador. Mientras que, el algoritmo del estimador MCCC, en forma paralela, construye dos matrices G de dimensiones 7936×3 para el MAB de primer orden y cuatro matrices de dimensiones 3969×5 para el MAB de segundo orden, una por cada coding. Posteriormente aplica la ecuación 2.13 y finalmente determina el valor del estimador MCCC, promediando los valores de los $\hat{\theta}_i$ ($i = 1$ y 2 - primer orden $i = 1, 2, 3$ y 4 - segundo orden).

A continuación se presentan en las Tablas 2.14 y 2.15 los resultados del cálculo del Speed up y del cociente entre el tiempo de ejecución del MCC y MCCC para cada MAB de distinto orden.

Tabla 2.14: Speed-up del MCCC para el MAB del primer y segundo orden: media \pm desvío estándar.

Casos	Orden 1	Orden 2
(i)	1.0763 \pm 0.0916	1.4053 \pm 0.1520
(ii)	1.0738 \pm 0.0689	1.4180 \pm 0.1478
(iii)	1.0799 \pm 0.0798	1.4281 \pm 0.1478

Tabla 2.15: Resultados del cálculo del cociente entre el tiempo de ejecución del MCC y MCCC para el MAB del primer y segundo orden: media \pm desvío estándar.

Casos	Orden 1	Orden 2
(i)	1.8999 \pm 0.1355	4.7566 \pm 0.7445
(ii)	1.9021 \pm 0.1381	4.7813 \pm 0.6079
(iii)	1.9159 \pm 0.1348	4.7399 \pm 0.6413

Cabe mencionar, que en los algoritmos paralelos la sincronización de los procesos tiene un costo computacional y en este caso se puede apreciar en el speed up del MCCC en el MAB de primer orden (ver Tabla 2.14).

Resultados de la clasificación

Con respecto a la aplicación real, los resultados de las medidas consideradas se reportan en la tabla 2.16. Los valores de Kappa revelan un nivel de acuerdo “moderado” cuando se utiliza el primer orden, y un nivel “sustancial” en el modelo de segundo orden, entre la imagen de verdad básica y los mapas de clase obtenidos por MCC y MCCC.

Tabla 2.16: Resultados de la clasificación: Imagen satelital Pléiades-1A de alta resolución: Viñado.

Orden	Método de Clasificación	T	FG (%)	$\hat{\kappa}$	IC de κ (95%)
Primer	MCC	1.5	80.62	0.5862	(0.5845,0.5878)
	MCCC		80.60	0.5858	(0.5841, 0.5874)
Segundo	MCC	1.6	87.78	0.7459	(0.7445,0.7472)
	MCCC		87.96	0.7499	(0.7486,0.7513)

El comportamiento de los métodos clasificadores fue similar.

Para realizar una comparación visual ver las Figuras 2.4.3 y 2.4.3, donde se presentan los mapas de clasificación obtenidos para MAB de primer y segundo orden, respectivamente

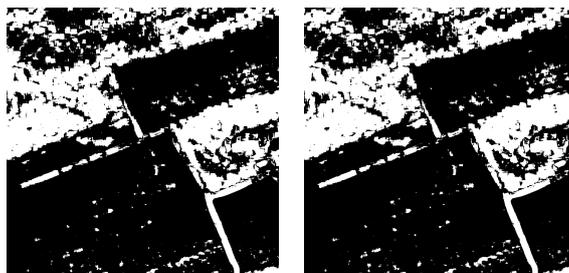


Figura 2.13: Resultado de la clasificación según MCC y MCCC, respectivamente (MAB de primer orden).

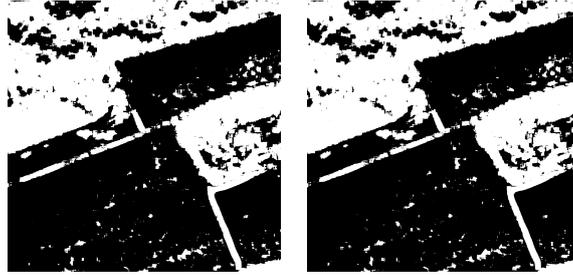


Figura 2.14: Resultado de la clasificación según MCC y MCCC, respectivamente (MAB de segundo orden).

2.4.4. Algunos resultados de los estimadores para un MAB de cuarto y quinto orden ¹.

A continuación se presentan los resultados del análisis descriptivos alcanzados para los estimadores MCC y MCCC, para un MAB de cuarto y quinto orden, además se informa el speed-up de los algoritmos para ambas casos.

Para un MAB de cuarto orden (ver Figura 2.15), se debe considerar agregar sucesivamente a η_s (2.2) los siguientes términos

$$\beta_5 \frac{x_{s-v_5} + x_{s+v_5}}{M} + \beta_6 \frac{x_{s-v_6} + x_{s+v_6}}{M},$$

mientras que para un MAB de quinto orden, además de los anteriores, los siguientes dos términos

$$\beta_7 \frac{x_{s-v_7} + x_{s+v_7} + x_{s-v_8} + x_{s+v_8}}{M} + \beta_8 \frac{x_{s-v_9} + x_{s+v_9} + x_{s-v_{10}} + x_{s+v_{10}}}{M}.$$

siendo $v_5 = (0, 2)$, $v_6 = (2, 0)$, $v_7 = (-2, 1)$, $v_8 = (-1, 2)$, $v_9 = (1, 2)$ y $v_{10} = (2, 1)$.

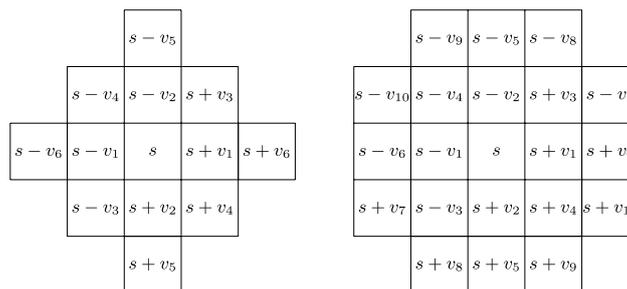


Figura 2.15: Sistemas de vecindades del pixel s , cuarto orden y quinto orden, respectivamente.

Se pueden obtener modelos de orden superior agregando los términos respectivos de una manera similar más allá de los de la Figura 2.15.

En las Figuras 2.16 y 2.17 se presentan las texturas generadas con los valores de los vectores de los parámetros de los casos (i), (ii) y (iii) bajo estudio del MAB de cuarto y quinto orden,

¹Winkler (1995) [79] los denomina MAB de tercer y cuarto orden, en lugar de cuarto y quinto orden, respectivamente.

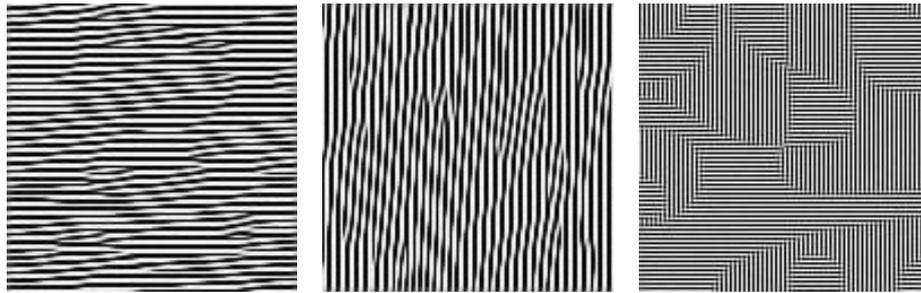


Figura 2.16: Tres ejemplos de realizaciones de MAB de cuarto orden usando los vectores de parámetros especificados: (i) $\theta = (-2, 1, -0.5, 0.5, -0.5, 2, -0.5)$, (ii) $\theta = (-2, -0.5, 1, 0.5, -0.5, -0.5, 2)$ y (iii) $\theta = (-0.5, -0.5, -0.5, -0.5, -0.5, 1.5, 1)$, respectivamente.

respectivamente

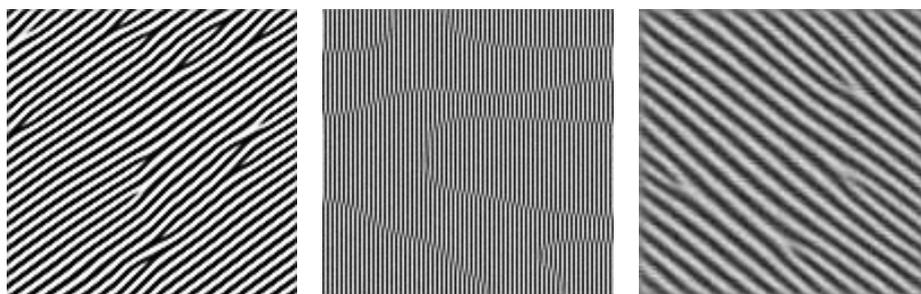


Figura 2.17: Tres ejemplos de realizaciones de MAB de quinto orden usando los vectores de parámetros especificados: (i) $\theta = (-2.5, 0.5, -0.5, 2, -0.5, 0.5, -0.5, 1, -0.5)$, (ii) $\theta = (-2, -0.9, 0.9, 0, 0, 0.5, 0.5, 0, 0.5)$ y (iii) $\theta = (-2, 1, 0, 0, 0, 0.5, 0.5, -0.5, 0.5)$, respectivamente.

Para la implementación del método de estimación MCCC del vector de parámetros de estos modelos, se consideró el patrón de codificación que se muestra en la Figura 2.18.

1	2	3	1	2	3	1	2
4	5	6	4	5	6	4	5
7	8	9	7	8	9	7	8
1	2	3	1	2	3	1	2
4	5	6	4	5	6	4	5
7	8	9	7	8	9	7	8
1	2	3	1	2	3	1	2
4	5	6	4	5	6	4	5

Figura 2.18: Patrón de codificación según sistema de vecinos de cuarto y quinto orden orden (9 codings).

En las tablas 2.17 y 2.18 se reportan la media, el desvío estándar, el coeficiente de variación y el ECM para cada uno de los parámetros de los estimadores bajo estudio. Además, observando las Figuras 2.19 y 2.21, podemos concluir que ambos estimadores presentan una performance comparables, misma conclusión obtenida para el MAB de primer y segundo orden.

Con respecto a la evaluación de los algoritmos, considerando una imagen de entrada de tamaño 128×128 píxeles, el algoritmo del MCC, según la ecuación 2.10, computa una única matriz G de dimensión 15376×7 y 15376×9 , para el MAB de cuarto y quinto orden, respectivamente.

Tabla 2.17: Resultados experimentales: media, desviación estándar (DE), coeficiente de variación (CV %) y ECM de las estimaciones MCC, MCCC en los MAB de cuarto orden.

Parámetro	Caso (i)							
	β_0	β_1	β_2	β_3	β_4	β_5		
Valor verdadero	-2	1	-0.5	0.5	-0.5	-0.5		
Método de estimación								
MCC	Media	-1.95221	0.92544	-0.54113	0.53893	-0.54533	2.10330	-0.52903
	DE	0.02455	0.03058	0.12071	0.06440	0.06433	0.03344	0.00809
	CV %	1.25759	3.30505	22.30749	11.95023	11.79671	1.59002	1.52946
	ECM	0.00288	0.00649	0.01623	0.00565	0.00618	0.01178	0.00090
MCCC	Media	-1.95227	0.92535	-0.54118	0.53899	-0.54535	2.10344	-0.52900
	DE	0.02472	0.03086	0.12102	0.06465	0.06449	0.03376	0.00815
	CV %	1.26660	3.33532	22.36354	11.99451	11.82624	1.60521	1.54204
	ECM	0.00288	0.00652	0.01631	0.00569	0.00620	0.01183	0.00090
Caso (ii)								
Valor verdadero	-2	-0.5	1	0.5	-0.5	-0.5	2	
MCC	Media	-1.95047	-0.55096	0.92514	0.54168	-0.53926	-0.52955	2.10335
	DE	0.02580	0.13342	0.03222	0.07037	0.07160	0.00792	0.03483
	CV %	1.32326	24.21686	3.48289	12.99102	13.27842	1.49706	1.65615
	ECM	0.00311	0.02036	0.00663	0.00667	0.00665	0.00093	0.01189
MCCC	Media	-1.95031	-0.55107	0.92518	0.54155	-0.53921	-0.52954	2.10332
	DE	0.02587	0.13364	0.03230	0.07083	0.07166	0.00797	0.03486
	CV %	1.32653	24.25230	3.49170	13.07959	13.28992	1.50677	1.65763
	ECM	0.00313	0.02043	0.00663	0.00673	0.00666	0.00093	0.01188
Caso (iii)								
Valor verdadero	-0.5	-0.5	-0.5	-0.5	-0.5	1.5	1	
MCC	Media	-0.67730	-0.52170	-0.52154	-0.45911	-0.45971	1.57004	1.06931
	DE	0.20080	0.06809	0.06837	0.06394	0.06680	0.04724	0.03764
	CV %	29.64733	13.05311	13.10931	13.92887	14.53188	3.00927	3.52025
	ECM	0.07167	0.00509	0.00512	0.00575	0.00607	0.00713	0.00621
MCCC	Media	-0.68175	-0.52094	-0.52076	-0.45841	-0.45896	1.57086	1.06996
	DE	0.20172	0.06844	0.06870	0.06458	0.06682	0.04765	0.03830
	CV %	29.58943	13.13954	13.19301	14.08929	14.55944	3.03369	3.58014
	ECM	0.07364	0.00511	0.00514	0.00589	0.00614	0.00728	0.00635

Luego, mediante la ecuación 2.11 determina el valor del estimador. Mientras que, el algoritmo del estimador MCCC, en forma paralela, construye nueve matrices G : cuatro de dimensión 1681×7 , cuatro de dimensión 1725×7 y una matriz de 1765×7 para el MAB de cuarto orden y nueve matrices: cuatro de dimensión 1681×9 , cuatro de dimensión 1725×9 y una matriz de 1765×9 para el MAB de quinto orden. Posteriormente aplica la ecuación 2.13 y finalmente determina el valor del estimador MCCC, promediando los valores de los $\hat{\theta}_i$.

El speed-up de los algoritmos paralelos, se muestran en la siguiente la Tabla 2.19, siendo los casos (i), (ii) y (iii) bajo estudio, los que se presentan en las Figuras 2.16 y 2.17 para el MAB de cuarto orden y quinto orden, respectivamente.

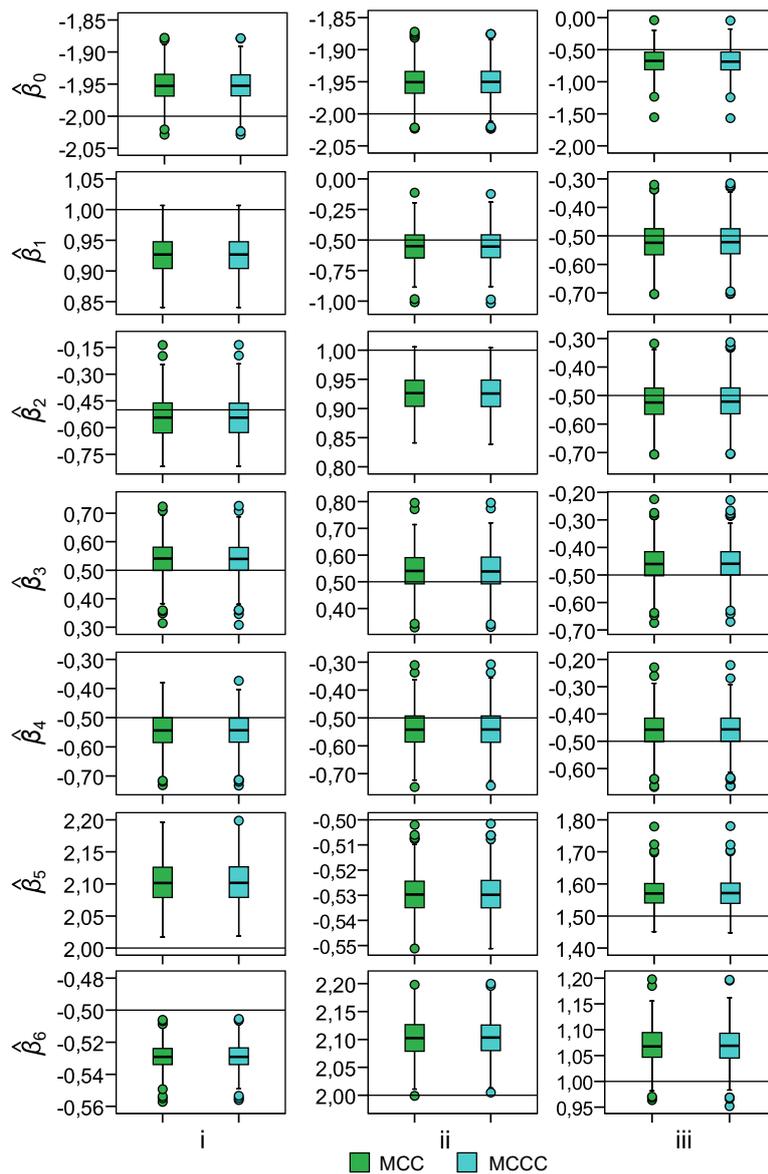


Figura 2.19: Los boxplots de cada columna representan las estimaciones de los parámetros de los modelos de cuarto orden, para los casos (i), (ii) y (iii), respectivamente. La línea horizontal indica el valor del parámetro especificado.

Todos los algoritmos son compilados por MATLAB R2015b y se empleó un servidor de alto desempeño, cuya configuración es: Servidor marca Supermicro AS-4042G-6RF, $4 \times$ AMD Opteron 6168 (48 procesadores), 128 GB \times RAM DDR3 ECC Registrada y 12TB HDD.

Tabla 2.18: Resultados experimentales: media, desviación estándar (DE), coeficiente de variación (CV %) y ECM de las estimaciones MCC, MCCC en los MAB de quinto orden.

Parámetro	Caso (i)									
	β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	
Valor verdadero	-2.5	0.5	-0.5	2	-0.5	0.5	-0.5	1	-0.5	-0.5
Método de estimación										
MCC	Media	-2.52249	0.82128	-0.66658	2.03769	-0.25800	0.36379	-0.57355	0.90891	-0.50995
	DE	0.04247	0.05871	0.04637	0.02473	0.07829	0.05215	0.04972	0.02942	0.01273
	CV %	1.68376	7.14862	6.95766	1.21394	30.34447	14.33646	8.67032	3.23727	2.49726
	ECM	0.00230	0.10666	0.02989	0.00203	0.06467	0.02126	0.00787	0.00916	0.00026
MCCC	Media	-2.52297	0.82122	-0.66659	2.03850	-0.25765	0.36362	-0.57394	0.90861	-0.50967
	DE	0.04272	0.05917	0.04689	0.02490	0.07897	0.05293	0.05008	0.02974	0.01281
	CV %	1.69346	7.20543	7.03473	1.22168	30.64953	14.5574	8.72689	3.27314	2.51448
	ECM	0.00234	0.10668	0.02994	0.00210	0.06495	0.02139	0.00797	0.00923	0.00025
Valor verdadero	-2	-0.9	0.9	0	0	0.5	0.5	0	0	0.5
MCC	Media	-2.03694	-0.90059	0.90710	0.00082	-0.00108	0.50598	0.50944	0.00290	0.50472
	DE	0.17539	0.06069	0.06203	0.05068	0.05391	0.05577	0.05301	0.04637	0.04384
	CV %	8.61064	6.73954	6.83863	6127.18919	4946.32321	11.02220	10.40599	1595.67818	8.68612
	ECM	0.03206	0.00367	0.00389	0.00256	0.00290	0.00314	0.00289	0.00215	0.00194
MCCC	Media	-2.04448	-0.90040	0.90836	0.00218	0.00030	0.50683	0.51030	0.00309	0.50536
	DE	0.17688	0.06088	0.06243	0.05093	0.05367	0.05636	0.05347	0.04686	0.04415
	CV %	8.65202	6.76240	6.87300	2330.01837	17652.32810	11.12080	10.47808	1515.24336	8.73700
	ECM	0.03320	0.00370	0.00395	0.00259	0.00287	0.00321	0.00295	0.00220	0.00197
Valor verdadero	-2	1	0	0	0	0.5	0.5	-0.5	-0.5	0.5
MCC	Media	-2.01097	1.00381	0.00049	0.00291	-0.00132	0.50402	0.50420	-0.50357	0.50200
	DE	0.01428	0.03154	0.03596	0.03383	0.03739	0.02872	0.02739	0.01628	0.02242
	CV %	0.71044	3.14224	7318.17848	1160.92787	2832.33266	5.69872	5.43229	3.23393	4.46751
	ECM	0.00032	0.00100	0.00129	0.00115	0.00139	0.00083	0.00076	0.00027	0.00050
MCCC	Media	-2.01117	1.00371	0.00017	0.00293	-0.00128	0.50426	0.50457	-0.50364	0.50204
	DE	0.01432	0.03137	0.03617	0.03395	0.03744	0.02880	0.02748	0.01631	0.02248
	CV %	0.71245	3.12583	21099.8891	1155.67336	2920.56811	5.71141	5.44734	3.24013	4.47849
	ECM	0.00032	0.00099	0.00130	0.00115	0.00140	0.00084	0.00077	0.00027	0.00050

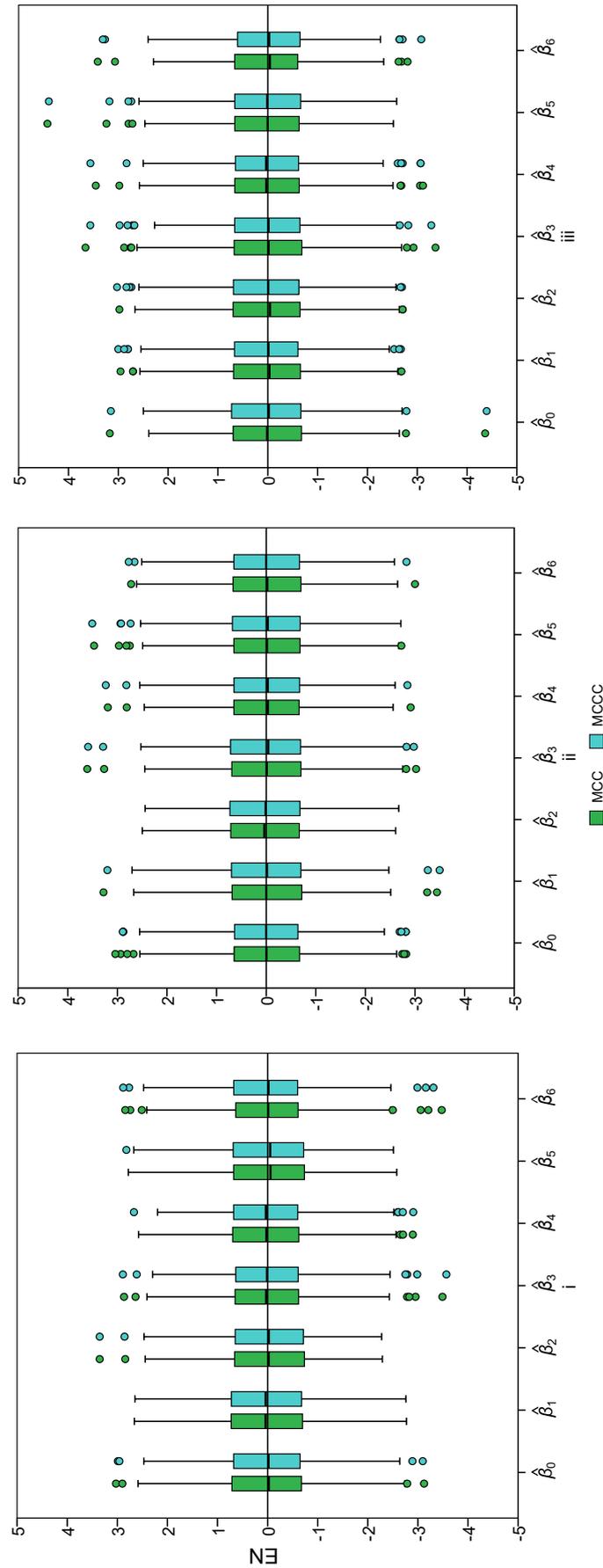


Figura 2.20: Los boxplots de cada fila representan la distribución de las estimaciones normalizadas (EN) de los parámetros de los modelos de cuarto orden, para los casos (i), (ii) y (iii), respectivamente.

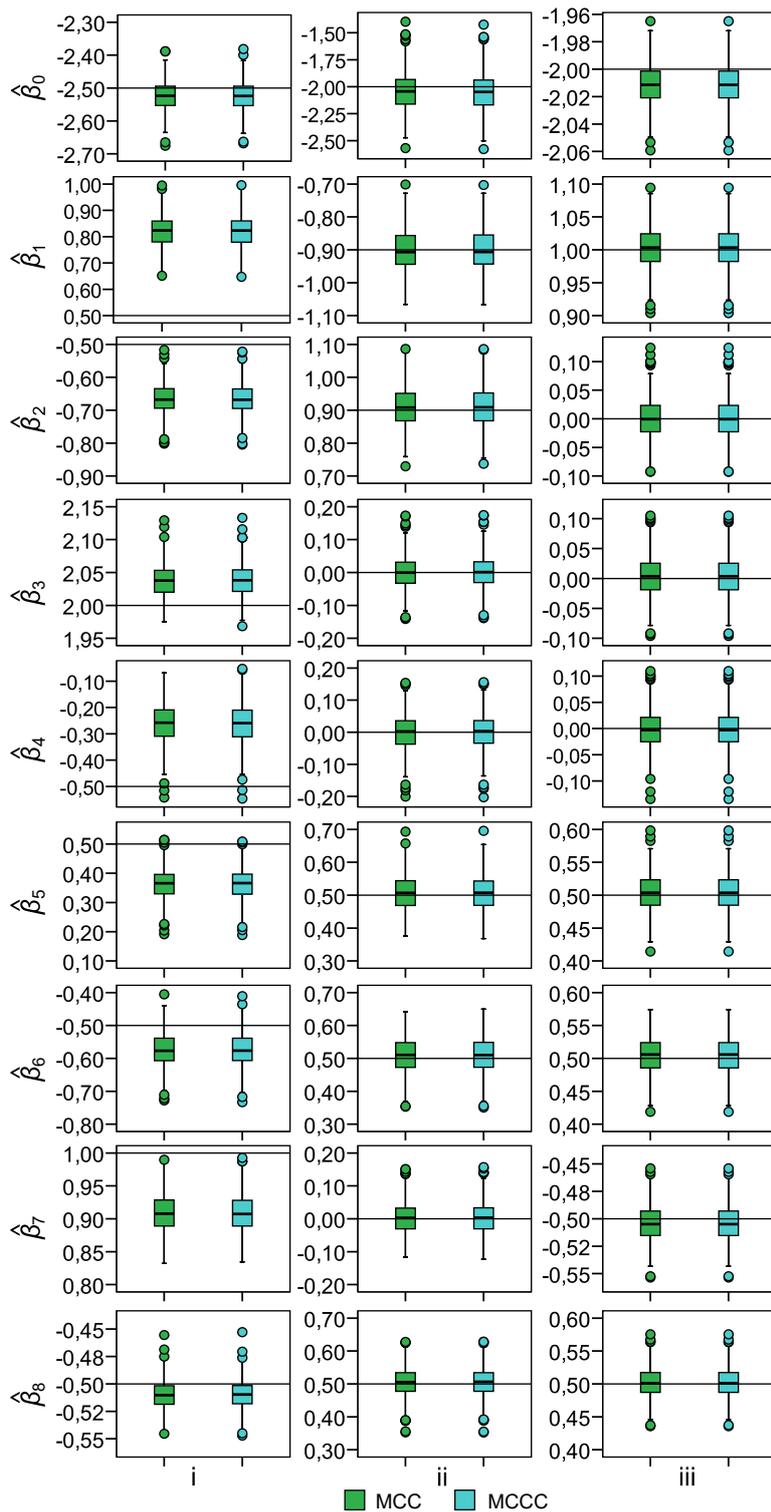


Figura 2.21: Los boxplots de cada columna representan la distribución de las estimaciones de los parámetros de los modelo de quinto orden, para los casos (i), (ii) y (iii), respectivamente. La línea horizontal indica el valor del parámetro especificado.

Tabla 2.19: Speed-up del MCCC para el MAB de cuarto y quinto orden: media \pm desvío estándar.

Casos	Orden 4	Orden 5
(i)	1.2941 \pm 0.1810	1.4947 \pm 0.1766
(ii)	1.2752 \pm 0.2113	1.4522 \pm 0.1533
(iii)	1.2899 \pm 0.2398	1.4550 \pm 0.1782

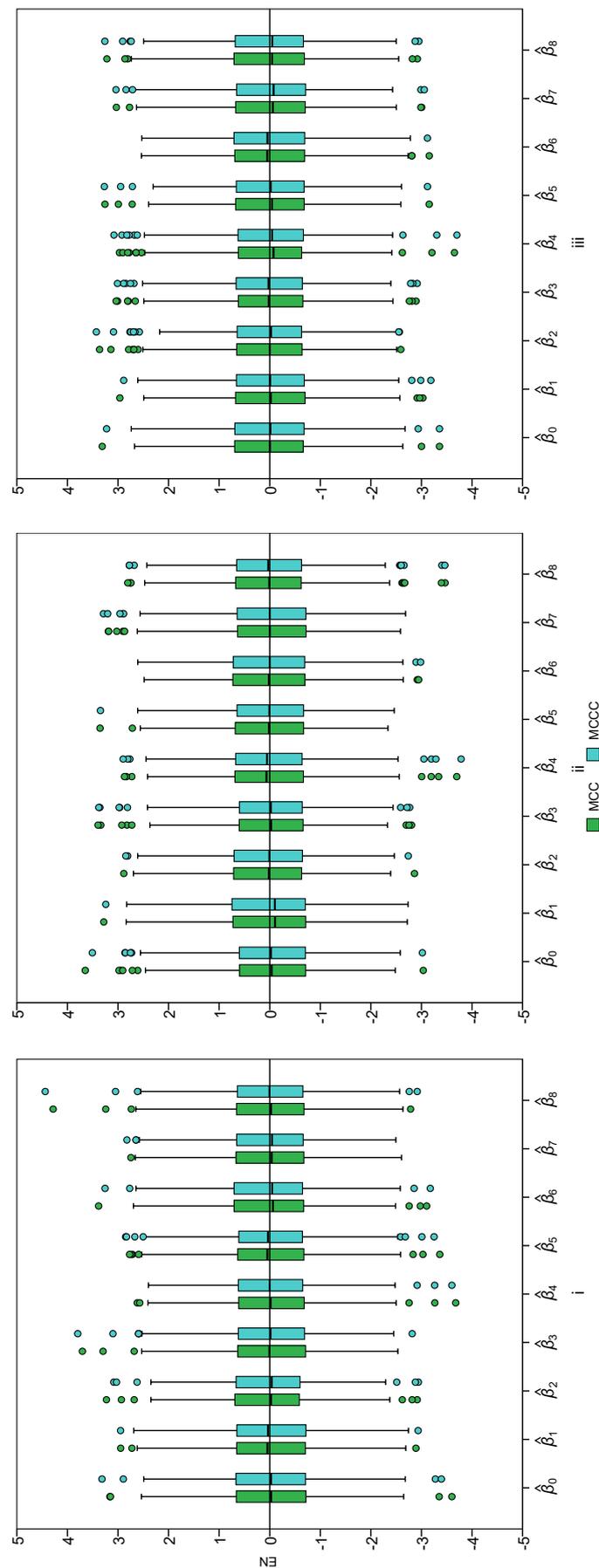


Figura 2.22: Los boxplots de cada fila representan las estimaciones normalizadas (EN) de los parámetros de los modelos de quinto orden, para los casos (i), (ii) y (iii), respectivamente.

Tabla 2.20: Resultados del cálculo del cociente entre el tiempo de ejecución del MCC y MCCC para el MAB de cuarto y quinto orden: media \pm desvío estándar.

Casos	Orden 4	Orden 5
(i)	8.7792 \pm 1.3118	11.3670 \pm 1.6760
(ii)	8.6463 \pm 1.3896	10.7064 \pm 2.1326
(iii)	8.8116 \pm 1.3404	10.7174 \pm 2.2389

Capítulo 3

Conclusión y estudios futuros

En este capítulo se exponen, en forma breve, las principales conclusiones obtenidas del trabajo desarrollado y se plantean los estudios futuros a realizar.

3.1. Conclusión

En esta primer parte de la tesis, se propone un teorema de descomposición del potencial de Gibbs basada en las contribuciones de las configuraciones en conjuntos de ubicaciones no vecinas, llamados *codings*, que divide a la grilla en conjuntos finitos disjuntos. Esta propuesta, esencialmente, explota la independencia condicional bajo el modelo GMRF entre los elementos de estos conjuntos de codificaciones y, por lo tanto, permite la computación paralela. Luego, se propone un nuevo método de estimación de los parámetros de un modelo textural, el modelo auto-binomial (MAB) que utiliza un esquema de codificación, que se llamó mínimos cuadrados condicionales coding (MCCC) y fue implementado en un algoritmo paralelo. Los experimentos muestran que el proceso de estimación realizado con MCCC tiene un tiempo de ejecución más bajo conservando las características del método mínimo cuadrado condicional (MCC): la simplicidad en el cálculo y una buena estimación. También se obtuvo similar comportamiento al testear su sensibilidad a la presencia de ruido y su performance en una aplicación real, la clasificación supervisada de una imagen satelital.

3.2. Estudios futuros

La popularidad de la arquitectura multinúcleo y el aumento de los recursos de cómputo en dichas plataformas brindan una nueva oportunidad para acelerar la velocidad de procesamiento de los algoritmos. Este hecho sugiere avanzar en el uso de técnicas de programación paralela en el contexto del proceso de estimación, continuando en este caso, en el estudio y performance del

método de estimación propuesto MCCC, para modelos MAB de ordenes superiores y abordar la problemática planteada en la industria textil, señalada en el reciente trabajo de Silvestre-Blanes et al. (2019) [66], la detección de defectos textiles, donde el principal objeto de estudio es la textura de la tela.

La siguiente Figura 3.1 muestra tres imágenes de telas provenientes de fábricas textiles distintas, donde se puede apreciar las diferentes texturas y tramas; mientras que en las Figuras 3.2 y 3.2, se puede observar una tela con un defecto fácilmente perceptible: grieta en la trama y otro defecto pequeño: bola borrosa, junto con sus respectivos mapas de clases.

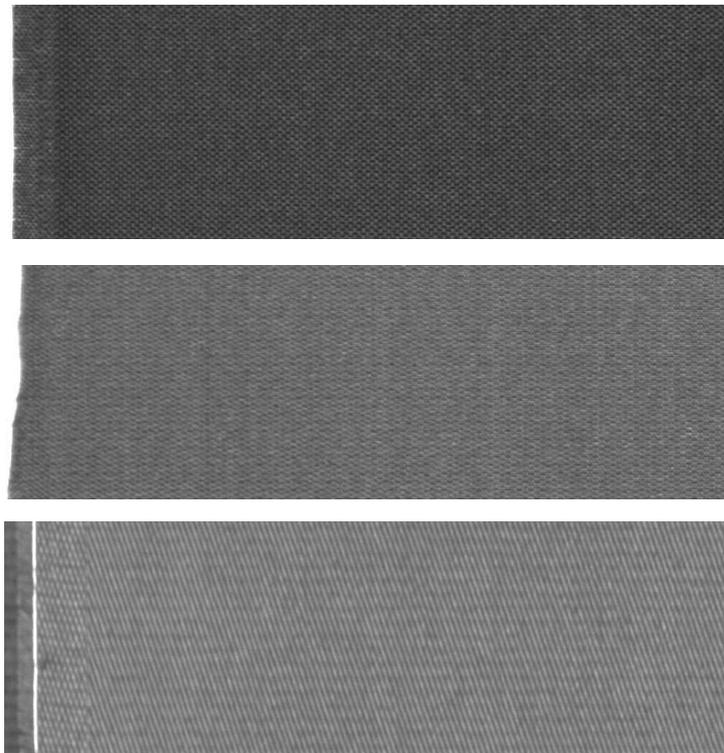


Figura 3.1: Tres ejemplos de imágenes de telas perteneciente a la base de imágenes *Autex* de fábricas textiles diferentes (00 - 01 y 06) (4096×256)



Figura 3.2: Imagen 0080_036_01 perteneciente a la base de imágenes *Autex*: imagen original y su mapa de clase (defecto: grieta en la trama)

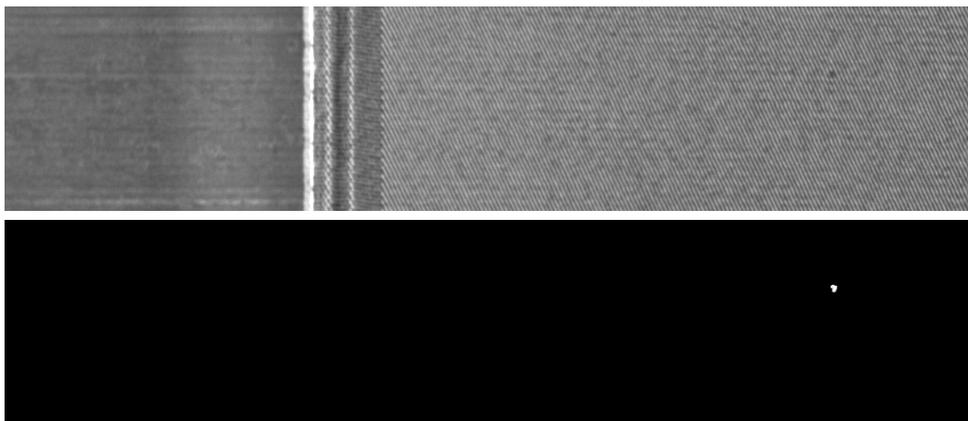


Figura 3.3: Imagen 0024_019_02 perteneciente a la base de imágenes *Autex*: imagen original y su mapa de clase (defecto: bola borrosa)

Parte II

Calidad de una imagen

Introducción

El objetivo de la evaluación de la calidad de una imagen (image quality assessment - IQA) es modelar la percepción del sistema visual humano (human visual system - HVS) para evaluar la calidad subjetiva de la misma. Dependiendo de la disponibilidad de una imagen de referencia con “calidad perfecta”, IQA puede clasificarse en: con referencia completa, donde la imagen de referencia es totalmente accesible para evaluar la imagen distorsionada; con referencia reducida, cuando se dispone de información parcial de la imagen de referencia y sin referencia, cuando no hay acceso a una imagen de referencia [77]. Para el primer caso, IQA con referencia completa, resulta adecuado hablar de fidelidad de la imagen en lugar de calidad de la imagen, ya que el objetivo es medir la similitud entre dos imágenes. Mientras que en los casos de las IQA con referencia reducida y sin referencia, se modela la IQA subjetiva solo a partir del conocimiento de una imagen distorsionada. Esta situación se encuentra más cerca del mundo real donde la imagen original no está disponible.

Las propiedades estadísticas del entorno visual natural juegan un papel fundamental en la evolución, el desarrollo y la adaptación de HVS. Una característica importante sobre las señales de imágenes naturales es que están altamente estructuradas. Por “señal estructurada”, se entiende que en las muestras de señal exhiben fuertes dependencias entre ellas, especialmente cuando son espacialmente próximas. Estas dependencias llevan información importante sobre la estructura de los objetos en la escena visual. La hipótesis principal de la evaluación de calidad de imagen basada en similitud estructural es que el HVS está altamente adaptado para extraer información estructural del campo visual, y por lo tanto, la medición de la similitud estructural (o distorsión) debe proporcionar una buena aproximación a la calidad de imagen percibida [80].

En estos últimos años, hubo un creciente interés en el desarrollo de medidas de calidad objetivas. Estas medidas juegan un rol importante en una amplia gama de aplicaciones del campo de procesamientos de imágenes, tales como adquisición de imágenes, restauración, visualización y fusión de imágenes, entre otras.

La fusión de imágenes es el proceso de combinar información disponible de dos o más imágenes de una escena en una sola imagen compuesta, también llamada imagen fusionada, que es

más informativa y adecuada tanto para la percepción visual como para el procesamiento por computadora. El objetivo de la fusión de imágenes es reducir la incertidumbre y minimizar la redundancia en la salida, mientras se maximiza la información relevante para una aplicación particular de interés. La evaluación de calidad de diferentes esquemas de fusión de imágenes se lleva a cabo tradicionalmente mediante evaluaciones subjetivas [68]. Aunque este método es confiable, es costoso y demasiado lento para las aplicaciones del mundo real. Dado que la evaluación objetiva en este contexto es una tarea difícil debido a los diferentes escenarios de aplicación y a la falta una imagen de referencia, se han propuesto varias medidas objetivas de calidad de imagen que son consistentes con la percepción visual humana para predecir automáticamente la calidad de fusión de la imagen.

De acuerdo a Liu et al. (2012) [44], se clasificaron en cuatro grupos según sus características:

- medidas basadas en la teoría de la información
- medidas basadas en características de la imagen
- medidas de fusión inspiradas en la percepción humana
- medidas basadas en similitudes estructurales de la imagen

En el contexto de las medidas basadas en la similitud estructural de la imagen, se desarrollaron la medida de Piella [56], la medida de Cvejic [19] y la medida de Yang [82]. Estas medidas de calidad de fusión utilizan el índice de calidad de imagen universal (Q) [74] como base. No requieren una imagen de referencia, son de propósito general y se basan en un enfoque de ventana deslizante, teniendo en cuenta las medidas locales para estimar cuánta información relevante contenida en las imágenes de entrada está presente en la imagen fusionada.

En esta segunda parte de la tesis, se presenta una nueva medida de calidad objetiva de imagen en el contexto de fusión, basada en el índice similitud estructural de la imagen. Para su validación se testea su propósito general, la estabilidad y su capacidad discriminativa, como así también la correlación con la opinión subjetiva. Por último, se presenta una potencial aplicación de la misma, en el marco de la obtención de una fotografía de alta calidad de pinturas.

Capítulo 4

Medidas de calidad

Esta sección se proporciona una descripción de las medidas de similitud estructural de la imagen y luego se presenta las medidas basadas en la similitud estructural para la fusión de imágenes.

En lo que sigue, se considera $x = \{x_{i,j} | 1 \leq i \leq M, 1 \leq j \leq N\}$ e $y = \{y_{i,j} | 1 \leq i \leq M, 1 \leq j \leq N\}$, con $M, N \in \mathbb{N}$, como las imágenes original y de prueba, respectivamente.

4.1. Medidas de similitud estructural de la imagen

La información estructural es una de los componentes importantes que afectan la calidad visual de las imágenes digitales. Estas medidas de similitud estructural de la imagen se basan en el supuesto de que el sistema visual humano está altamente adaptado para extraer información estructural del campo de visión. La información estructural en una imagen esta caracterizada por aquellos atributos que representan la estructura de los objetos en la escena, independientemente de la luminosidad y el contraste promedio.

4.1.1. Índice de calidad de imagen universal (Q)

Wang y Bovik (2002) [74] introdujeron el índice de calidad de imagen universal Q y se define de la siguiente manera:

$$Q(x, y) = \frac{4\bar{x}\bar{y}s_{xy}}{(\bar{x}^2 + \bar{y}^2)(s_x^2 + s_y^2)}, \quad (4.1)$$

donde \bar{x} y \bar{y} son los valores promedio de la muestra de las imágenes x y y , s_x , s_y y s_{xy} son las desviaciones y la covarianza de la muestra, respectivamente.

Este índice se puede descomponer en los siguientes tres factores diferentes:

$$Q(x, y) = l(x, y) \cdot c(x, y) \cdot s(x, y) = \frac{2\bar{x}\bar{y}}{\bar{x}^2 + \bar{y}^2} \cdot \frac{2s_x s_y}{s_x^2 + s_y^2} \cdot \frac{s_{xy}}{s_x s_y}. \quad (4.2)$$

El primer factor en (4.2) mide la similitud entre la luminancia de x e y , y el segundo factor mide la similitud relacionada con el contraste entre las imágenes y el tercer factor mide el grado de correlación lineal¹ entre x e y .

Teniendo en cuenta que las señales de imagen generalmente no son estacionarias y las distorsiones de la imagen pueden variar en el espacio, es más apropiado medir el índice de calidad de imagen Q sobre las regiones locales y luego combinar los diferentes resultados en una sola medida. Por lo tanto, Wang y Bovik (2002) [74] proponen utilizar un enfoque de ventana deslizante: comenzando desde la esquina superior izquierda de las dos imágenes x , y , una ventana deslizante de un tamaño fijo (bloque por bloque) sobre toda la imagen hasta llegar a la esquina inferior derecha. Finalmente, el índice general de calidad de imagen Q se determina promediando todos los índices de calidad locales $Q(x, y|w)$ calculado en todas las ventanas deslizantes $w \in W$

$$Q = \sum_{w \in W} \frac{Q(x, y|w)}{|W|}, \quad (4.3)$$

con W la familia de todas las ventanas y $|W|$ es la cardinalidad de W .

4.1.2. Índice de similitud estructural (SSIM)

Para medir las distorsiones estructurales de dos imágenes x e y , Wang et al. propuso un índice de similitud estructural (SSIM) [76], definido de la siguiente manera:

$$\begin{aligned} \text{SSIM}(x, y) &= [l(x, y)]^\alpha \cdot [c(x, y)]^\beta \cdot [s(x, y)]^\gamma \\ &= \left(\frac{2\bar{x}\bar{y} + C_1}{\bar{x}^2 + \bar{y}^2 + C_1} \right)^\alpha \cdot \left(\frac{2s_x s_y + C_2}{s_x^2 + s_y^2 + C_2} \right)^\beta \cdot \left(\frac{s_{xy} + C_3}{s_x s_y + C_3} \right)^\gamma, \end{aligned} \quad (4.4)$$

Los parámetros α , β y γ ajustan la importancia relativa de los tres componentes. Las constantes C_1 , C_2 y C_3 , denominadas constantes de estabilidad, se incluyen para evitar la inestabilidad cuando los denominadores están muy cerca de cero y para adaptarse a la percepción del HVS. Para simplificar la expresión (4.4), Wang et al. establecieron $\alpha = \beta = \gamma = 1$ y $C_3 = C_2/2$. Esto da como resultado una forma específica del índice SSIM:

$$\text{SSIM}(x, y) = \frac{2\bar{x}\bar{y} + C_1}{\bar{x}^2 + \bar{y}^2 + C_1} \cdot \frac{2s_x s_y + C_2}{s_x^2 + s_y^2 + C_2}. \quad (4.5)$$

En [76], los autores consideran $C_1 = (K_1 L)^2$ y $C_2 = (K_2 L)^2$, donde L es el rango dinámico de los valores de píxel (255 para imágenes en escala de grises de 8 bits) y K_1 y K_2 son constantes positivas pequeñas. La medida SSIM usa la configuración de parámetros: $K_1 = 0,01$ y $K_2 = 0,03$.

¹También llamado coeficiente de correlación de Pearson.

Estos valores son algo arbitrarios, pero Wang et al. descubrieron que en sus experimentos, el rendimiento del algoritmo de índice SSIM era poco sensible a las variaciones de estos valores.

El índice Q , definido en (4.1), corresponde a un caso particular del índice SSIM, si $C_1 = C_2 = 0$.

Para evaluar la calidad general de la imagen, se utiliza el índice medio SSIM (MSSIM), que consiste en el cálculo promediado de los índices locales $SSIM(x, y|w)$ en $w \in W$,

$$MSSIM = \sum_{w \in W} \frac{SSIM(x, y|w)}{|W|}, \quad (4.6)$$

La filosofía general detrás del índice SSIM es distinguir las distorsiones estructurales de las distorsiones no estructurales. Las distorsiones estructurales son las que más afectan la percepción de degradación de una imagen, mientras que las distorsiones no estructurales solo la afectan ligeramente. Los dos primeros factores del índice SSIM (l y c) explican las distorsiones no estructurales, mientras que las distorsiones estructurales están representadas por una pérdida de correlación lineal.

Como se explica en [76], el término de luminancia del índice SSIM está relacionado con la Ley de Weber [71]: la percepción del cambio de un estímulo es proporcional a la intensidad del estímulo. Por lo tanto, la Ley de Weber puede expresarse de la siguiente manera:

$$\frac{\Delta I}{I} = C$$

donde I es la intensidad, ΔI es el cambio de intensidad y C es una constante. Dada una imagen x de intensidad de luminancia I_x , la intensidad I_y de la luminancia de una imagen y para que la diferencia sea perceptible debe ser

$$\frac{I_y - I_x}{I_x} = C$$

Esto implica que para cualquier distorsión apenas perceptible (just-noticeable distortion-JND²), la relación I_y/I_x es constante. Si la constante de estabilidad es cero, esta propiedad se verifica para la primera componente del índice SSIM. En efecto,

$$\frac{2\bar{x}\bar{y}}{\bar{x}^2 + \bar{y}^2} = k \Leftrightarrow \frac{\bar{y}}{\bar{x}} = \frac{1 \pm \sqrt{1 - k^2}}{k}$$

de esta manera, la relación de intensidad de luminancia es constante para cada nivel de distorsión. Dado que la propiedad se mantiene no solo para JND, sino también para distorsiones por encima

²JND se refiere a la distorsión máxima que el sistema visual humano (HVS) no puede percibir

del umbral, el índice SSIM extiende la Ley de Weber.

La Ley de Weber no solo se aplica cuando $I_x = \bar{x}$, la luminancia, sino también cuando $I_x = s_x$, el contraste de la imagen. Si las constantes de estabilidad son cero, se llega a la misma conclusión, el cociente de contrastes s_x/s_y es constante para un nivel de distorsión de SSIM constante.

Las constantes C_1, C_2 y C_3 representan el umbral de percepción. De hecho, la Ley de Weber no se cumple para intensidades bajas, ya que hay un cambio mínimo absoluto de intensidad que debe observarse. Tener en cuenta la similitud con el modelo de normalización divisiva [70], donde hay una constante en el denominador.

La utilización del enfoque de la ventana deslizante w está relacionada con el funcionamiento del HVS, donde la mayoría de los receptores se concentran en la fovea y el movimiento ocular es necesario para obtener una imagen completa del entorno circundante. En [76], se toma una ventana gaussiana truncada de tamaño 11×11 con desviación estándar de 1.5 píxeles.

4.1.3. Una clase de métricas basadas en SSIM

En esta sección se presenta una clase de métricas para imágenes basadas en el índice SSIM, la cual considera los dos factores de este índice como componentes de un par ordenado. Para más detalles consultar Brunet et al. (2012) [6].

El error cuadrático medio (mean squared error - MSE) es la métrica más difundida y ampliamente utilizada en el marco del procesamiento de imágenes, debido a que es simple de calcular, tiene un claro significado físico y posee varias propiedades matemáticas convenientes en el contexto de optimización. Sin embargo, es conocido que las medidas basadas en norma L^2 , como por el ejemplo MSE, no necesariamente correlacionan bien con la percepción humana de la calidad de una imagen [75].

El siguiente ejemplo extraído de [74], evidencia esta falla del MSE para propósitos de IQA. El mismo consiste en tomar varias imágenes distorsionadas con igual MSE de la imagen original, obtenidas a partir de diferentes tipos de distorsiones. Como se puede observar, la calidad percibida en cada imagen difiere, mientras que MSE predice erróneamente una calidad similar. Por otro lado, el índice SSIM exhibe un mejor acuerdo con la opinión humana respecto a la calidad de las imágenes (ver Figura 4.1).

A diferencia del MSE, el índice SSIM no es una métrica pues no verifica la desigualdad triangular. Sin embargo, es posible definir una clase de métricas basadas en él, que serán presentadas a continuación.

En lo que sigue, sea \mathbb{R}_+^N el espacio de bloques de señal/imagen N -dimensionales no negativos,



Figura 4.1: Ejemplo Lena. todas las imagenes poseen $MSE = 225$. Primera fila: imagen original, segunda y tercera fila: imagen distorsionada con desplamiento medio (Mean Shift) (0,9894), sal y pimienta (0,6494), AWGN (0,3891) y borroso (0,3461), respectivamente. Entre paréntesis figura el valor de SSIM.

es decir, $x \in \mathbb{R}_+^N$ implica que $x = (x_1, x_2, \dots, x_N)$, con $x_k \geq 0, 1 \leq k \leq N$. También se considera que la distancia L^2 entre dos de estas señales $x, y \in \mathbb{R}_+^N$ es el error cuadrático medio (root mean square error - RMSE) habitual, denotado como sigue, $\|x - y\|_2 = \left[\frac{1}{N} \sum_{k=1}^N (x_k - y_k)^2 \right]^{1/2}$

A partir de la definición original del índice SSIM (4.4), se obtiene la fórmula simplificada (4.5), la cual posee solo dos componentes, $S_1 = \frac{2\bar{x}\bar{y}+C_1}{\bar{x}^2+\bar{y}^2+C_1}$ y $S_2 = \frac{2s_x s_y + C_2}{s_x^2 + s_y^2 + C_2}$, y es frecuentemente usada en las aplicaciones. La primer componente S_1 mide la similaridad de los valores medios, \bar{x} e \bar{y} , de x e y , respectivamente. Su forma funcional fue elegida originalmente en un esfuerzo por adaptarse a la ley de percepción de Weber [71]. El componente S_2 es una combinación de la correlación y una medida de distorsión de contraste (similitud entre las variaciones) entre x e y . Su forma funcional sigue la idea de la normalización divisiva [70]. Dado que x e $y \in \mathbb{R}_+^N$, esto sigue que $0 \leq S_1 \leq 1$ y $S_1(x, y) = 1$ si y solo si $\bar{x} = \bar{y}$. Además, notar que $-1 \leq S_2 \leq 1$ y $S_2 = 1$ si y solo si $x - \bar{x} = y - \bar{y}$. Esto implica que $-1 \leq \text{SSIM} \leq 1$ y que, para señales no-negativas, $\text{SSIM} = 1$ si y sólo si $x = y$. (Un valor negativo de SSIM implica que x e y están correlacionadas negativamente). Esto sugiere que la función

$$T(x, y) = 1 - \text{SSIM}(x, y)$$

puede actuar como un tipo de función distancia, dado que $x = y$ implica que $T(x, y) = 0$. Notar que $0 \leq T(x, y) \leq 2$.

Si se examina la componente S_1 se tiene

$$\begin{aligned} 1 - S_1(x, y) &= 1 - \frac{(2\bar{x}\bar{y} + C_1)}{(\bar{x}^2 + \bar{y}^2 + C_1)} \\ &= \frac{|\bar{x} - \bar{y}|^2}{\bar{x}^2 + \bar{y}^2 + C_1}. \end{aligned} \quad (4.7)$$

El lado derecho de (4.7) puede ser visto como una distancia L^2 cuadrada normalizada entre los valores medios \bar{x} y \bar{y} . Para S_2 , se tiene

$$\begin{aligned} 1 - S_2(x, y) &= 1 - \frac{2s_x s_y}{s_x^2 + s_y^2 + C_2} \\ &= \frac{s_x^2 + s_y^2 - 2s_x s_y}{s_x^2 + s_y^2 + C_2}. \end{aligned} \quad (4.8)$$

Un caso particular $\bar{x} = \bar{y} = 0$

$$1 - S_2(x, y) = \frac{\|x - y\|^2}{\|x\|^2 + \|y\|^2 + \frac{N-1}{N}C_2}. \quad (4.9)$$

el cual es además una distancia L^2 cuadrada normalizada entre x e y . Las ecuaciones (4.7) y (4.9) sugieren que es natural considerar métricas basadas en SSIM las cuales operan en una descomposición de señales en sus componentes medias y media-cero.

Para definir esta clase de métricas basadas en el índice SSIM, es necesario el siguiente Lema 4.1.1, que establece una descomposición ortogonal del espacio de señal/imagen.

Lema 4.1.1 *Sea $\{X_m\}_{m=1}^M$ una descomposición ortogonal de \mathbb{R}^N , i.e., $\bigcup_{m=1}^M X_m = \mathbb{R}^N$, $X_i \perp X_j \forall i \neq j$. Entonces, cada elemento x de \mathbb{R}^N puede ser expresado unicamente como $x = \sum_{m=1}^M x_m$ donde $x_m = P_m(x)$ son las proyecciones de x en X_m .*

Una clase de métricas bidimensionales

A partir de las ecuaciones (4.7) y (4.9), y considerando el Lema 4.1.1, es razonable descomponer \mathbb{R}_+^N en la siguiente suma directa $\mathbb{R}_+^N = \mathbb{R}_1 \oplus \mathbb{R}_2$, donde $\mathbb{R}_1 = \text{gen}(1) = \text{gen}(1, 1, \dots, 1)$ y $\mathbb{R}_2 = \mathbb{R}_1^\perp$.

Sea d_1 una métrica en \mathbb{R} y d_2 una métrica en \mathbb{R}^{N-1} . Luego, para dos elementos $x, y \in \mathbb{R}^N$, se define:

$$d := (d_1(\bar{x}, \bar{y}), d_2(x - \bar{x}, y - \bar{y})) \in \mathbb{R}^2. \quad (4.10)$$

Es claro que $x = y$ implica que $d = 0$. El próximo resultado muestra que d puede ser usado para definir una métrica en \mathbb{R}^N .

Teorema 4.1.1 *Sea $\|\cdot\|$ una norma en \mathbb{R}^2 que satisface la siguiente propiedad de incremento en \mathbb{R}_+^2 . Para cualquier $a \in \mathbb{R}_+^2$ y cualquier par ordenado $b = (b_1, b_2)$ con $b_1, b_2 > 0$,*

$$\|a + b\| \geq \|a\|, \quad (4.11)$$

Entonces para d definido en (4.10),

$$d(x, y) := \|d\|, \quad (4.12)$$

es una métrica en \mathbb{R}^N .

Para demostrar que $d(x, y)$ en (4.12) es métrica, es fácil ver que $d(x, y) = d(y, x)$ (simetría), $d(x, y) > 0$ (positividad), $d(x, y) = 0$ si y sólo si $x = y$ (positividad estricta), resta ver si $d(x, y)$ satisface la desigualdad triangular, i.e., para cualquier $x, y \in \mathbb{R}^N$, $d(x, y) \leq d(x, z) + d(z, y)$, para poder satisfacer esta propiedad, es necesario asumir que d_1 y d_2 son métricas y que la norma $\|\cdot\|$ satisface la propiedad de incremento.

Nota 4.1.1 La propiedad de incremento (4.11) también es válida para la norma-ponderada L^p , por ejemplo,

$$\|x\| = \left[\sum_{i=1}^N w_{k,p} |x_k|^p \right]^{1/p}, \quad (4.13)$$

donde $w_{k,p} > 0$ para $1 \leq k \leq N$. Pero (4.11) no se cumple para todas las normas.

La métrica normalizada relevante para SSIM

Teorema 4.1.2 Para $M \geq 1$, sea $\|\cdot\|_2$ la norma L^2 en \mathbb{R}^M . Entonces $C \geq 0, \bar{d}: \mathbb{R}^M \times \mathbb{R}^M \rightarrow \mathbb{R}$, dada por

$$\bar{d}(x, y) := \frac{\|x - y\|}{\sqrt{\|x\|^2 + \|y\|^2 + C}}, \quad (4.14)$$

es una métrica

Demostración:

Este teorema se demostró para el caso $C = 0$ en [37]. Ver [6], para $C > 0$.

Construcción de la métrica basada en SSIM

Ahora es posible definir la métrica basada en SSIM. La métrica normalizada \bar{d} es utilizada en cada subespacio definidos anteriormente (Ver Lema 4.1.1).

Dados x y $y \in \mathbb{R}^N$, se define el siguiente vector de métricas

$$d(x, y) = (d_1(\bar{x}, \bar{y}), d_2(x - \bar{x}, y - \bar{y})) \in \mathbb{R}^2, \quad (4.15)$$

donde

$$\begin{aligned} d_1(\bar{x}, \bar{y}) &= \bar{d}(\bar{x}, \bar{y}) = \frac{\|x - y\|}{\sqrt{\bar{x}^2 + \bar{y}^2 + C_1}} \\ d_2(x - \bar{x}, y - \bar{y}) &= \bar{d}(x - \bar{x}, y - \bar{y}) = \frac{\|(x - \bar{x}) - (y - \bar{y})\|_2}{\sqrt{\|x - \bar{x}\|_2^2 + \|y - \bar{y}\|_2^2 + \frac{(N-1)}{N}C_2}} \end{aligned} \quad (4.16)$$

Sea $\|\cdot\|$ cualquier norma en \mathbb{R}^2 que satisface la propiedad de incremento definida en el Teo-

rema 4.1.1. A partir de este resultado, se tiene la métrica en \mathbb{R}^N :

$$D(x, y) = \left\| \left(\bar{d}(\bar{x}, \bar{y}), \bar{d}(x - \bar{x}, y - \bar{y}) \right) \right\|, \quad (4.17)$$

En el caso que $\|\cdot\| = \|\cdot\|_p$, la norma ponderada L^p en \mathbb{R}^2 , con $p \geq 1$, la métrica explícitamente es

$$\begin{aligned} D_p(x, y) &= \left\| \left(\bar{d}(\bar{x}, \bar{y}), \bar{d}(x - \bar{x}, y - \bar{y}) \right) \right\|_p \\ &= \left(w_{1,p} [\bar{d}(\bar{x}, \bar{y})]^p + w_{2,p} [\bar{d}(x - \bar{x}, y - \bar{y})]^p \right)^{1/p} \end{aligned} \quad (4.18)$$

Para el caso $p = 2$ se obtiene

$$D_2 = \sqrt{\bar{d}^2(\bar{x}, \bar{y}) + \bar{d}^2(x - \bar{x}, y - \bar{y})}, \quad (4.19)$$

Validación perceptiva de la modificación de SSIM

En aplicaciones prácticas, la métrica obtenida a partir de la modificaciones del índice SSIM, es esencialmente equivalente al índice original SSIM. Esta afirmación se muestra vía un experimento perceptual.

Teniendo en cuenta el Teorema 4.1.1, para el caso particular de un vector de métricas dado en (4.15), una norma ponderada l^p , como en (4.13), con $w = (1, 1)$ y $p = 2$ se obtiene

$$\begin{aligned} D_2 := \|d(x, y)\|_2 &= \sqrt{d_1^2(\bar{x}, \bar{y}) + d_2^2(x - \bar{x}, y - \bar{y})} \\ &= \sqrt{2 - S_1(x, y) - S_2(x, y)} \end{aligned}$$

La métrica D_2 puede ser vista como una estimación de orden inferior de $\sqrt{1 - \text{SSIM}}$. Ver Brunet et al.(2012) [6].

Las distorsiones más típicas de una imagen son ruido aditivo, difuminado y compresión de imagen que generalmente preserva el promedio local de las imágenes, tal que $\bar{x} \cong \bar{y}$.

Para la comparación entre D_2 , $\sqrt{1 - \text{SSIM}}$ y SSIM, se consideraron las bases: TID2008 [59] (1700 imágenes distorsionadas, ver Figura 4.2), y LIVE [65] (779 imágenes distorsionadas). El coeficiente de correlación de Pearson resultante entre D_2 y $\sqrt{1 - \text{SSIM}}$ es 0.9998 y 0.9986, respectivamente. Mientras que al comparar los puntajes de D_2 y SSIM con la opinión subjetiva, el coeficiente de correlación de Spearman resultante es 0.9 y 0.7 para la base TID2008 y LIVE, respectivamente; mientras que el coeficiente de correlación de Kendall es 0.7 y 0.5 para la base TID2008 y LIVE, respectivamente.

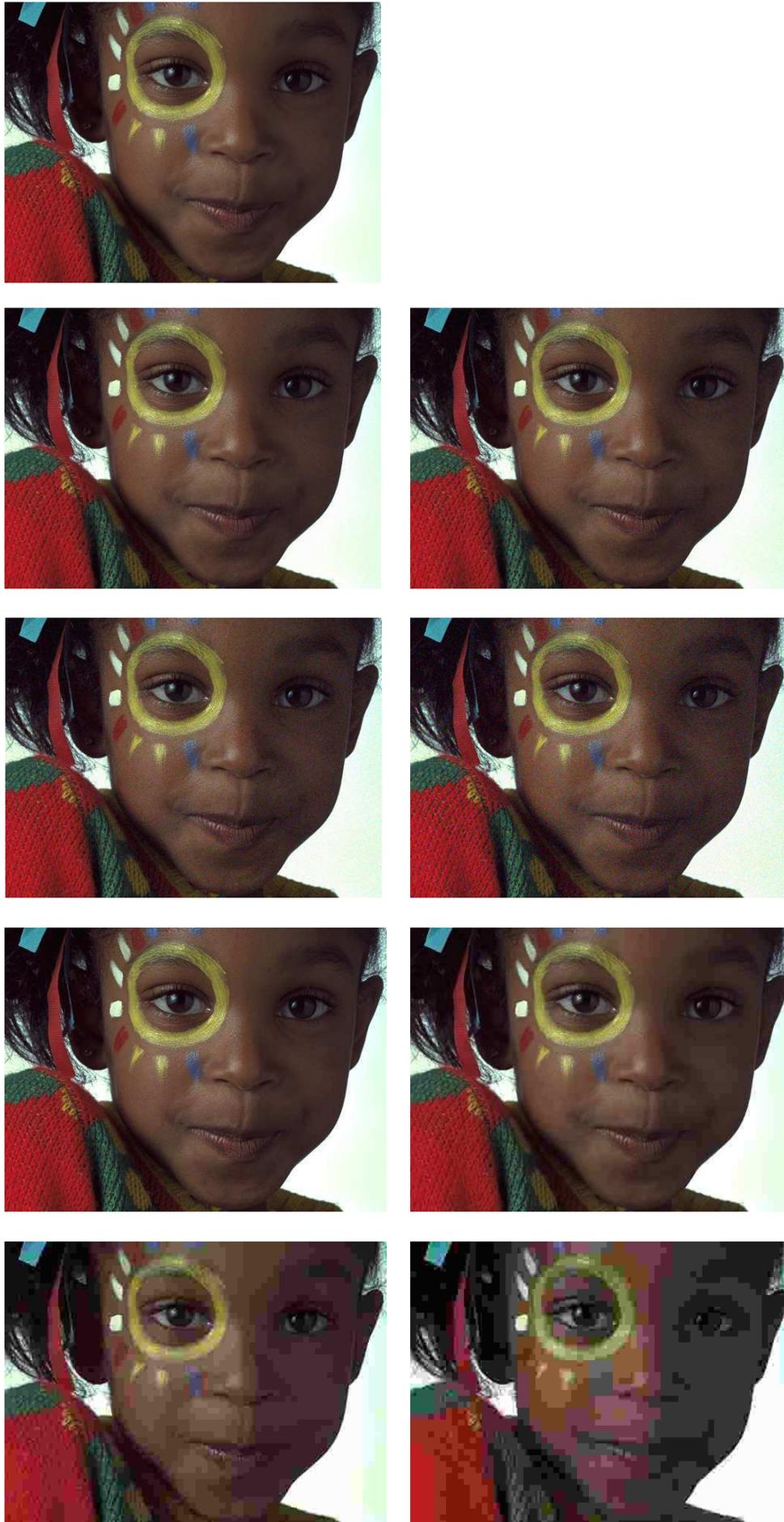


Figura 4.2: Ejemplo de la base de imágenes TID2008. Primera fila: imagen original, segunda y tercera fila: imagen distorsionada con AWGN, tercer y cuarta fila: imagen distorsionada por compresión JPEG en diferentes niveles.

4.2. Medidas de calidad en fusión de imágenes

Esta sección se proporciona una descripción general de las medidas basadas en la similitud estructural en el contexto de fusión de imágenes.

4.2.1. Medidas de Piella (Q_S , Q_W y Q_E)

Piella y Heijmans (2003) [56] propusieron tres medidas de calidad de fusión basadas en el índice Q . En estas medidas, $Q(x, f|w)$ y $Q(y, f|w)$ se calculan entre las imágenes de origen (x e y) y la imagen fusionada (f) en una ventana deslizante w .

La primera medida es la siguiente:

$$Q_S(x, y, f) = \frac{1}{|W|} \sum_{w \in W} [\lambda(w) Q(x, f|w) + (1 - \lambda(w)) Q(y, f|w)], \quad (4.20)$$

donde el peso $\lambda(w)$ se define como

$$\lambda(w) = \frac{s(x|w)}{s(x|w) + s(y|w)}. \quad (4.21)$$

El $s(x|w)$ y $s(y|w)$ son las saliencias locales de las dos imágenes de entrada x e y dentro de la ventana w , respectivamente. Reflejan la relevancia local de la imagen de origen dentro de la ventana w , y puede depender del contraste, la nitidez o la entropía. En la implementación de Piella, $s(\cdot|w)$ es la varianza de la imagen dentro de la ventana w .

El segundo, *índice de calidad de fusión ponderado*, le da más peso a aquellas ventanas donde la saliencia de las imágenes de entrada es mayor. Estos corresponden a áreas que probablemente sean partes perceptualmente importantes de la escena subyacente. Se define como:

$$Q_W(x, y, f) = \sum_{w \in W} c(w) [\lambda(w) Q(x, f|w) + (1 - \lambda(w)) Q(y, f|w)], \quad (4.22)$$

donde $\lambda(w)$ se calcula como en (4.21) y el coeficiente $c(w)$ se calcula como

$$c(w) = \frac{C(w)}{\sum_{w' \in W} C(w')}, \quad (4.23)$$

con

$$C(w) = \max\{s(x|w), s(y|w)\}, \quad (4.24)$$

En la implementación de Piella, $C(w)$ se define como en (4.24). También sugirió otra forma de calcularlo, como $C(w) = s(x|w) + s(y|w)$.

Observación 4.2.1 Si el denominador de c en (4.23) es cero, se establece $c = 1/|W|$, por lo que en este caso: $Q_W = Q_S$.

Finalmente, presentaron el *índice de calidad de fusión dependiente del borde* [56]. Esta medida tiene en cuenta algunas características del sistema visual humano, como la información de borde. Se obtiene de la siguiente manera:

$$Q_{E_1}(x, y, f) = Q_W(x, y, f) \cdot Q_W(x', y', f')^\alpha, \quad (4.25)$$

Piella sugirió otra variante de esta métrica en [54], definida de la siguiente manera:

$$Q_{E_2}(x, y, f) = Q_W(x, y, f)^{1-\alpha} \cdot Q_W(x', y', f')^\alpha, \quad (4.26)$$

En (4.25) y (4.26), $Q_W(x', y', f')$ es el Q_W calculado con las imágenes de “borde” x' , y' y f' , y α es un parámetro que pondera la información de contribución de borde, $\alpha \in [0, 1]$. Cuanto más cerca esté α de uno, más importante es la imagen de borde.

4.2.2. Medida de Cvejic (Q_C)

Cvejic et al. definió una medida de rendimiento de fusión [19] que tiene en cuenta la similitud local entre la imagen de origen y la imagen fusionada dentro de la misma posición espacial, de la siguiente manera:

$$Q_C = \sum_{w \in W} \text{sim}(x, y, f|w) \cdot Q(x, f|w) + (1 - \text{sim}(x, y, f|w)) \cdot Q(y, f|w), \quad (4.27)$$

con

$$\text{sim}(x, y, f|w) = \begin{cases} 0, & \text{if } \frac{s_{xf}}{s_{xf} + s_{yf}} < 0, \\ \frac{s_{xf}}{s_{xf} + s_{yf}}, & \text{if } 0 \leq \frac{s_{xf}}{s_{xf} + s_{yf}} \leq 1, \\ 1, & \text{if } \frac{s_{xf}}{s_{xf} + s_{yf}} > 1. \end{cases} \quad (4.28)$$

Este factor de ponderación depende de la similitud en el dominio espacial entre la fuente y la imagen fusionada. Cuanto mayor sea la similitud entre la imagen de origen y la imagen fusionada, mayor será el factor de ponderación correspondiente.

4.2.3. Medida de Yang (Q_Y)

Yang et al. propuso otra forma de usar SSIM para la evaluación de fusión [82]. En esta medida, la similitud estructural local entre las imágenes de origen $\text{SSIM}(x, y|w)$ se utiliza como

medida de coincidencia. Para aquellas ventanas cuyo $SSIM(x, y|w)$ es igual o mayor que un umbral dado (para distinguir información redundante), el promedio ponderado de $SSIM(x, f|w)$ y $SSIM(y, f|w)$ se toma, de lo contrario, se toma el más grande de los dos.

$$Q_Y(x, y, f|w) = \begin{cases} \lambda(w) SSIM(x, f|w) + (1 - \lambda(w)) SSIM(y, f|w), \\ \quad \text{if } SSIM(x, y|w) \geq 0,75, \\ \max\{SSIM(x, f|w), SSIM(y, f|w)\}, \\ \quad \text{if } SSIM(x, y|w) < 0,75. \end{cases} \quad (4.29)$$

El peso local $\lambda(w)$ está definido en (4.21).

La medida de calidad de fusión, Q_Y , se calcula promediando todos los valores de $Q_Y(x, y, f|w)$ sobre toda la imagen,

$$Q_Y(x, y, f) = \sum_{w \in W} \frac{Q_Y(x, y, f|w)}{|W|}. \quad (4.30)$$

Tener en cuenta que, todas estas medidas de calidad de fusión tienen un rango dinámico de $[-1, 1]$. Cuanto más cercano sea el valor a 1, mayor será la calidad de la imagen compuesta.

En este ejemplo [58], se puede comparar el rendimiento de las medidas de fusión anteriormente presentadas; donde se utiliza un par de imágenes de enfoque múltiple: Reloj, ver Figura 4.3, y se consideran los métodos de fusión clásicos en la literatura: Pirámide laplaciana (Laplacian Pyramid - LP) [83], Pirámide de razón (Ratio Pyramid - RP)[83], Transformada de wavelet discreta (Discrete Wavelet Transform - DWT) [41] y DWT invariante a traslaciones (Shift Invariant DWT - SIDWT) [60], implementado en “MATLAB Image Fusion Toolbox” [61] (Ver Figura 4.4).

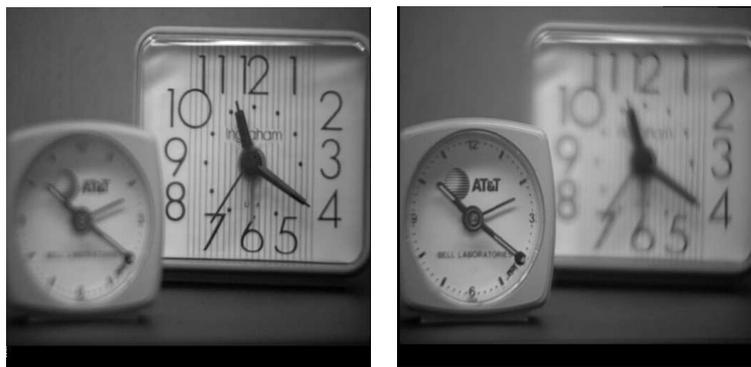


Figura 4.3: Imagen Reloj (foco a derecha y foco a izquierda).

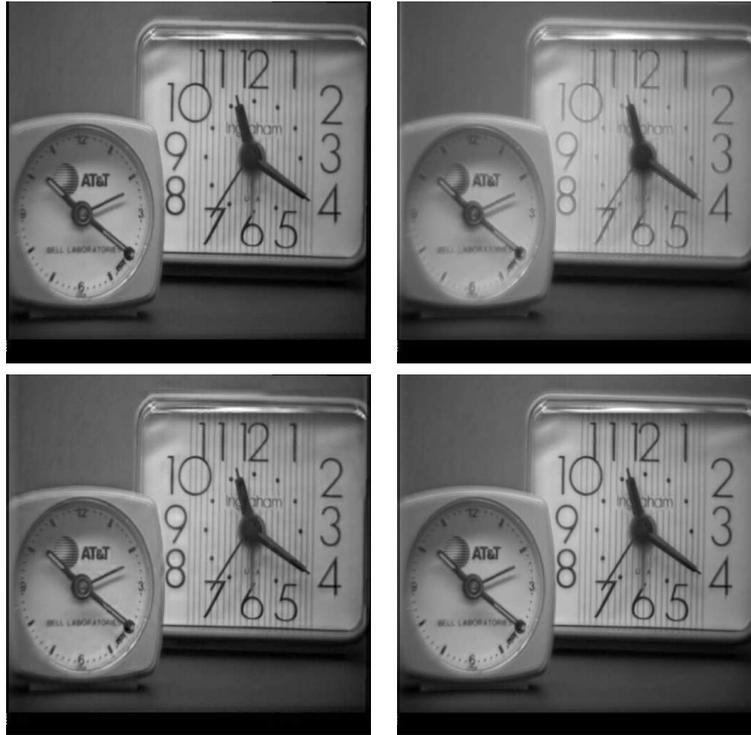


Figura 4.4: Imágenes fusionada obtenida mediante los métodos de fusión: LP, RP, DWT y SIDWT, respectivamente.

Los resultados se muestran en la siguiente Tabla 4.1 junto con los mapas de evaluación de calidad que se muestran en la Figura 4.5.

Tabla 4.1: Evaluaciones objetivas de diferentes medidas de calidad de fusión de imágenes para una imagen de reloj fusionada.

Medidas	Métodos			
	LP	RP	DWT	SIDWT
Q_S	0.8220	0.7440	0.7824	0.8366
Q_W	0.9272	0.7878	0.9139	0.9217
Q_{E_1}	0.8188	0.4774	0.8025	0.8027
Q_{E_2}	0.9048	0.6910	0.8958	0.8959
Q_C	0.8284	0.7564	0.7919	0.8368
Q_Y	0.8816	0.7879	0.8461	0.8853

4.2.4. Versión extendida

A diferencia de las medidas propuestas por Piella, las medidas de Yang y Cvejic no poseen la versión para múltiples imágenes de entrada. En 2009, Piella presentó una versión extendida de la medida ponderada de calidad de fusión Q_W [55] para medir el rendimiento de un proceso

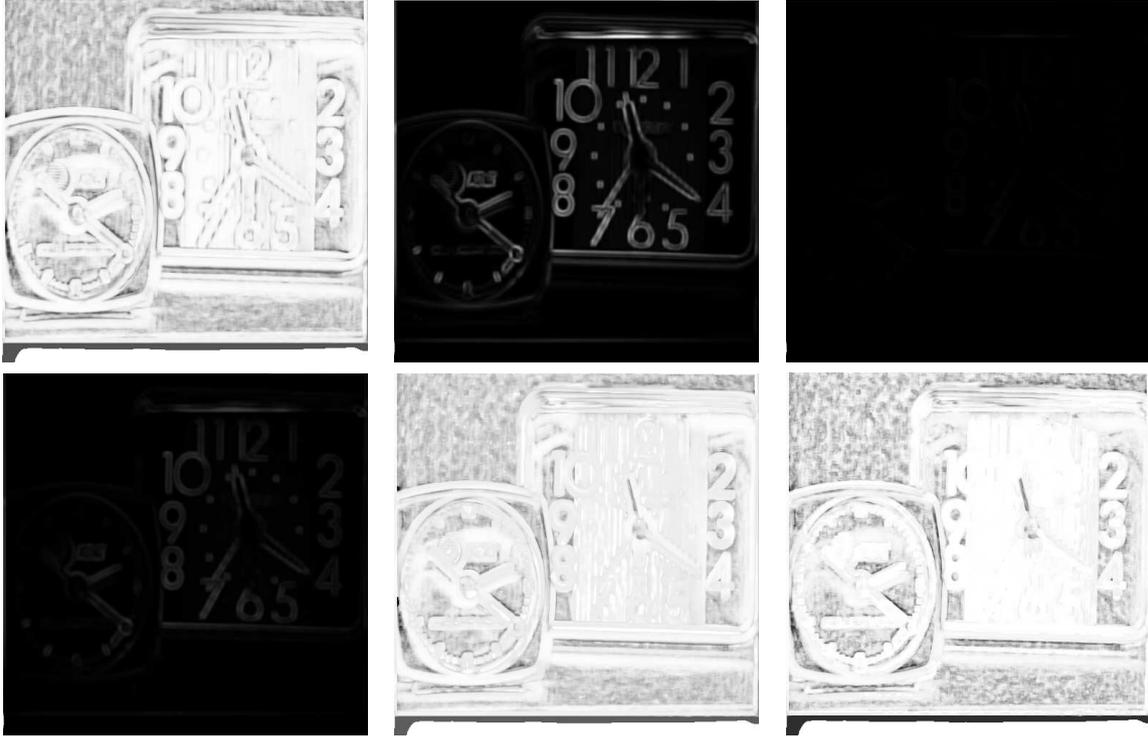


Figura 4.5: Mapas de calidad de las medidas de fusión: Q_S , Q_W , Q_{E_1} , Q_{E_2} , Q_C y Q_Y de la imagen fusionada obtenida por el método de fusión LP, respectivamente. El brillo indica la magnitud de la métrica de fusión local.

de fusión que considera múltiples imágenes de entrada. Se definió como

$$Q_W(I, f) = \sum_{n=1}^K \sum_{w \in W} c(w) \lambda_n(w) \text{SSIM}(I_n, f|w), \quad (4.31)$$

donde $I = (I_1, \dots, I_K)$ es la imagen de valores múltiples, cuyos componentes $I_n, n = 1, \dots, K$ son las imágenes de entrada, f es la imagen fusionada, $\lambda_n(w) = \frac{s(I_n|w)}{\sum_{n=1}^K s(I_n|w)}$ el peso local, con $s(I_n|w)$ la saliencia local, $c(w) = \frac{C(w)}{\sum_{w' \in W} C(w')}$, con $C(w) = \max\{s(I_1|w), \dots, s(I_K|w)\}$. A partir de esta, se deduce Q_S

$$Q_S(I, f) = \frac{1}{|W|} \sum_{n=1}^K \sum_{w \in W} c(w) \lambda_n(w) \text{SSIM}(I_n, f|w),$$

donde $|W|$ es el cardinal de W .

Luego, Q_E se define considerando las expresiones 4.25 o 4.26.

La primera aplicación de esta versión fue realizada en [8], en la implementación de un algoritmo cuyo objetivo era determinar la fusión de imágenes de enfoque múltiple a partir de varias imágenes policromáticas capturadas para microscopios automatizados.

Capítulo 5

Evaluación de la calidad de similitud estructural a escala múltiple

En este capítulo se propone y testea una nueva medida de calidad de imágenes. Ver Martinez et al. 2016 [48].

5.1. Índice de calidad de similitud estructural multiescala (MS-SIM)

La expresión del SSIM en (4.5), no tiene en cuenta la distancia de visión del observador. Una forma sencilla de considerar la distancia visual es ajustar la resolución de la imagen para que el ángulo de visión por píxel permanezca constante. Es decir, si la distancia de visualización se duplica, la resolución debería reducirse a la mitad. Por este motivo, en 2003, Wang, Simoncelli y Bovik [78] estudian una alternativa al índice de calidad SSIM, basado en el análisis multiescala. La descomposición a múltiples escalas es una forma apropiada de incorporar detalles de la imagen en diferentes resoluciones [78] y proporciona una representación conveniente de las estructuras geométricas locales [55].

El sistema aplica iterativamente un filtro de paso bajo seguido de un muestreo descendente de las imágenes filtradas por un factor de 2. Indexamos las imágenes de entrada como escala 1, y la escala más alta como escala L , obtenida después de $L - 1$ iteraciones. En una escala dada l , el sistema calcula el contraste s y los factores de correlación lineal c del índice SSIM en el sistema de parche W . El factor de luminancia solo se calcula en la escala L . La puntuación final

del parche se realiza combinando la información intraescala de la siguiente manera:

$$\text{MSSIM}(x, y) = [l(x_L, y_L)]^{\alpha_L} \prod_{j=1}^L [c(x_j, y_j)]^{\gamma_j} \cdot [s(x_j, y_j)]^{\beta_j}, \quad (5.1)$$

donde c mide el grado de correlación lineal entre x e y , l mide la similitud entre la luminancia de x e y , y s mide la similitud relacionada con el contraste entre imágenes de parche (4.4). Con el fin de evaluar la calidad general de la imagen, los índices locales MSSIM se promedian utilizando un enfoque de ventana deslizante.

Wang, Simoncelli y Bovic (2003) [78] realizaron un estudio subjetivo para determinar la importancia relativa de la información recopilada de cada escala. Los parámetros obtenidos donde $\beta_1 = \gamma_1 = 0,0448$, $\beta_2 = \gamma_2 = 0,2856$, $\beta_3 = \gamma_3 = 0,3001$, $\beta_4 = \gamma_4 = 0,2363$ y $\alpha_5 = \beta_5 = \gamma_5 = 0,1333$ para cinco escalas.

5.2. Medida de calidad propuesta (MS- Q_W)

Ma et al.(2015) [45] consideró una extensión de MSSIM al problema de evaluación de la calidad en el contexto de fusión de imágenes, para el caso específico de datos MEF (en inglés multiple exposure fusion): de múltiples imágenes y múltiples exposiciones. En ese caso, la luminancia no era una característica muy relevante, ya que podía introducir ruido en la imagen fusionada, mientras que la determinación de la estructura es de suma importancia. La estructura del parche de la imagen fusionada debe representar la estructura de todos los parches de entrada. Con esta idea en mente, definieron una imagen estructural y la compararon con la imagen fusionada en cada etapa del sistema multiescala MSSIM. La medida final promedia los puntajes calculados dentro de cada escala y multiplica estos valores medios entre escalas. La correlación con la opinión subjetiva se reporta alta para esta medida dentro de su base de datos.

Hassen et al. (2015) [29] también consideraron la resolución múltiple para detectar cambios en la coherencia de fase local que, a su vez, se relacionan con el desenfoque y la nitidez reducida en la imagen fusionada. Definen una medida para evaluar la calidad de la imagen obtenida por la fusión de una colección de imágenes de múltiples exposiciones y múltiples enfoques. La medida es producto de tres factores, uno que evalúa el contraste, otro que evalúa la nitidez al estudiar la coherencia de fase local, y el tercero que evalúa la similitud estructural al estimar la correlación. La medida final es bastante compleja, pero también logró una alta correlación con las puntuaciones subjetivas proporcionadas en sus experimentos.

La medida propuesta pretende ser universal, ciega y más simple que estos intentos anteriores, sin comprometer la calidad final. El objetivo es proporcionar una medida que pueda incluirse

fácilmente en algoritmos que seleccionan el conjunto más pequeño de imágenes, o la cantidad de información, necesaria para lograr una calidad determinada. Por lo tanto, se propone una medida de fidelidad a múltiples escalas, MS- Q_W , para la evaluación de la calidad de imagen cuyo diagrama del sistema se ilustra en la Figura 5.1. Tomando la colección de imágenes I y la imagen fusionada f como entrada, el sistema aplica iterativamente un filtro paso bajo seguido de un muestreo descendente por un factor de 2. Se indexa las imágenes de entrada como escala 1, y la escala más alta en la descomposición como escala L , que se obtiene después de iteraciones $L - 1$. La medida de fusión general a escala múltiple se obtiene combinando la medida $Q_W(I, f)$ definida en (4.2.4) en diferentes escalas utilizando la siguiente ecuación:

$$\text{MS-}Q_W(I, f) = \prod_{l=1}^L [Q_W(I_l, f_l)]^{\beta_l}, \quad (5.2)$$

La comparación de la estructura local se aplica utilizando el enfoque de ventanas deslizantes

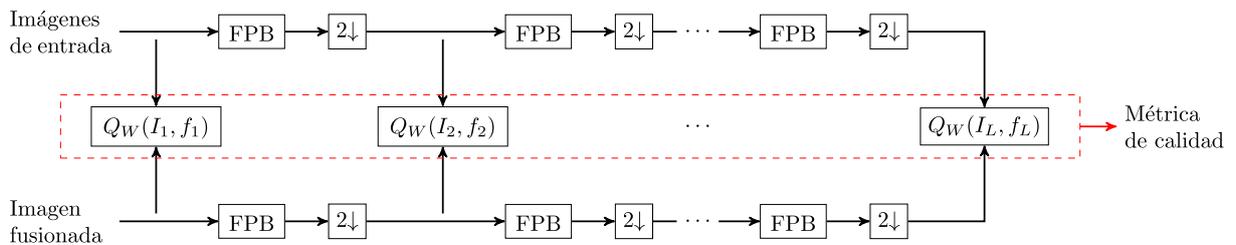


Figura 5.1: Sistema de medida de fusión de imagen a escala múltiple. FPB: Filtro paso bajo; 2 ↓: Muestreo descendente por un factor de 2. Nuestra propuesta se obtiene aplicando Q_W en cada escala, con $L = 5$.

en toda la imagen y las imágenes escaladas de la descomposición de resolución múltiple, lo que resulta en los mapas de calidad que se muestran en la Figura 5.2, donde el brillo indica una mejor calidad. Estos mapas indican cómo se conservan los detalles estructurales en cada ubicación espacial, en cada escala. La figura retrata ejemplos de dos conjuntos clásicos de dos imágenes de origen de enfoque múltiple, “Pepsi” y “Disk”, y su imagen fusionada correspondiente creada por el SIDWT. Se sabe que los mapas de calidad predicen la preservación de la estructura local o las distorsiones y la pérdida de detalles [45]. La Figura 5.2 se puede observar que Q_W es capaz de separar la información de estructura de ambas imágenes de entrada, errores de detalles finos en el código de barras de la segunda imagen de entrada (foco a derecha) en la escala $l = 1$ y detalles de estructura general de la lata y la etiqueta en la escala $l = 3$. Lo mismo ocurre con la imagen del disco, la pérdida de detalles se refleja bien en los mapas de calidad de Q_W . Esta propiedad se vuelve más importante cuando se fusionan muchas imágenes de origen con diferentes detalles, particularmente cuando una de las imágenes de origen tiene un defecto grave. La imagen fusionada heredará dicho defecto, y al ser comparada con las imágenes



Figura 5.2: Imágenes de entrada multi-foco: “Pepsi” y “Disk” (foco a izquierda y foco a derecha) y la imagen fusionada obtenida por el método SIDWT junto a sus respectivos mapas de calidad para $L = 1, 3$ y 5. El brillo indica la magnitud de $MS-Q_W$.

fuentes no defectuosas, debe tener una puntuación baja. El conjunto que produce la mejor imagen fusionada en términos de calidad $MS-Q_W$ es el que no tiene la imagen de entrada defectuosa. Exploramos esta idea más a fondo en la sección de aplicaciones 5.6.

5.3. Validación

Para validar el sistema de calidad de fusión de imágenes propuesto, se llevaron a cabo tres experimentos. Cada uno de ellos, aborda una característica relevante en la medida de calidad de fusión MS- Q_W ($L = 5$):

- propósito general,
- comportamiento estable y capacidad discriminativa,
- correlación con la opinión subjetiva.

Se compara, además, su rendimiento en los niveles de descomposición: $L = 1$ y 3 . Notar que si el nivel de descomposición L se establece en 1 , esta medida se convierte en la medida extendida de Piella Q_W .

Todos los parámetros se heredan de publicaciones anteriores para permitir una comparación adecuada: tamaño de ventana de 8×8 píxeles, $C_1 = C_2 = 0$, $s(\cdot|w)$ variancia de la imagen dentro de la ventana w , número de escalas: $L = 1, 3$ y 5 y pesos de escala dados por $\beta_1 = 0.0448$, $\beta_2 = 0.2856$, $\beta_3 = 0.3001$, $\beta_4 = 0.2363$ y $\beta_5 = 0.1333$ [78].

A continuación, se reportan los detalles de los experimentos realizados:

5.3.1. Primer experimento: evaluación de propósito general

En el primer experimento, se muestra el rendimiento de la medida propuesta, MS- Q , como herramienta de propósito general. Para el estudio se consideran cuatro escenarios distintos de aplicación de imágenes clásicas, compuestos de sólo dos imágenes de entrada, utilizados ampliamente en la literatura, [55, 72, 31, 53]. Estos son: “Disk” (dos imágenes de enfoque múltiple, por ejemplo, cámara de visión - 512×512), “Medical” (una tomografía computarizada y una imagen de resonancia magnética - 256×256), “Remote” (par de bandas de una imagen de detección remota multibanda - 512×512), “Tank” (una imagen infrarroja y una imagen visual, utilizada para tareas de vigilancia - 354×472).

En las Figuras 5.3, 5.4, 5.5 y 5.6 se exhiben los conjuntos de imágenes con sus contrapartes de fusión. Los algoritmos clásicos considerados son: Pirámide laplaciana (LP), Pirámide de razón (RP), Transformación de wavelet discreta (DWT) y DWT invariante de cambio (SIDWT), implementado en la “MATLAB Image Fusion Toolbox”, proporcionada por Rockinger [61], donde se pueden encontrar detalles de estos métodos y referencias. Sus desempeños fueron probados subjetivamente y aceptados en la literatura, [56, 55]. Al ser todos los algoritmos de escala múltiple, se realizó una descomposición de tres niveles, con los coeficientes de aproximación de

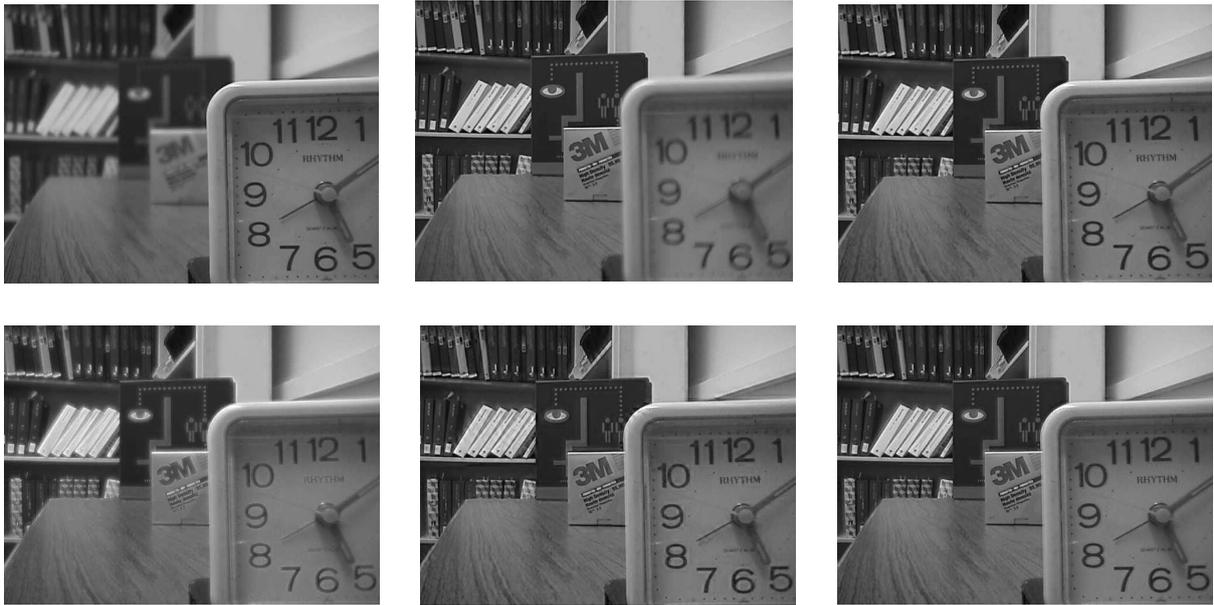


Figura 5.3: Imágenes de entrada: “Disk” e imágenes fusionadas obtenidas por los métodos de fusión: LP, RP, DWT y SIDWT, respectivamente

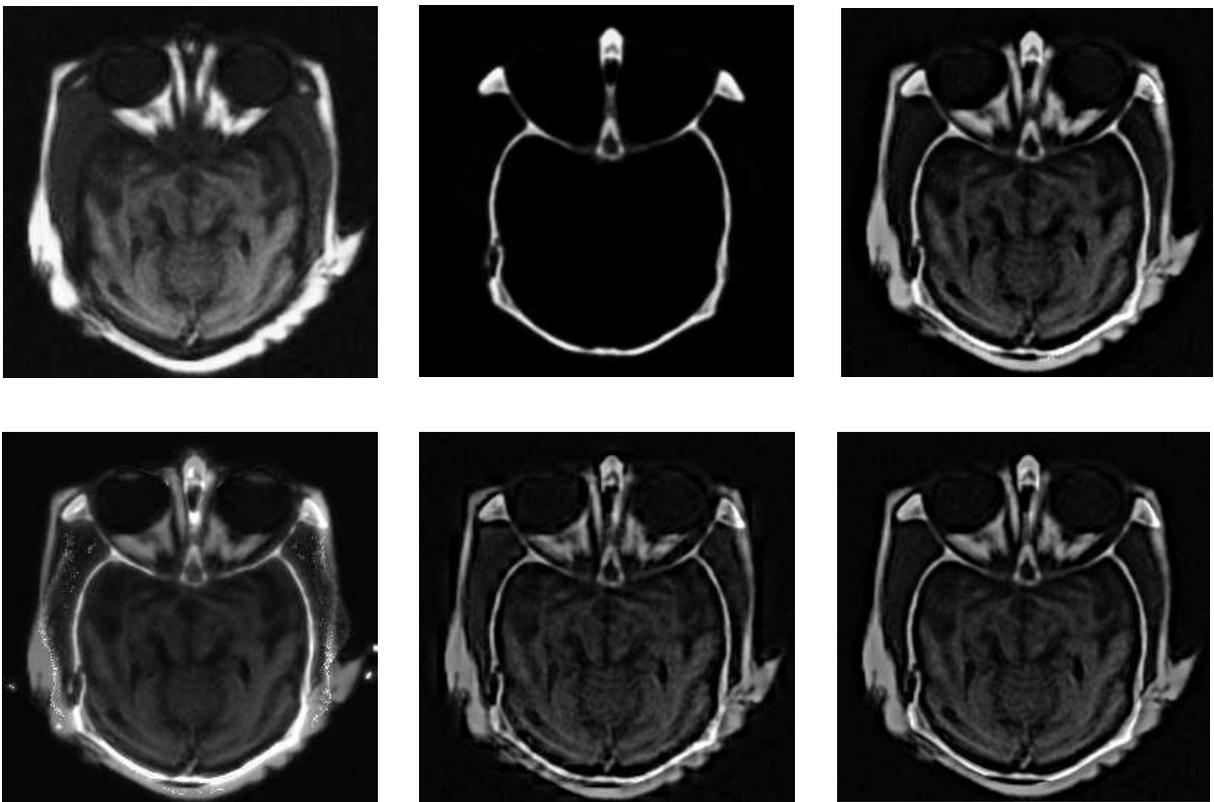


Figura 5.4: Imágenes de entrada: “Medical” e imágenes fusionadas obtenidas por los métodos de fusión: LP, RP, DWT y SIDWT, respectivamente

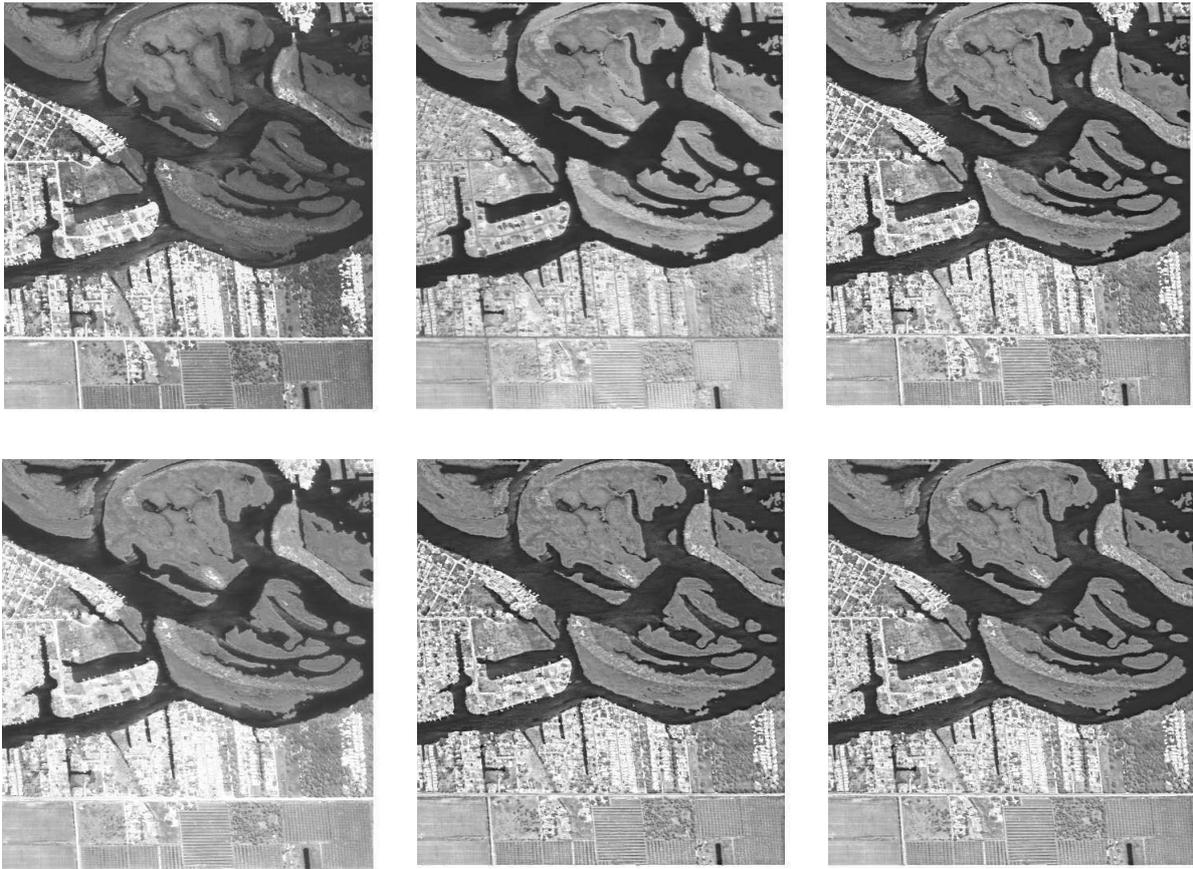


Figura 5.5: Imágenes de entrada: “Remote” e imágenes fusionadas obtenidas por los métodos de fusión: LP, RP, DWT y SIDWT, respectivamente

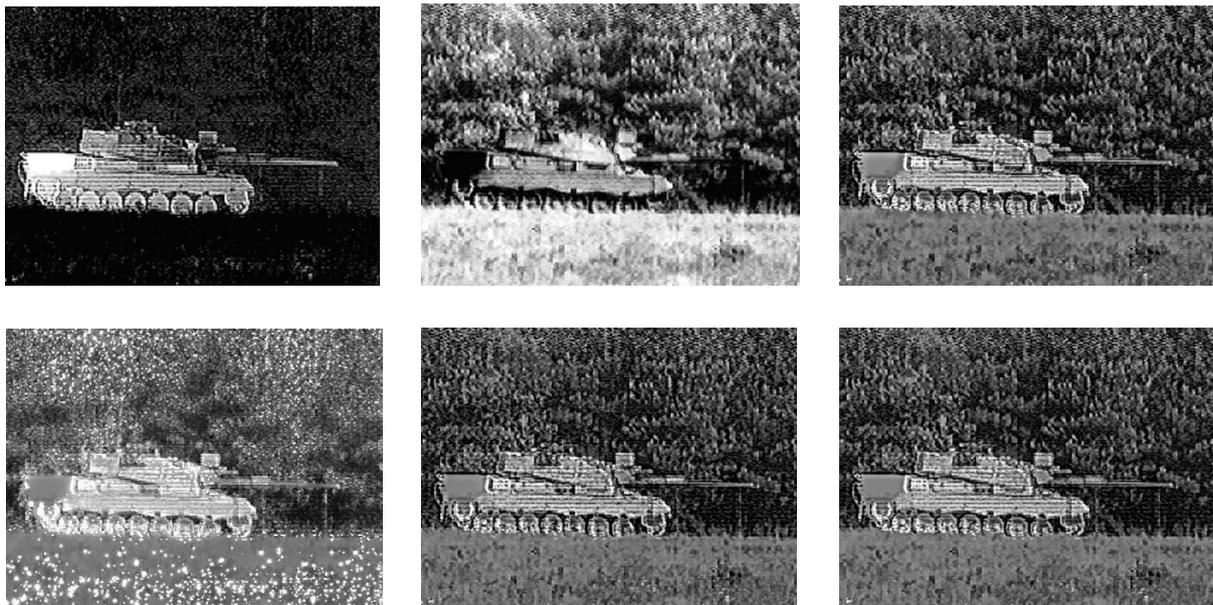


Figura 5.6: Imágenes de entrada: “Tank” e imágenes fusionadas obtenidas por los métodos de fusión: LP, RP, DWT y SIDWT, respectivamente

las dos imágenes de entrada promediadas y los valores absolutos más grandes de la subbanda seleccionada más alta.

5.3.2. Segundo experimento: estabilidad y capacidad discriminativa de la medida

En este experimento, se emplea la base de datos “TNO UN Camp”, constituida por 32 pares de imágenes de entrada de vigilancia multimodal (infrarrojas y visuales - 360×270). Estas imágenes se fusionaron con los mismos métodos de fusión utilizados en el experimento anterior. Ver Figura 5.7.

El segundo experimento persigue dos propósitos:

- estabilidad de la medida: mediante el cálculo de la varianza, se analiza la estabilidad de la medida de calidad de fusión propuesta considerando diferentes imágenes de entrada, en un mismo contexto. Un valor bajo de varianza, indica una buena estabilidad de la medida para un método de fusión específico [44].
- capacidad discriminativa de la medida para la determinación del algoritmo de fusión con mejor performance: Liu et al. (2017) [43] recomiendan el uso de análisis estadísticos no paramétricos para la comparación de los algoritmos de fusión. Por tal motivo, se aplica la prueba no paramétrica de Friedman [23] que permite comparar la calidad de las imágenes fusionadas por los métodos de fusión bajo estudio. Este test establece como hipótesis nula que la calidad de las imágenes fusionadas obtenidas por los distintos métodos de fusión es la misma contra la hipótesis alternativa de que por lo menos una no lo es. Se supone que los datos son muestras independientes de distribuciones continuas que son idénticas. El diseño de esta experiencia se podría asimilar a un diseño en bloques completos aleatorizados. Cada par de imágenes a fusionar puede ser considerado como un bloque. La estrategia de bloqueo de cada par de imágenes de la base se realiza con el fin de distinguir la variación debida a las mismas, de la variación aleatoria o error experimental. A cada par de imágenes se le aplican 4 tratamientos, es decir, se fusionan con cada uno de los 4 métodos. Para realizar la prueba de Friedman se aplica la transformación rango a la variable en estudio, medida de calidad, dentro de cada bloque. Los procedimientos para realizar comparaciones múltiples adecuados para este caso, se basan en las medias de los rangos por tratamiento y en la varianza de los rangos según se describe en Conover (1999) [16].



Figura 5.7: Un ejemplo del conjunto de 32 pares de imágenes de: “TNO UN Camp” e imágenes fusionadas obtenidas por los métodos de fusión: LP, RP, DWT y SIDWT, respectivamente.

5.3.3. Tercer experimento: evaluación de acuerdo con la opinión subjetiva

La perspectiva subjetiva es ampliamente aceptada como el enfoque más preciso para medir la calidad, ya que el ojo humano es el receptor final de la mayoría, sino de todos los sistemas de comunicación visual. Para ello, el tercer experimento estudia el grado de acuerdo entre la medida propuesta con evaluaciones subjetivas. En el mismo, se utilizó la base de datos MEF Waterloo IVC 2015 desarrollada por Ma et al. [45] con motivo de realizar una evaluación comparativa. Consiste en diecisiete secuencias de imágenes de entrada de alta calidad que cubren diversos escenarios: vistas interiores y exteriores, paisajes naturales y construcciones hechas por el hombre. Estas son: “Ballons” ($339 \times 512 \times 9$), “Belgium house” ($384 \times 512 \times 9$), “Lamp1” ($384 \times 512 \times 15$), “Candle” ($364 \times 512 \times 10$), “Cave” ($384 \times 512 \times 4$), “Chinese garden” ($340 \times 512 \times 3$), “Farm house” ($341 \times 512 \times 3$), “House” ($340 \times 512 \times 4$), “Kluki” ($341 \times 512 \times 3$), “Lamp2” ($342 \times 512 \times 6$), “Landscape” ($441 \times 512 \times 3$), “Light house” ($340 \times 512 \times 3$), “Madison capitol” ($384 \times 512 \times 30$), “Memorial” ($512 \times 381 \times 16$), “Office” ($340 \times 512 \times 6$), “Tower” ($512 \times 341 \times 3$) y “Venice” ($341 \times 512 \times 3$); ver Figura 5.8. Todas las secuencias de imágenes de entrada contienen al menos 3 imágenes, con diferentes niveles de exposición. También incluye puntuaciones subjetivas que evalúan la calidad de las salidas de ocho algoritmos MEF: Mertens07, Li12, ShutaoLi12, Li13, Global Energy Weighted, Gu12, Raman09 y Local Energy Weighted ¹. En la Figura 5.9 se muestra un ejemplo de secuencia de imágenes de entrada de esta base y en la Figura 5.10, sus correspondientes imágenes fusionadas obtenidas por los distintos métodos mencionados anteriormente.

La evaluación subjetiva de la calidad de las imágenes fusionadas se desarrolló en el interior

¹Se utiliza la notación de los algoritmos MEF que aparece en [45].

de una oficina, con un nivel de iluminación normal, sin cielorraso ni suelos reflectantes. Todas las imágenes se muestran en un monitor LCD, con una resolución de 2560×1600 píxeles con Truecolor (32 bits) a 60 Hz. El monitor se calibró de acuerdo con las recomendaciones de ITU-T BT.500 [21]. Para la presentación de las imágenes en la pantalla, se utilizó una ventana de figura de MATLAB. En la evaluación se muestran al sujeto, las 8 imágenes fusionadas de una misma secuencia simultáneamente. Estas son ubicadas espacialmente en orden aleatorio. El estudio adoptó una estrategia de puntuación de calidad de múltiples estímulos, sin mostrar la secuencia de referencia. Participaron en el experimento, un total de 25 observadores novatos, 15 hombres y 10 mujeres de entre 22 y 30 años. Los sujetos pueden acercarse o alejarse de la pantalla para mejorar su visualización. Las calificaciones de los sujetos se registraron con lápiz y papel. La duración de una sesión se limitó a un máximo de 30 minutos, para minimizar la influencia de la fatiga visual. Para cada conjunto de imágenes fusionadas, el sujeto debe otorgar una puntuación entera que refleje mejor la calidad de percepción de cada una de ellas. La puntuación varía de 1 a 10, donde 1 denota la peor calidad y 10 la mejor. Después del experimento dos sujetos fueron considerados outliers y se removieron del estudio, con lo cual solo se consideraron las respuestas de 23 observadores.

En el análisis del nivel de acuerdo entre la evaluaciones subjetivas y los puntajes obtenidos por las medidas de calidad de fusión de imágenes, se utilizaron dos criterios ampliamente aceptados, el coeficiente de correlación de los rangos de Spearman (CC de Spearman)² y el coeficiente de correlación de los rangos de Kendall (CC de Kendall)³. Ambos criterios oscilan entre -1 a 1. Estos números representan un perfecto desacuerdo o acuerdo entre la medida de calidad y la opinión subjetiva, respectivamente. A continuación se detallan las características principales de cada una de ellos.

5.4. Resultados experimentales

5.4.1. Resultados del primer experimento: evaluación de propósito general

Los resultados de la evaluación cuantitativa correspondientes se reportan en la Tabla 5.1. En esta prueba, las medidas de escalas múltiples asignan los valores más altos a LP y SIDWT, seguidos de DWT, como se esperaría, según la literatura, ver [55, 67] para más detalles. Las peores puntuaciones corresponden al método RP. Este comportamiento se mantiene en todas las escalas. En general, los resultados también son consistentes con las comparaciones visuales

²El coeficiente de correlación de Spearman se define como el coeficiente de correlación de Pearson entre las variables de rango.

³Sea $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ un conjunto de observaciones de las variables aleatorias conjuntas x e y respectivamente, se define el coeficiente de correlación de Kendall como: $\tau = \frac{2}{n(n-1)} \sum_{i < j} \text{sgn}(x_i - x_j) \text{sgn}(y_i - y_j)$.

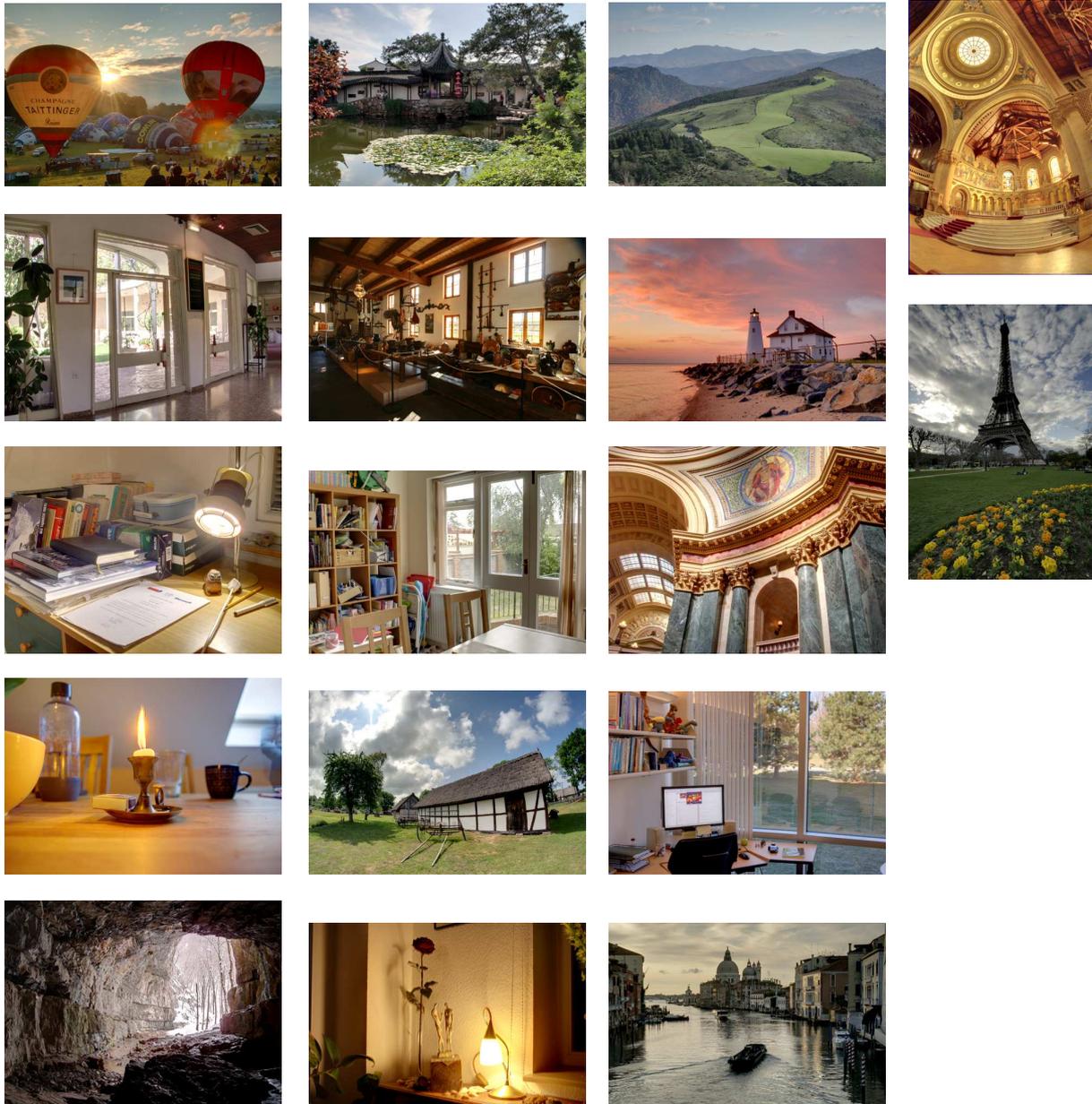


Figura 5.8: Base de imágenes MEF Waterloo IVC 2015: “Ballons”, “Belgium house”, “Lamp1”, “Candle”, “Cave”, “Chinese garden”, “Farm house”, “House”, “Kluki”, “Lamp2”, “Landscape”, “Light house”, “Madison capitol”, “Office, Venice”, “Memorialz “Tower”.

subjetivas, ver Figuras 5.3, 5.4, 5.5 y 5.6.

5.4.2. Resultados del segundo experimento: estabilidad y capacidad discriminativa de la medida

Estabilidad

De acuerdo a los valores de varianza reportados en la Tabla 5.2, se puede observar diferencias de comportamiento al cambiar el nivel de descomposición de la medida propuesta. La Figura 5.11 (primer panel) muestra los valores numéricos informados en la Tabla 5.2 para una revisión rápida.

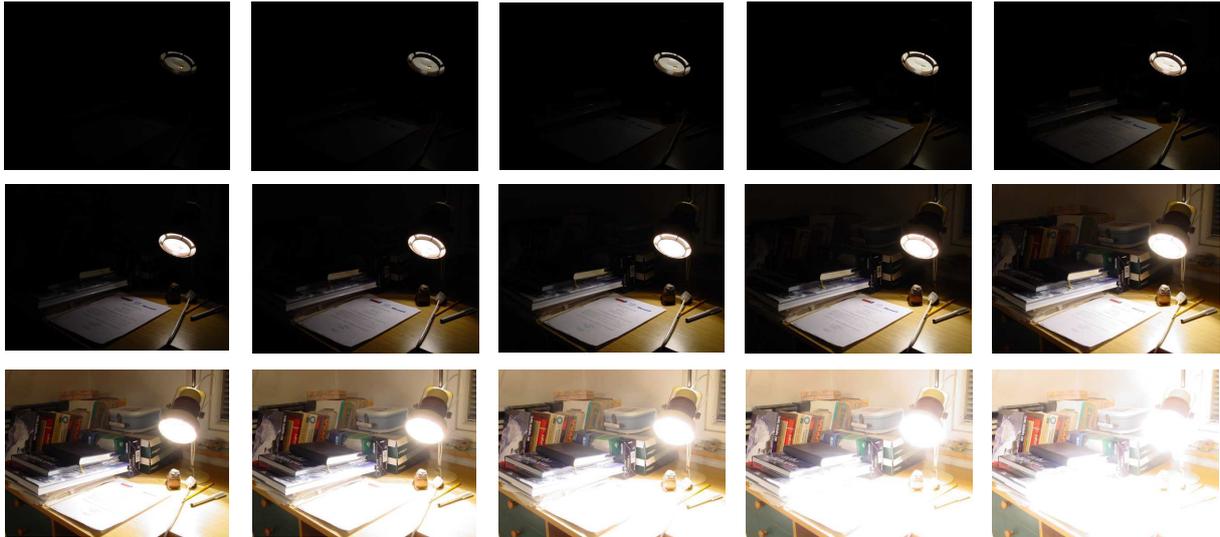


Figura 5.9: Un ejemplo de las 17 secuencias de imágenes de alta calidad (subexpuestas, sobreexpuestas y entre casos) de la base de datos MEF Waterloo IVC 2015: "Lamp1".

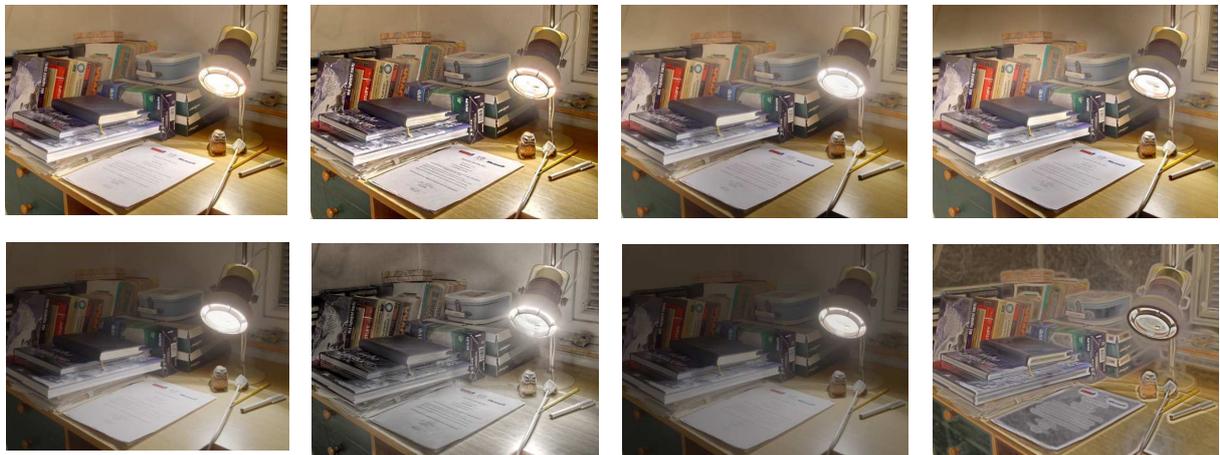


Figura 5.10: Imágenes fusionadas obtenidas por diferentes algoritmos MEF, ordenadas según los rangos subjetivos medios (MSR): Mertens07, Li12, ShutaoLi12, Li13, Global Energy Weighted, Gu12, Raman09 y Local Energy Weighted, respectivamente.

El valor más bajo de variancia de $MS-Q_W$ se alcanza con el nivel más alto de descomposición $L = 5$. Este mismo hecho puede apreciarse cuando se comparan los paneles restantes de la Figura 5.11, donde las medidas se discriminan por cada uno de los pares de imágenes que componen la base "TNO UN Camp".

Capacidad discriminativa

Si se observa la Figura 5.11, las curvas de las medidas para LP, son ligeramente más altas que las curvas de SIDWT, para todos los niveles de descomposición. Si se considera el puntaje máximo de las medidas de calidad, LP es el mejor método de fusión para esta base, independientemente del nivel de descomposición, y en segundo lugar, pero muy próximo, SIDWT. Por lo tanto, es importante estudiar si esta diferencia es estadísticamente significativa y así, conocer la capacidad

Tabla 5.1: Evaluaciones objetivas de las imágenes fusionadas en diferentes escenarios de aplicación de imágenes. Las puntuaciones máximas por columna se resaltan en negrita

Imagen	Método de Fusión	MS- Q_W		
		$L = 1$ Piella's [56]	$L = 3$	$L = 5$
"Disk"	LP	0.9171	0.9301	0.9496
	RP	0.7943	0.8713	0.9030
	DWT	0.9036	0.9235	0.9470
	SIDWT	0.9203	0.9373	0.9558
"Medical"	LP	0.7889	0.7176	0.6771
	RP	0.6577	0.5763	0.5845
	DWT	0.7147	0.6456	0.6156
	SIDWT	0.7504	0.6744	0.6384
"Remote"	LP	0.8656	0.9061	0.9199
	RP	0.8137	0.8744	0.8948
	DWT	0.8505	0.8894	0.9080
	SIDWT	0.8725	0.9058	0.9196
"Tank"	LP	0.6481	0.6484	0.5999
	RP	0.2509	0.3146	0.3498
	DWT	0.6095	0.5832	0.5432
	SIDWT	0.6464	0.6255	0.5789

Tabla 5.2: Varianza de las medidas de calidad de fusión de la base de datos "TNO UN Camp" (32 imágenes). Los mejores resultados se destacan en negrita.

Método	MS- Q_W		
	$L = 1$ Piella's[56]	$L = 3$	$L = 5$
LP	3.8104E-05	5.3833E-05	2.8654E-05
RP	2.3001E-05	1.7657E-05	1.6518E-05
DWT	2.8163E-05	2.4889E-05	1.2909E-05
SIDWT	2.5976E-05	3.4924E-05	1.8982E-05

discriminativa de la medida de calidad propuesta MS- Q_W ($L = 5$).

A través de la prueba de Friedman se puede concluir que existen diferencias altamente significativas en la calidad de las imágenes fusionadas obtenidas por los distintos métodos de fusión bajo estudio ($p < 1E - 04$). Mediante las comparaciones post hoc se concluye que las calidades de las imágenes fusionadas mediante los métodos son significativamente diferentes entre sí ($p < 0.05$).

5.4.3. Resultados del tercer experimento: evaluación de acuerdo con la opinión subjetiva

Ma et. al (2015) [45] definió una medida específica para los datos de MEF Waterloo IVC 2015 y comparó su rendimiento con otras 9 medidas de calidad objetivas para fusión de imágenes de propósito general : métrica de preservación de información de borde (Edge information preservation metric - EIP) [81], métrica de entropía de Tsallis (Tsallis entropy metric-TE) [18], métrica

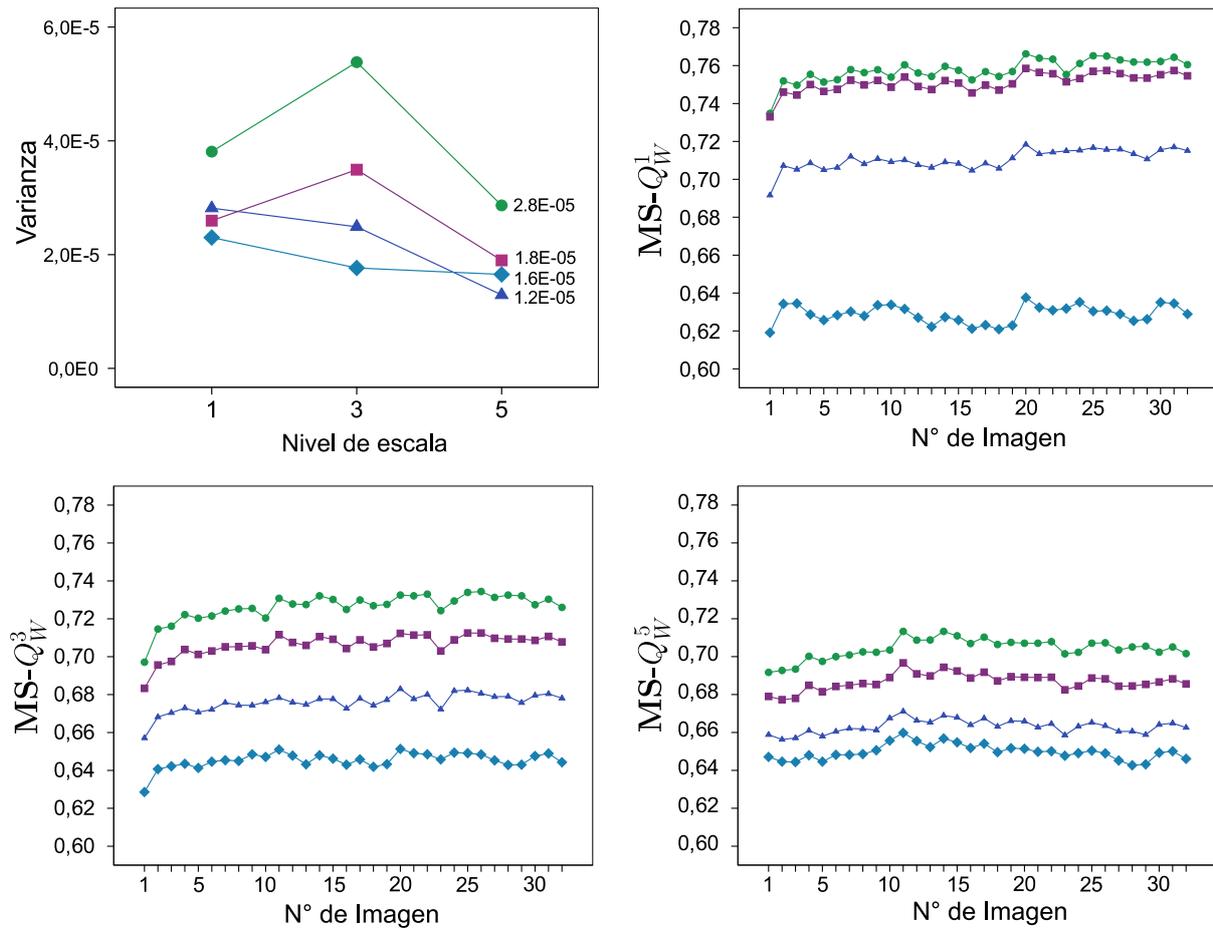


Figura 5.11: Gráfico de líneas de la varianza de $MS-Q_W$ en “TNO UN Camp”, según el nivel de escala de $MS-Q_W$; y performance de la métrica, calculada en imágenes fusionadas por LP (verde \circ), RP (cian oscuro \diamond), DWT (azul \triangle) y SIDWT (malva \square), respectivamente.

de resistencia de borde (Edge strength metric - ES) [13], métrica de frecuencia espacial (Spatial frequency metric - SF) [84], métrica de preservación de bordes de escala múltiple (Multi-scale edge preservation metric - MEP) [72], métrica de información mutua normalizada (Normalized mutual information metric - NMI) [30], métrica de correlación no lineal (nonlinear correlation metric - NC) [73], métrica de preservación de contraste ponderado por saliencia (Saliency weighted contrast preservation metric - SWCP) [14], incluidas las medidas de Piella a escala única que obtuvieron los mejores resultados. Como se utilizó la misma base de datos y los mismos parámetros en las medidas de Piella, no se implementaron las otras 8 medidas, pero utilizamos sus resultados para ampliar nuestro estudio de comparación.

De acuerdo con las Tablas 5.3 (CC de Spearman) y la Tabla III de Ma et. al (2015) [45], nuestra medida propuesta $MS-Q_W$, $L = 5$ supera significativamente las métricas de fusión de propósito general de 9 objetivos, en el contexto MEF. El valor promedio mínimo de CC de Spearman, considerando estas 9 métricas, fue -0.5612 y el valor máximo fue 0.6198, mientras que $MS-Q_W$, $L = 5$, la métrica mostró un valor promedio del CC de Spearman de 0.8060. Las

Tabla 5.3: Coeficientes de correlación de Kendall y Spearman entre $MS-Q_W$ y la opinión subjetiva, para todas las imágenes perteneciente a la base MEF Waterloo IVC 2015. Los mejores resultados se destacan en negrita.

Imágenes de entrada	CC de Kendall			CC de Spearman		
	$L = 1$ Piella's[8]	$L = 3$	$L=5$	$L = 1$ Piella's[8, 55]	$L = 3$	$L = 5$
“Ballons”	0.4286	0.6429	0.7143	0.6429	0.8095	0.9048
“Belgium house”	0.4728	0.8365	0.8365	0.6228	0.9102	0.9461
“Lamp1”	0.5000	0.7857	0.9286	0.6905	0.8571	0.9762
“Candle”	0.5000	0.5714	0.8571	0.6667	0.7143	0.9524
“Cave”	0.4286	0.5714	0.7143	0.5715	0.7619	0.8810
“Chinese garden”	0.2143	0.5000	0.7143	0.5000	0.6667	0.8810
“Farm house”	0.5000	0.5000	0.5000	0.5952	0.6190	0.6905
“House”	0.2143	0.5000	0.7143	0.4048	0.7143	0.8810
“Kluki”	0.0714	0.4286	0.7857	0.1190	0.6429	0.9286
“Lamp2”	0.2857	0.3571	0.2857	0.5952	0.6190	0.5714
“Landscape”	0.2857	0.4286	0.3551	0.4524	0.6667	0.5238
“Light house”	0.0714	0.2857	0.3551	0.1905	0.5714	0.5714
“Madison capitol”	0.4286	0.7143	0.5000	0.5238	0.8333	0.7143
“Memorial”	0.5714	0.6429	0.7143	0.7143	0.8095	0.8810
“Office”	0.2965	0.4447	0.5189	0.4579	0.6506	0.7109
“Tower”	0.5714	0.7143	0.7857	0.6190	0.8333	0.8571
“Venice”	0.1091	0.4001	0.6910	0.2036	0.6467	0.8264
Promedio	0.3499	0.5485	0.6460	0.5041	0.7251	0.8060

mejores correlaciones de $MS-Q_W$ con la opinión subjetiva en comparación con las 10 medidas objetivas de evaluación de calidad de fusión de imágenes, logradas en la base de datos MEF, se resaltan en negrita (Tabla 5.3).

Se determinaron los valores del CC de Kendall considerando la medida de calidad de fusión específica de Ma et al. (2015) [45] y utilizando la opinión subjetiva sobre la calidad de las imágenes fusionadas de la base de datos MEF Waterloo IVC 2015. Esta medida particular de calidad, proporciona puntajes ligeramente mejores en algunas secuencias de imágenes de esta base de datos. En la Tabla 5.3 (CC de Kendall), están resaltadas en negrita, las mejores correlaciones de $MS-Q_W$ con las evaluaciones subjetivas logradas en imágenes fusionadas de exposición múltiple, en comparación con la métrica de Ma et al. La medida propuesta $MS-Q_W$, $L = 5$, obtuvo un valor promedio de CC de Kendall de 0.6460.

5.5. Otros resultados

En esta sección se extiende el sistema multi-escala propuesto para el resto de las medidas de Piella.

La expresión de $MS-Q_S$ y $MS-Q_E$ son:

$$MS-Q_S(I, f) = \prod_{l=1}^L [Q_S(I_l, f_l)]^{\beta_l} \quad MS-Q_E(I, f) = \prod_{l=1}^L [Q_E(I_l, f_l)]^{\beta_l},$$

siendo I una colección de imágenes y f la imagen fusionada.

A continuación se reportan, en las tablas 5.4 y 5.5, los resultados de CC de Kendall y Spearman entre las medidas de calidad de fusión y la opinión subjetiva.

Tabla 5.4: Coeficientes de correlación de Kendall y Spearman entre $MS-Q_S$ y la opinión subjetiva, para todas las imágenes perteneciente a la base MEF Waterloo IVC 2015. Los mejores resultados se destacan en negrita.

Imágenes de entrada	CC de Kendall			CC de Spearman		
	$L = 1$ Piella's [8]	$L=3$	$L=5$	$L = 1$ Piella's [8, 55]	$L = 3$	$L = 5$
"Ballons"	0.6429	0.7143	0.7857	0.8095	0.8571	0.9286
"Belgium house"	0.6910	0.6910	0.8365	0.8623	0.8503	0.8982
"Lamp1"	0.7857	0.8571	0.8571	0.9048	0.9286	0.9286
"Candle"	0.7143	0.7143	0.7143	0.8810	0.8810	0.8810
"Cave"	0.3571	0.4286	0.5714	0.4762	0.5238	0.7619
"Chinese garden"	0.4286	0.5714	0.7143	0.6429	0.7381	0.8571
"Farm house"	0.6429	0.7857	0.7857	0.8333	0.9048	0.9048
"House"	0.4286	0.6429	0.7857	0.5714	0.8333	0.9048
"Kluki"	0.2143	0.5000	0.7143	0.3333	0.7143	0.8571
"Lamp2"	0.5000	0.3571	0.2857	0.6905	0.6190	0.5952
"Landscape"	0.2143	-0.0714	-0.1429	0.3333	0.0000	-0.1429
"Light house"	0.6429	0.2143	0.0000	0.8333	0.3095	0.0476
"Madison capitol"	0.6429	0.7857	0.7143	0.8095	0.8810	0.8810
"Memorial"	0.6429	0.7143	0.7857	0.8333	0.8810	0.9048
"Office"	0.3706	0.5189	0.4447	0.4699	0.6627	0.5784
"Tower"	0.6429	0.7857	0.7857	0.7857	0.9048	0.9048
"Venice"	0.4001	0.5455	0.6183	0.6108	0.7425	0.7785
Promedio	0.5271	0.5738	0.6033	0.6871	0.7195	0.7335

Con respecto a los costos computacionales, los algoritmos de todas las medidas $MS-Q$ son lineales con respecto al número de píxeles en la imagen y el nivel de descomposición de la escala. Por ejemplo, para la secuencia "Lamp1" que incluye una secuencia de imágenes de entrada con quince niveles de exposición (384×512) y ocho imágenes fusionadas, ver Figuras 5.9 y 5.10, toma un tiempo de computo general de aproximadamente 12.55 segundos, 14.05 segundos y 22.65 segundos, para $MS-Q_S$, $MS-Q_W$ y $MS-Q_E$ (con $L = 5$), respectivamente. La computadora utilizada se configuró con Intel Core i5 @1.80 GHz con 4 GB de RAM y un código MATLAB no optimizado. Como el número de imágenes de entrada de la secuencia suele ser diferente, no se proporciona un tiempo promedio.

Se puede decir que, todas las medidas de fusión $MS-Q$ reflejan mejor la calidad resultante de la imagen fusionada a una velocidad razonable que muestra un buen potencial para aplicaciones como la selección automática de parámetros, además de la evaluación del rendimiento y la

Tabla 5.5: Coeficientes de correlación de Kendall y Spearman entre $MS-Q_E$ y la opinión subjetiva, para todas las imágenes perteneciente a la base MEF Waterloo IVC 2015. Los mejores resultados se destacan en negrita.

Imágenes de entrada	CC de Kendall			CC de Spearman		
	$L = 1$ Piella's [8]	$L = 3$	$L=5$	$L = 1$ Piella's[8, 55]	$L = 3$	$L = 5$
"Ballons"	0.2857	0.5000	0.5714	0.4524	0.7143	0.7381
"Belgium house"	0.4728	0.5455	0.4001	0.6228	0.6707	0.5868
"Lamp1"	0.3571	0.6429	0.7143	0.5238	0.8095	0.8333
"Candle"	0.3571	0.5000	0.5000	0.5476	0.6667	0.6667
"Cave"	0.5000	0.5000	0.5714	0.7143	0.7143	0.7619
"Chinese garden"	0.2143	0.2857	0.4286	0.3571	0.5238	0.6190
"Farm house"	0.2857	0.2143	0.1429	0.4286	0.3571	0.1429
"House"	0.2143	0.4286	0.5000	0.4048	0.6429	0.7143
"Kluki"	0.0000	0.2143	0.2857	-0.0476	0.2619	0.3095
"Lamp2"	0.3571	0.3571	0.4286	0.6667	0.6667	0.6905
"Landscape"	0.1429	0.2143	0.0000	0.1429	0.2857	0.0476
"Light house"	0.0714	0.2857	0.4286	0.1905	0.5714	0.6667
"Madison capitol"	0.3571	0.4286	0.4286	0.4762	0.5238	0.6190
"Memorial"	0.5000	0.6429	0.7143	0.6667	0.7619	0.8095
"Office"	0.3706	0.3706	0.4447	0.4940	0.4940	0.6627
"Tower"	0.5000	0.4286	0.4286	0.5714	0.5714	0.5714
"Venice"	0.1818	0.3273	0.3273	0.3114	0.4671	0.4791
Promedio	0.3039	0.4050	0.4303	0.4425	0.5707	0.5834

comparación de algoritmos de fusión.

5.6. Aplicación

En esta sección se presenta una aplicación de la medida propuesta en el contexto de fotografías de pinturas. Ver Martinez et al. (2019) [50].

5.6.1. Obtención de fotografías de alta calidad de pinturas por fusión de imágenes

Recientemente, Haro et al. (2012) [28] propusieron un método de fusión para obtener una imagen de alta calidad de una pintura a partir de muchas imágenes de la misma, tomada con una cámara de mano, sin ninguna configuración particular, en un entorno no controlado. Este método de fusión de múltiples imágenes es un sistema complejo que involucra muchos problemas de procesamiento de imágenes. Se supone que todas las imágenes son fotografías de la misma pintura y que cada subconjunto de imágenes son tomadas desde el mismo punto de vista es consecutivo (en adelante, llamado ráfaga). Además, el algoritmo supone que la primera fotografía (imagen de referencia) es una vista frontal de toda la pintura, que puede haberse recortado manualmente. El algoritmo de fusión se divide en los siguientes pasos principales:

- detectar y agrupar las diferentes ráfagas de imágenes, utilizando el registro de imágenes y la detección de cambios en el punto de vista.
- fusionar cada ráfaga en una solo punto de vista con poco ruido y poco desenfoque de movimiento aplicando una versión del método del promedio después del registro, con los pesos proporcionales al nivel de nitidez de la imagen.
- eliminar defectos fusionando diferentes vistas, para obtener la imagen fusionada final.

Los autores identifican tres problemas que surgen cuando se toma una foto de una pintura en un museo: ruido, reflejos y resolución limitada [7]. Estas perturbaciones están presentes en fotografías profesionales (expertas) y aficionadas (inexpertas). Por lo tanto, no todas las imágenes tomadas en dicho entorno tienen una contribución positiva al proceso de fusión, por lo que es importante detectar qué subconjunto de la colección de imágenes produce la imagen fusionada de más alta calidad.

Una búsqueda rápida podría ser tomar todos los subconjuntos posibles, obtener la imagen fusionada de cada uno de ellos y ranquearlos según la medida de calidad. Dado que no todas las secuencias de imágenes tienen la misma cantidad, se propone un procedimiento de exclusión simple “todas menos una” que consiste en seleccionar el mejor subconjunto de imágenes, calificando la imagen fusionada obtenida con toda la secuencia menos una imagen con la métrica $MS-Q_W$. Si el mejor resultado obtenido todavía tiene defectos, el procedimiento puede aplicarse nuevamente.

Según Haro et al. (2012) [28], una buena imagen de una pintura se define como una imagen nítida, sin ruido ni reflejos de vidrio. Como se ha visto en los ejemplos anteriores, las medidas existentes que evalúan la calidad de la fusión de la imagen generalmente se basan en la medición de la transferencia de fidelidad de la característica (por ejemplo, bordes u otra información) de las imágenes de entrada en la imagen fusionada. En este caso, los defectos pueden ser características fuertes que los métodos mantienen como buena información. Por lo tanto, es importante eliminar la imagen defectuosa de la colección para garantizar una buena imagen fusionada. Para esta aplicación, se examinan dos ejemplos: una colección de 4 imágenes de la pintura de Van Dongen y una colección de 8 imágenes de la pintura de Latour, tomadas del conjunto de datos y la demostración disponible en [7]. Dos tipos de defectos aparecen en las imágenes de estas pinturas. Dado que la pintura de Van Dongen tiene un vidrio protector, las imágenes tomadas pueden tener defectos bien estructurados producidos por reflejos. Además, esta colección solo tiene cuatro imágenes de 1056×782 píxeles (el peor caso de información disponible, ya que el método de fusión necesita al menos tres imágenes) (ver Figura 5.12). En el segundo ejemplo, las imágenes

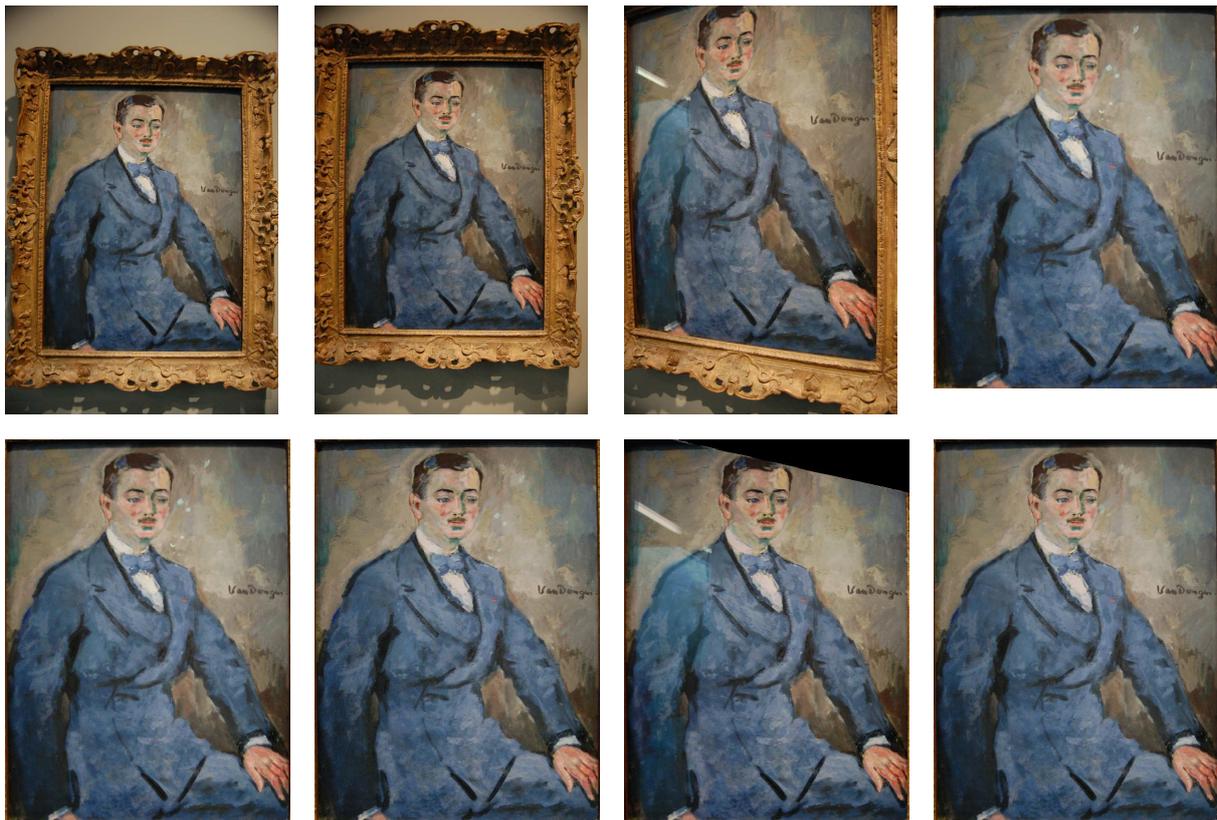


Figura 5.12: Ejemplo Van Dongen. Primera fila: imágenes originales: $i00$, $i01$, $i02$ y $i03$ (imagen de referencia), respectivamente. Segunda fila: las mismas imágenes después del registro. La imagen de referencia no se puede tomar del conjunto. Por lo tanto, solo se puede dejar una imagen a la vez. La imagen $i02$ tiene un sector faltante debido al ángulo en el que se tomó la imagen.

de la pintura de Latour (589×816 píxeles) en la colección tienen un resaltado difuso, creando un ruido brillante moteado. Estos reflejos no estructurados se deben al reflejo de lámparas o de ventanas en la escena. Cada imagen se tomó de una de las ocho ráfagas seleccionadas del conjunto original de 87 imágenes (no disponible). Cada ráfaga corresponde a un punto de vista diferente, moviéndose a las regiones más destacadas (ver Figura 5.13).

El experimento “todas menos una” consiste en aplicar el sistema de fusión de Buades et al. (2015) a subconjuntos de la colección de imágenes de cada pintura, obtenidos al eliminar una imagen. El sistema comprende tres técnicas diferentes para la eliminación de reflejos: mediana de gradientes en cada componente (component-wise median of gradients-CMG), mediana del vector RGB (median vector of RGB vectors-MV) y mínimo del vector RGB (component-wise minimum of RGB vectors-CMV) con una salida de las tres imágenes fusionadas distintas, una para cada técnica. La calidad de dichos resultados se calcula con la medida propuesta $MS-Q_W$, usando $L = 5$.

En ambos ejemplos, las imágenes de entrada se cargan secuencialmente, de acuerdo con el número de la imagen, como aparece en los ejemplos de Buades et al.(2015).

Las puntuaciones $MS-Q_W$ obtenidas de los subconjuntos de las imágenes de Van Dongen

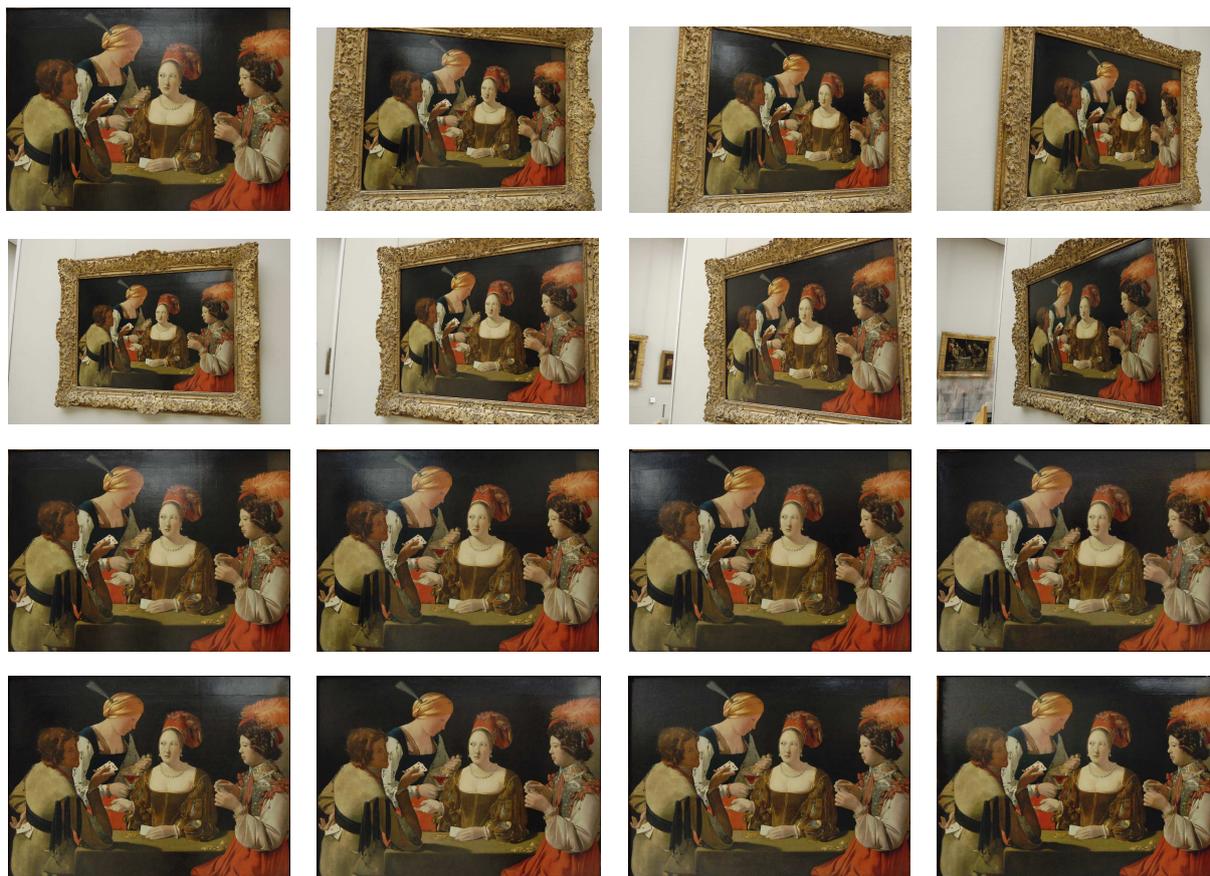


Figura 5.13: Ejemplo Latour. Imágenes originales en las primeras dos filas: i_{00} (imagen de referencia), i_{01} , ..., i_{06} y i_{07} , respectivamente. En las últimas dos filas, las mismas imágenes después del registro, respectivamente. Es importante notar los reflejos espurios presentes en las imágenes. Los algoritmos de fusión están diseñados para eliminar estos artefactos.

se muestran en la Tabla 5.6. Se nota que el subconjunto sin la imagen de entrada i_{00} y el subconjunto sin la imagen i_{01} producen puntajes más bajos que la imagen hecha fusionando toda la colección. Sin embargo, al eliminar la imagen i_{02} , la salida aumenta su calidad, eliminando todas las perturbaciones y dando un mejor resultado que la imagen hecha al fusionar toda la colección, ya que la pieza faltante de la imagen i_{02} incluye un borde en la salida fusionada. Además, el subconjunto sin la imagen i_{01} produce tres imágenes fusionadas con un nivel de calidad similar, ya que las imágenes restantes son de la misma ráfaga y tienen las mismas perturbaciones. Por lo tanto, la característica robusta del CMG parece pasar desapercibida.

Tabla 5.6: La pintura de Vandongen. Evaluación de calidad para imágenes compuestas de colecciones que se muestran en la Figura 5.12. Los mejores resultados se resaltan en negrita.

Imagen	Técnica	Imagen omitida			
		i_{00}	i_{01}	i_{02}	Todas
"Vandongen"	CMG	0.8737	0.8797	0.9910	0.8803
	MV	0.8675	0.8753	0.9912	0.8763
	CMV	0.8444	0.8541	0.9897	0.8522

Una inspección visual en el peor de los casos, subconjunto que consta de imágenes $i01$, $i02$ y $i03$ (imagen de referencia), ver las Figuras 5.14 y la Figura 5.15, indicando que la imagen $i00$ ayuda a eliminar un reflejo moteado. Además, la presencia de la imagen $i02$ en el proceso de fusión produce una línea diagonal no deseada en la parte superior izquierda de la salida del método CMV para la eliminación de reflejos. Estos hallazgos visuales corroboran la evaluación de calidad ciega realizada por el sistema propuesto.

El segundo ejemplo, la pintura Latour, tiene 8 imágenes en la secuencia, pero la primera debe estar en todos los subconjuntos, ya que es la imagen de referencia. Los puntajes obtenidos fusionando subconjuntos realizados dejando una imagen a la vez, se muestran en la Tabla 5.7. La puntuación más alta corresponde al subconjunto sin la imagen de entrada $i07$. La imagen $i07$ representa un punto de vista lateral, con reflejos que cubren grandes regiones de la imagen. La puntuación más baja corresponde al subconjunto sin imagen $i01$ (punto de vista frontal), que es claramente necesario para equilibrar el proceso de fusión.

Tabla 5.7: Evaluación de calidad para las imágenes compuestas realizadas con subconjuntos de imágenes tomadas de la pintura de Latour. Los mejores resultados se resaltan en negrita.

Imagen	Técnica	Imagen omitida							Todas
		$i01$	$i02$	$i03$	$i04$	$i05$	$i06$	$i07$	
"Latour"	CMG	0.9586	0.9608	0.9626	0.9610	0.9635	0.9610	0.9665	0.9645
	MV	0.9580	0.9597	0.9618	0.9604	0.9627	0.9598	0.9652	0.9641
	CMV	0.9451	0.9511	0.9510	0.9513	0.9532	0.9496	0.9584	0.9541

Una evaluación visual del peor de los casos, un subconjunto que consta de $i00$ (imagen de referencia), $i02$, $i03$, $i04$, $i05$, $i06$, $i07$, que se muestra en la Figura 5.16 y las Figuras 5.18 y 5.19, revelan artefactos puntuales en el fondo y líneas verticales y diagonales indeseables cerca de la parte superior de la imagen y sobre la mano del jugador infiel. Los mismos artefactos se encuentran en la imagen fusionada obtenida por el método CMV en el mejor de los casos, subconjunto que consta de $i00$ (imagen de referencia), $i01$, $i02$, $i03$, $i04$, $i05$, $i06$.

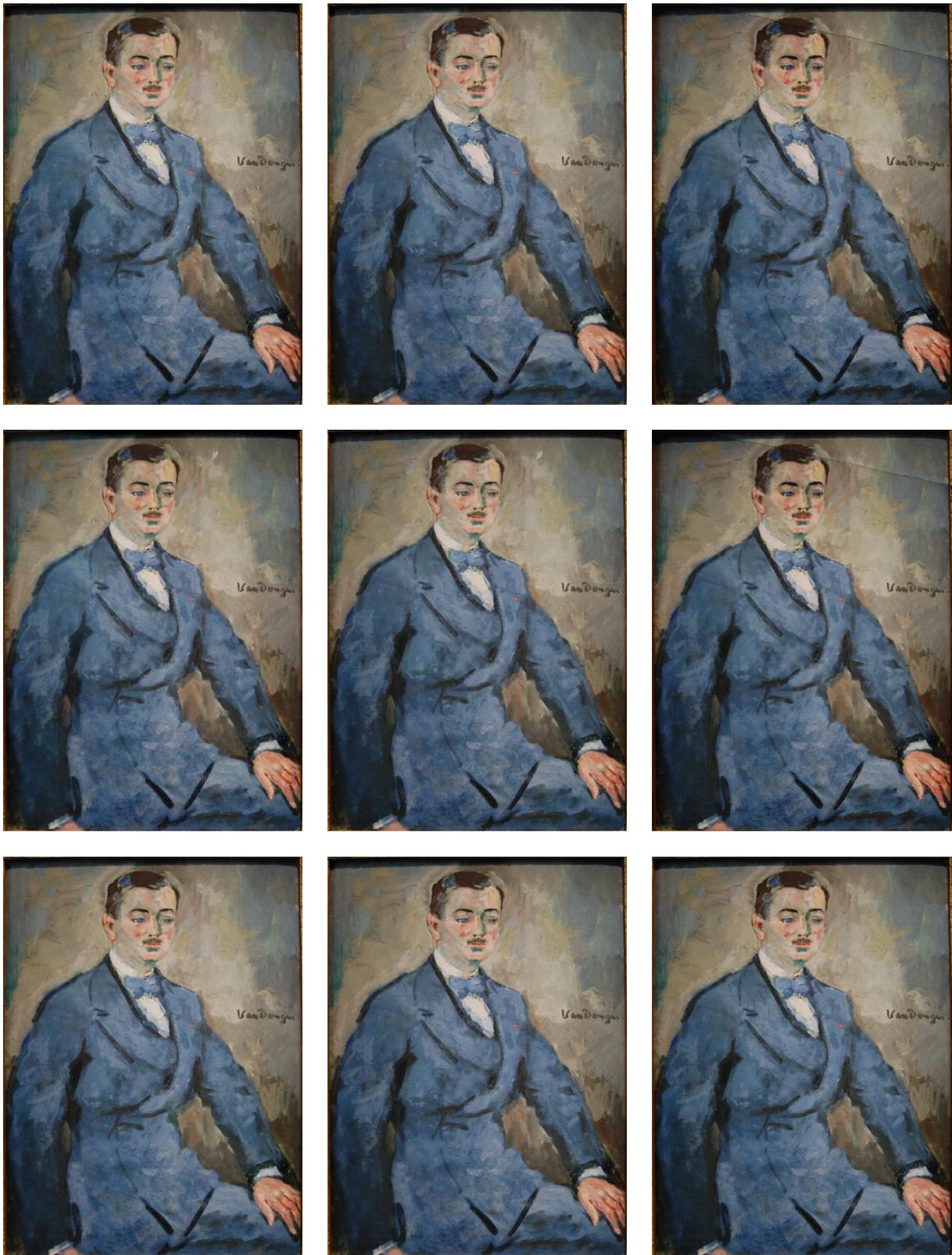


Figura 5.14: Ejemplo de Van Dongen. Imágenes fusionadas obtenidas mediante la aplicación de CMG, MV y CMV, respectivamente. Primera fila: imágenes fusionadas considerando todas las imágenes disponibles. Segunda fila: las peores imágenes fusionadas, obtenidas al omitir la imagen $i00$. Tercera fila: las mejores imágenes fusionadas, obtenidas al omitir la imagen $i02$.

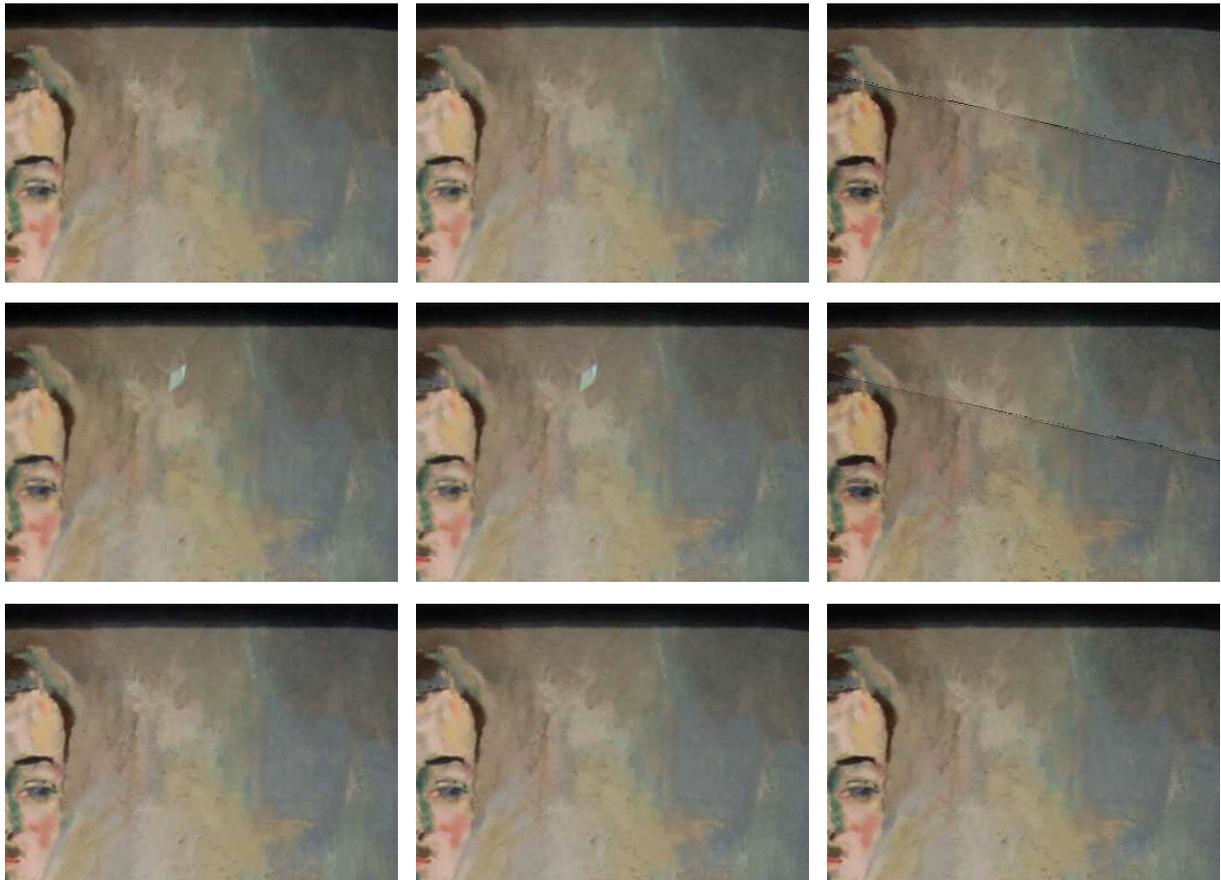


Figura 5.15: Ejemplo de Van Dongen (Detalles). Imágenes fusionadas obtenidas mediante la aplicación de técnicas CMG, MV y CMV, respectivamente. Primera fila: detalle de las imágenes fusionadas considerando todas las imágenes disponibles. Segunda fila: detalle de las peores imágenes fusionadas obtenidas al omitir la imagen $i00$. Tercera fila: detalle de las mejores imágenes fusionadas obtenidas al omitir la imagen $i02$.



Figura 5.16: Latour ejemplo. Imágenes fusionadas obtenidas mediante la aplicación de técnicas CMG, MV y CMV, respectivamente. Primera fila: imágenes fusionadas considerando todas las imágenes disponibles. Segunda fila: imágenes fusionadas obtenidas por los métodos CMG, MV y CMV considerando el conjunto de 7 mejores imágenes de entrada (conjunto de imágenes de entrada que omite la imagen $i07$). Tercer fila: imágenes fusionadas considerando el conjunto de 7 peores imágenes de entrada (conjunto de imágenes de entrada que excluye la imagen $i01$).

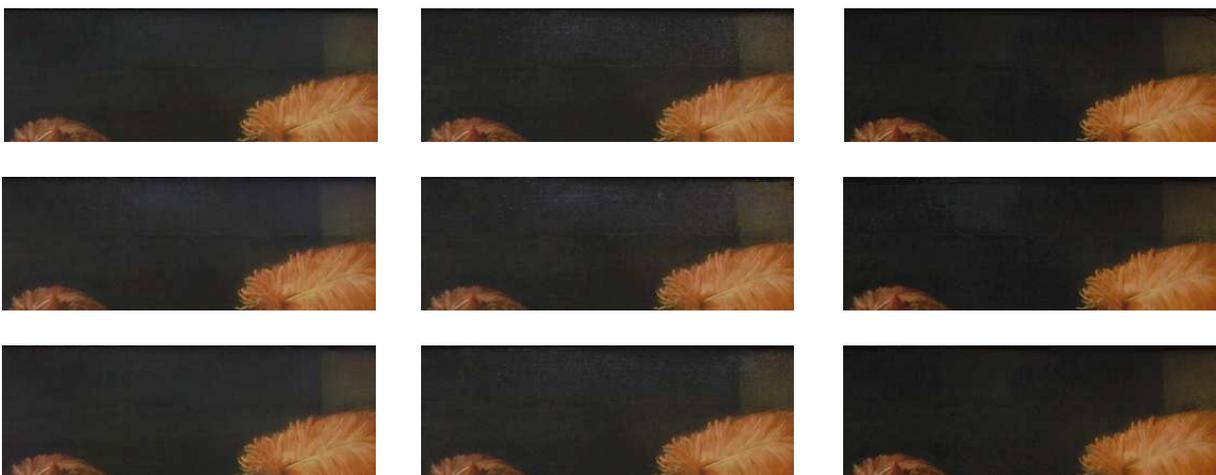


Figura 5.17: Ejemplo Latour (Detalles): pluma naranja. Imágenes fusionadas obtenidas mediante la aplicación de técnicas CMG, MV y CMV, respectivamente. Primera fila: detalle de las imágenes fusionadas considerando todas las imágenes disponibles. Segunda fila: detalle de las peores imágenes fusionadas obtenidas dejando fuera la imagen $i01$. Tercera fila: detalle de las mejores imágenes fusionadas obtenidas dejando fuera la imagen $i07$.

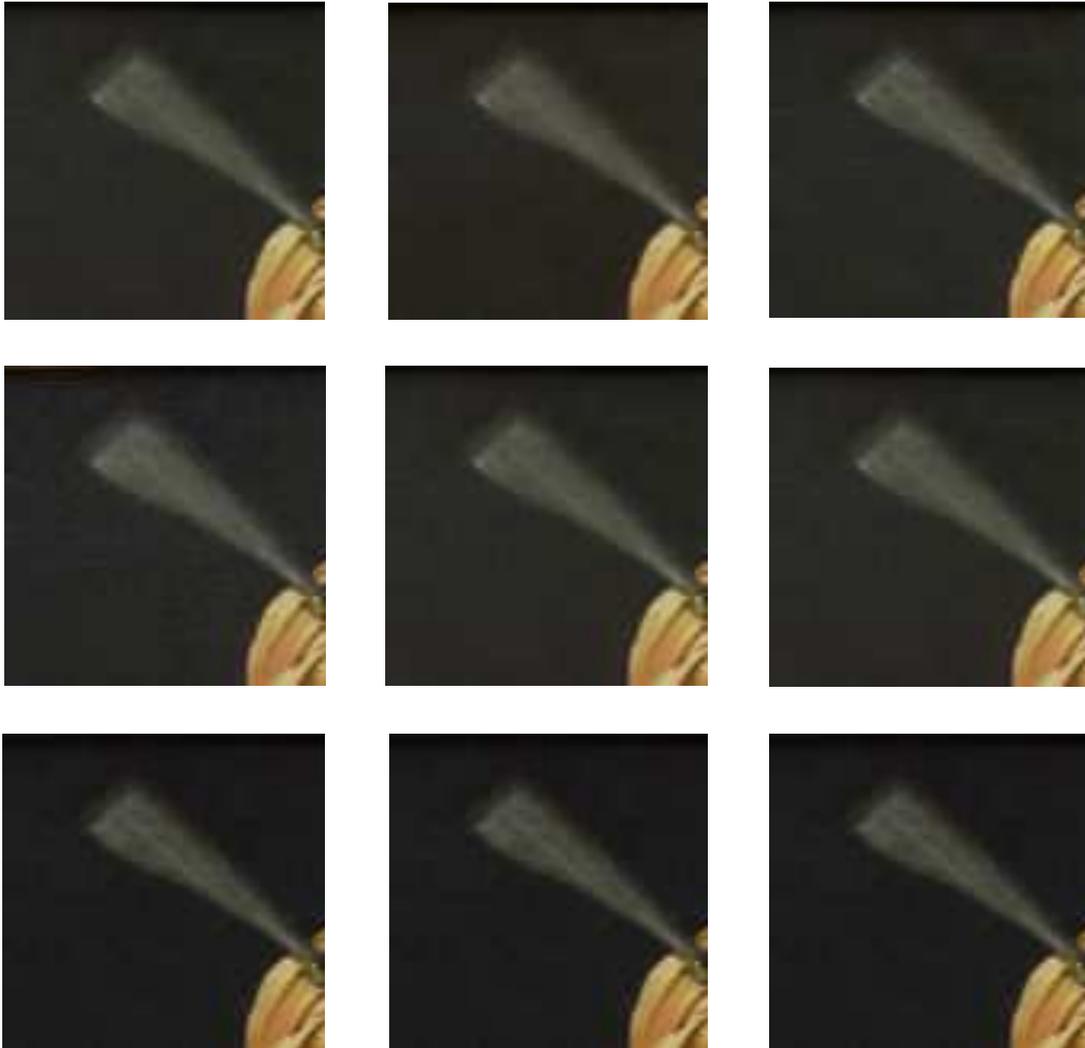


Figura 5.18: Ejemplo Latour (Detalles): pluma blanca. Imágenes fusionadas obtenidas mediante la aplicación de técnicas CMG, MV y CMV, respectivamente. Primera fila: detalle de las imágenes fusionadas considerando todas las imágenes disponibles. Segunda fila: detalle de las peores imágenes fusionadas obtenidas dejando fuera la imagen *i01*. Tercera fila: detalle de las mejores imágenes fusionadas obtenidas dejando fuera la imagen *i07*.

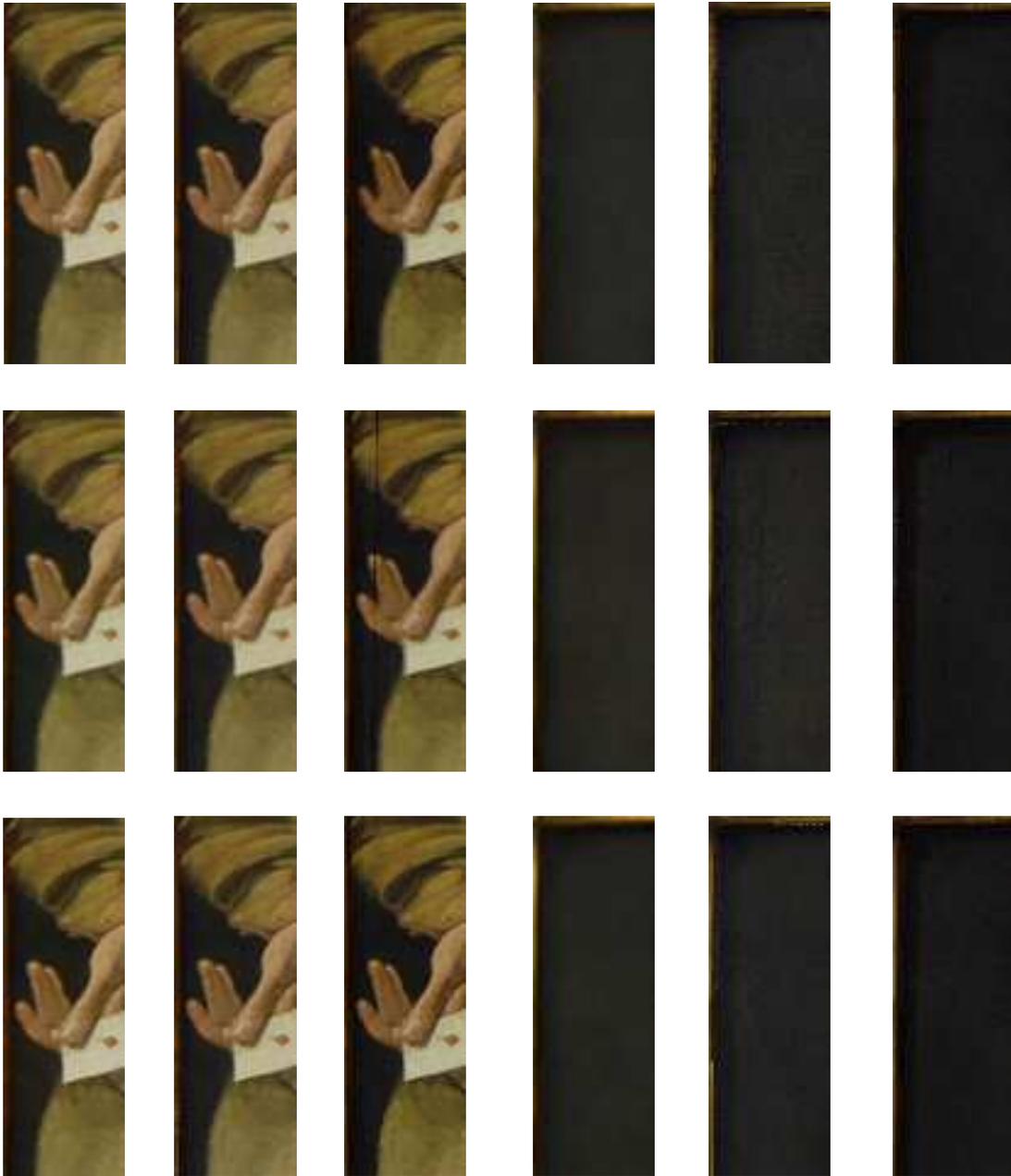


Figura 5.19: Ejemplo de Latour (Detalles): pequeña sección del jugador con cartas y esquina superior izquierda de la pintura. Imágenes fusionadas obtenidas mediante la aplicación de técnicas CMG, MV y CMV, respectivamente. Primera fila: detalle de las imágenes fusionadas considerando todas las imágenes disponibles. Segunda fila: detalle de las peores imágenes fusionadas obtenidas dejando fuera la imagen *i01*. Tercera fila: detalle de las mejores imágenes fusionadas obtenidas dejando fuera la imagen *i07*.

Capítulo 6

Conclusión y estudios futuros

En este capítulo se exponen las principales conclusiones obtenidas de la segunda parte de la tesis, y se presentan los estudios futuros.

6.1. Conclusión

En esta segunda parte de la tesis, se propone una medida de similitud estructural multi-escala, sin referencia y de propósito general para la evaluación objetiva de la calidad de la fusión de imágenes. El sistema de la medida propuesta consiste en el cálculo de la similitud entre la imagen fusionada y las imágenes de entrada, en cada una de las escalas, combinando los resultados a través de un promedio ponderado. Si se utilizan diferentes medidas de similitud en cada etapa del sistema se pueden crear otras medidas, adaptadas para aplicaciones específicas.

Para la validación de la medida propuesta, se consideró relevante analizar cuatro aspectos fundamentales: ser de propósito general, mediante su aplicación en diferentes escenarios de fusión; su estabilidad, por medio del cálculo de la varianza [44], utilizando imágenes de visión nocturna “TNO UN Camp”; su capacidad discriminativa, a través del empleo de pruebas estadísticas no-paramétricas para la comparación del rendimiento de los algoritmos de fusión [43]; y su correlación con las evaluaciones subjetivas, utilizando la base de datos de exposición múltiple MEF Waterloo IVC 2015, conservando la misma configuración de los parámetros utilizados por Ma et al. (2015) [45]. Se comparó, también el rendimiento de la medida propuesta con todas las medidas consideradas en este documento, fomentando la importancia de la investigación reproducible.

Recientemente, se abordó el problema de obtener una fotografía de alta calidad de una pintura [28], el cual implica la fusión de una gran cantidad de imágenes (ráfagas) del mismo objeto, algunas de ellas relevantes para este fin y otras solo con ruido y defectos. En este escenario, los

esquemas de fusión deben incorporar la selección automática de las imágenes de entrada para aumentar la calidad de la imagen fusionada. Para lograr este objetivo, se presentó un simple procedimiento de selección de exclusión utilizando la medida propuesta, el cual selecciona con éxito el mejor conjunto de imágenes de entrada en términos de la calidad de la imagen fusionada.

6.2. Estudios futuros

El reciente trabajo de Fang et al. (2020) [22] que presenta un estudio completo de evaluación de la calidad perceptiva de la fotografía obtenida por teléfonos inteligentes, motiva a continuar en la creación de una base de imágenes de pinturas capturadas por este tipo de dispositivos donde se considere la opinión subjetiva de la calidad de la imagen resultante del método de fusión propuesto por Haro et al. (2012) [28] y la evaluación de la performance de la medida de calidad propuesta en este contexto. Las siguientes imágenes de la Figuras 6.1 son dos ejemplos de prueba y fueron obtenidas mediante un teléfono Samsung Galaxy A20S. En la Figura 6.2, se presenta una imagen frontal de la pintura (candidata a imagen de referencia) y una ráfaga en dirección de arriba hacia abajo, cuyo defecto es el faltante de la parte superior de la pintura y la falta de nitidez, mientras que en la Figura 6.3, se muestra una toma desde el ángulo derecho y presenta una obstrucción producida por una persona.



Figura 6.1: Ejemplos de imágenes de pinturas: Benito Quinquela Martín. *Elevadores a pleno sol* (1945). Óleo sobre tela (200 cm. × 164 cm.). Museo Nacional de Bellas Artes (izquierda). Antonio Berni. *Manifestación* (1934). Temple sobre arpillera (180 cm. × 249.50 cm.). MALBA - Museo de Arte Latinoamericano de Buenos Aires (derecha).

Es por ello que se continuará trabajando en la calibración de los parámetros o modificación de algunos de los términos de la medida tal como sugiere el trabajo de Aja-Fernandez et al. (2006) [2]. Además, dado que la medida puede emplearse potencialmente en un amplio espectro de aplicaciones, como en el contexto de imágenes multi-exposición multi-foco [29], en imágenes binarias (mapas de bordes [25]) , y en imágenes satelitales [42], entre otras, se explorará estos problemas cuando estas bases de datos estén disponibles.



Figura 6.2: Ejemplo Quinquela (4160 × 3120). Imagen frontal (candidata a imagen de referencia) y ráfaga de imágenes en la dirección arriba hacia abajo (faltante parte superior de la pintura e imagen borrosa).



Figura 6.3: Ejemplo Berni. (4160 × 3120). Imagen frontal (candidata a imagen de referencia) y ráfaga de imágenes obstruidas por una persona.

Apéndice A

Algoritmos de simulación

A continuación se presenta el marco teórico necesario para construir una cadena de Markov de imágenes. Se tiene la intención de que esta cadena converja a una realización de un proceso bajo una distribución de Gibbs. Para más detalles ver Bustos y Ojeda (1998) [11]

A.1. Cadenas de Markov (de imágenes)

Sea $\Delta \in \mathcal{S}$ (soporte finito), sea $\Omega_* = E^\Delta \times \dots \times E^\Delta \times \dots = (E^\Delta)^{\mathbb{N} \cup \{0\}}$ (el conjunto de todas las sucesiones de imágenes en E^Δ , es decir $\omega \in \Omega_* \Rightarrow \omega = (\omega(i))_{i \geq 0}$ tal que $\omega(i) \in E^\Delta$) y sea $\mathcal{D}_* \doteq (\mathcal{E}^\Delta)^{\mathbb{N} \cup \{0\}}$ la σ -álgebra producto. Notar que E^Δ es un espacio finito. Fijada una función de densidad ρ y una sucesión de funciones de transición $(p_n)_{n \in \mathbb{N}}$ sobre $(E^\Delta, \mathcal{E}^\Delta)$. Sea $P(\cdot; \rho, (p_n))$ la única probabilidad sobre $(\Omega_*, \mathcal{D}_*)$ que cumple:

$$\begin{aligned} P(\{\omega \in \Omega_* / \omega(0) = x_0, \omega(1) = x_1, \dots, \omega(n) = x_n\}; \rho, (p_n)) \\ = \rho(x_0) p_1(x_0, x_1) \cdots p_n(x_{n-1}, x_n). \end{aligned}$$

Notación A.1.1 *Se pondrá:*

- $P(\cdot; x, (p_n))$ en lugar de $P(\cdot; \delta_x, (p_n))$, para cada $x \in E^\Delta$
- $P(\cdot; \rho, p)$ en lugar de $P(\cdot; \rho, (p_n))$ si $p_n = p$ para todo $n \in \mathbb{N}$

Sea $(X_n)_{n \geq 0}$ una sucesión de campos aleatorios \mathcal{D}_* -medibles tal que para cada $n \in \mathbb{N} \cup \{0\}$, $X_n : \Omega_* \rightarrow E^\Delta$ está definido por $X_n(\omega) = \omega(n)$ para cada $\omega \in \Omega_*$, y se cumple:

- $P(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n; \rho, (p_n)) = \rho(x_0) p_1(x_0, x_1) p_2(x_1, x_2) \cdots p_n(x_{n-1}, x_n)$,
luego también se cumple:

- $P(X_0 = x; \rho, (p_n)) = \rho(x)$ y $P(X_n = y | X_0 = x_0, \dots, X_{n-1} = x; \rho, (p_n)) = p_n(x, y) = P(X_n = y | X_{n-1} = x; \rho, (p_n))$, Es decir $(X_n)_{n \geq 0}$ es una $P(\cdot; \rho, (p_n))$ -cadena de Markov. A ρ se la llama densidad inicial y a las funciones de transición p_n kernels de Markov.

A.2. Gibbs sampler

Definición A.2.1 Sea $\Delta \in \mathcal{S}$ (soporte finito), sea π una densidad sobre $(E^\Delta, \mathcal{E}^\Delta)$ se dice que es de Gibbs si para cada $x_\Delta \in E^\Delta$:

$$\pi(x_\Delta) = \frac{\exp\left(-\sum_{\Lambda \in \mathcal{S}(\Delta)} \Phi_\Lambda(x_\Delta)\right)}{\sum_{z_\Delta \in E^\Delta} \exp\left(-\sum_{\Lambda \in \mathcal{S}(\Delta)} \Phi_\Lambda(z_\Delta)\right)}, \quad (\text{A.1})$$

donde $\Phi = (\Phi_\Lambda)_{\Lambda \in \mathcal{S}(\Delta)}$ es un potencial. Para cada $s \in \Delta$ se define su característica local π_s como la función definida sobre E^Δ , que para cada x_Δ cumple:

$$\pi_s(x_\Delta) = \frac{\exp\left(-\sum_{\Lambda \in \mathcal{S}_{\{s\}}(\Delta)} \Phi_\Lambda(x_\Delta)\right)}{\sum_{z_s \in E} \exp\left(-\sum_{\Lambda \in \mathcal{S}_{\{s\}}(\Delta)} \Phi_\Lambda(z_s x_{\Delta \setminus \{s\}})\right)},$$

La realización de una imagen proveniente de una densidad de Gibbs es intratable debido al alto número de sumandos en el denominador de la ecuación A.1 es un algoritmo que fijada una distribución de Gibbs, utiliza sus características locales para generar una sucesión de imágenes provenientes de una sucesión de campos aleatorios cuyas densidades convergen a la densidad de Gibbs.

Sea $(s_k)_{k \geq 1}$ una sucesión en Δ tal que $\#\{k/s_k\} = \infty$ para todo $s \in S$. Para cada $k \in \mathbb{N}$ sea p_k la función de transición sobre $(E^\Delta, \mathcal{E}^\Delta)$ dada por:

$$p_k(v, y) = \begin{cases} \pi_{s_k}(y) & \text{si } y_{\Delta \setminus \{s_k\}} = v_{\Delta \setminus \{s_k\}} \\ 0 & \text{si caso contrario} \end{cases}$$

Sea x_0 una imagen (inicial) cualquiera de E^Δ , se define ahora por inducción una sucesión $(X_n)_{n \geq 0}$ de campos aleatorios tal que sea una $P(\cdot; x_0, (p_n))$ -cadena de Markov. Sea X_0 la función constante que cumple $X_0(\omega) = x_0$, para todo $\omega \in \Omega_*$. Sea $k \geq 0$, se suponen definidas X_0, \dots, X_k , se define ahora X_{k+1} . Sea X_{k+1} definida sobre Ω_* tal que para cada $\omega \in \Omega_*$:

$$(X_{k+1}(\omega))_t = \begin{cases} (X_k(\omega))_t & \text{si } t \neq s_{k+1} \\ \eta_{s_{k+1}} & \text{si } t = s_{k+1} \end{cases}$$

donde $P(\eta_s = \zeta_s | X_k = v, X_{k-1} = \dots, X_0 = x_0; x_0, (p_k)) = \pi_s(\zeta_s v_{\Delta \setminus \{s\}})$.

Luego $(X_n)_{n \geq 0}$ es una $P(\cdot; x_\Delta, (p_n))$ –cadena de Markov como en la Sección A.1.

El teorema de Geman y Geman asegura que para todo $y \in E^\Delta$ se cumple:

$$P(X_n = y; x_0, (p_n)) \xrightarrow[n]{} \pi(y),$$

es decir, la densidad de X_n converge a la densidad de Gibbs.

Como casos particulares, al considerar un campo aleatorio $X : \omega \rightarrow E^S$ (S infinito) con distribuciones de Gibbs-Markov ($P_X \in \mathcal{G}(\Phi)$, con Φ ∂ –potencial), la densidad de X_Λ (condicional al entorno $x_{\partial\Lambda}$) es una distribución de Gibbs sobre E^Λ :

$$\pi(x_\Lambda) = p(x_\Lambda | x_{\partial\Lambda}) := P_{x_{\partial\Lambda}}(X_\Lambda = x_\Lambda) = \frac{e^{-\sum_{\Delta \in C_\Lambda} \Phi_\Delta(x_\Lambda x_{S \setminus \Lambda})}}{\sum_{z_\Lambda \in E^\Lambda} e^{-\sum_{\Delta \in C_\Lambda} \Phi_\Delta(z_\Lambda w_{S \setminus \Lambda})}}$$

para el potencial $\tilde{\Phi} = (\Phi_\Delta)_{\Delta \in \mathcal{S}(\Lambda^*)}$. Cuyas características locales cumplen:

$$\pi_s(x_\Lambda) = \frac{e^{-\sum_{\Delta \in C_{\{s\}}} \Phi_\Delta(x_{\{s\}} x_{S \setminus \Lambda})}}{\sum_{z_{\{s\}} \in E^{\{s\}}} e^{-\sum_{\Delta \in C_{\{s\}}} \Phi_\Delta(z_{\{s\}} w_{S \setminus \{s\}})}}$$

donde $x_{\partial\Lambda}$ está fijo y funciona como parámetro.

Bibliografía

- [1] A. Agresti. *An introduction to categorical data analysis*. Wiley, New York, 1996.
- [2] S. Aja-Fernandez, R. S. J. Estepar, C. Alberola-Lopez, and C. Westin. Image quality assessment based on local variance. In *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4815–4818, 2006.
- [3] J.R. Anderson, E.E. Hardy, J.T. Roach, and R.E. Witmer. *A Land use and land cover classification system for use with remote sensor data*. Geological Survey Professional Paper 964. U.S. Govt. Print. Off., 1976.
- [4] J. Besag. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society. Series B*, 36:192–236, 1974.
- [5] C.F. Borges. On the estimation of markov random field parameters. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 21(3):216–224, 1999.
- [6] D. Brunet, E. R. Vrscay, and Z. Wang. On the mathematical properties of the structural similarity index. *IEEE Transactions on Image Processing*, 21(4):1488–1499, 2012.
- [7] A. Buades, G. Haro, and E. Meinhardt. Obtaining high quality photographs of paintings by image fusion. *IPOL Journal*, 5, 2015.
- [8] A.M. Bueno, J. Alvarez Borrego, L. Acho, and M.C. Chávez Sánchez. Polychromatic image fusion algorithm and fusion metric for automatized microscopes. *Optical Engineering.*, 44(9):093201.1–093201.7, 2005.
- [9] O. Bustos and A. Guerrero. Breve introducción a la matemática de la estadística espacial. *Ensaio Matemáticos 20*, Sociedade Brasileira de Matemática, 2011.
- [10] O.H. Bustos, A.C. Frery, and S.M. Ojeda. Strong markov process in image modeling. *Brazilian Journal of Probability and Statistics*, 2:149–194, 1998.

- [11] O.H. Bustos and S.M. Ojeda. Campos Aleatorios Markovianos en Procesamiento de Imágenes. Trabajos de Matemática, Serie B 25/94, Facultad de Matemática, Astronomía y Física. Universidad Nacional de Córdoba, 1994.
- [12] O.H. Bustos, V. Rulloni, and J. Martinez. Distribuciones de Gibbs: Volumen 1. Trabajos de Matemática, Serie B 65/17, Facultad de Matemática, Astronomía y Física. Universidad Nacional de Córdoba, 2017.
- [13] H. Chen and K.P. Varshney. A human perception inspired quality metric for image fusion based on regional information. *Information Fusion*, 8(2 SPEC. ISS.):193–207, 4 2007.
- [14] Y. Chen and R.S. Blum. A new automated quality assessment algorithm for image fusion. *Image Vision Comput.*, 27(10):1421–1432, September 2009.
- [15] J. Cohen. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20(1):37, 1960.
- [16] W. Conover. *Practical nonparametric statistics*. Wiley series in probability and statistics. Wiley, New York, NY, 3. ed edition, 1999.
- [17] G. R. Cross and A. K. Jain. Markov random field texture models. *IEEE Trans. Pattern Anal. Mach. Intell.*, 5(1):25–39, 1983.
- [18] N. Cvejic, C.N. Canagarajah, and D.R. Bull. Image fusion metric based on mutual information and tsallis entropy. *Electronics Letters*, 42:626–627(1), May 2006.
- [19] N. Cvejic, A. Loza, D. Bull, and N. Canagarajah. A similarity metric for assessment of image fusion algorithms. *International Journal of Signal Processing*, 2(3):178–182, 2005.
- [20] C. de Franchis, E. Meinhardt-Llopis, J. Michel, J. M. Morel, and G. Facciolo. Automatic sensor orientation refinement of pléiades stereo images. In *2014 IEEE Geoscience and Remote Sensing Symposium*, pages 1639–1642, July 2014.
- [21] document Rec. ITU-R BT.500-12. Methodology for the subjective assessment of the quality of television pictures, 1993.
- [22] Y. Fang, H. Zhu, Y. Zeng, K. Ma, and Z. Wang. Perceptual quality assessment of smartphone photography. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020.
- [23] M. Friedman. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association*, 32(200):675–701, 1937.

- [24] H.O. Georgii. *Gibbs Measures and Phase Transitions*. De Gruyter Studies in Mathematics. De Gruyter, 1988.
- [25] J. Gimenez, J. Martinez, and A.G. Flesia. Unsupervised edge map scoring: A statistical complexity approach. *Computer Vision and Image Understanding*, 122:131 – 142, 2014.
- [26] S. W. Greenhouse and S. Geisser. On methods in the analysis of profile data. *Psychometrika*, 24(2):95–112, Jun 1959.
- [27] J. M. Hammersley and P. Clifford. Markov field on finite graphs and lattices. <http://www.statslab.cam.ac.uk/~grg/books/hammfest/hamm-cliff.pdf>, 1971.
- [28] G. Haro, A. Buades, and J. M. Morel. Photographing paintings by image fusion. *SIAM Journal on Imaging Sciences*, 5(3):1055–1087, 2012.
- [29] R. Hassen, Z. Wang, and M.M.A. Salama. Objective quality assessment for multiexposure multifocus image fusion. *IEEE Transactions on Image Processing*, 24(9):2712–2724, 2015.
- [30] M. Hossny, S. Nahavandi, and D. Creighton. Comments on information measure for performance of image fusion. *Electronics Letters*, 44:1066–1067(1), August 2008.
- [31] W. Huang and Z. Jing. Evaluation of focus measures in multi-focus image fusion. *Pattern Recognition Letters*, 28:493–500, 2007.
- [32] H. Huynh and L. S. Feldt. Estimation of the box correction for degrees of freedom from sample data in randomized block and split-plot designs. *Journal of Educational Statistics*, 1(1):69–82, 1976.
- [33] E. Ising. Beitrag zur theorie des ferromagnetismus. *Zeitschrift für Physik*, 31:253–258, 1997.
- [34] J. O. Johansson. Parameter-estimation in the auto-binomial model using the coding- and pseudo-likelihood method approached with simulated annealing and numerical optimization. *Pattern Recognition Letters*, 22(11):1233–1246, 2001.
- [35] M.S. Kaiser, S.N. Lahiri, and D.J. Nordman. Goodness of fit tests for a class of markov random field models. *The Annals of Statistics*, 40(1):104–130, 2012.
- [36] A. Kaplan, M.S. Kaiser, S.N. Lahiri, and D.J. Nordman. Simulating markov random fields with a conclave-based gibbs sampler. *Journal of Computational and Graphical Statistics*, 0(0):1–11, 2019.
- [37] M.S. Klamkin and A. Meir. Ptolemy’s inequality, chordal metric, multiplicative metric. *Pacific J. Math.*, 101(2):389–392, 1982.

- [38] V. Kumar. *Introduction to Parallel Computing*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2nd edition, 2002.
- [39] M. H. Kutner, C. J. Nachtsheim, J. Neter, and W. Li. *Applied Linear Statistical Models*. McGraw Hill, 5th edition, 2005.
- [40] S. R. Lele and J. K. Ord. Conditional Least Squares Estimation for Spatial Processes: Some Asymptotics Results. Technical Report 65, Pennsylvania State University, Dept. Statistics, 1986.
- [41] H. Li, B. S. Manjunath, and S. K. Mitra. Multisensor image fusion using the wavelet transform. *Graph. Models Image Process.*, 57(3):235–245, 1995.
- [42] J. Liu, J. Huang, S. Liu, H. Li, Q. Zhou, and J. Liu. Human visual system consistent quality assessment for remote sensing image fusion. *Journal of Photogrammetry and Remote Sensing*, 105:79 – 90, 2015.
- [43] Z. Liu, E. Blasch, and John V. Statistical comparison of image fusion algorithms: Recommendations. *Information Fusion*, 36:251–260, 2017.
- [44] Z. Liu, E. Blasch, Z. Xue, J. Zhao, R. Laganière, and W. Wu. Objective assessment of multi-resolution image fusion algorithms for context enhancement in night vision: A comparative study. *IEEE Transactions on Pattern Analysis and Machine Intelligence.*, 34(1):94–109, 2012.
- [45] K. Ma, K. Zeng, and Z. Wang. Perceptual quality assessment for multi-exposure image fusion. *IEEE Transactions on Image Processing*, 24(11):3345–3356, 2015.
- [46] J. Martinez and S. Pistonesi. Gibbs-potential decomposition based on coding sets. *To appear*, 2020.
- [47] J. Martinez, S. Pistonesi, and A. G. Flesia. Inference strategies for texture parameters. In A. Pardo and J. Kittler, editors, *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, pages 460–467, Cham, 2015. Springer International Publishing.
- [48] J. Martinez, S. Pistonesi, M. C. Maciel, and A. G. Flesia. Image fusion quality measure based on a multi-scale approach. In G. Bebis, editor, *Advances in Visual Computing. ISVC 2016*, pages 836–845, Cham, 2016. Springer International Publishing.
- [49] J. Martinez, S. Pistonesi, M. C. Maciel, and A. G. Flesia. Parameter estimation in a gibbs-markov field texture model based on a coding approach. In *2018 IEEE Statistical Signal Processing Workshop (SSP)*, pages 105–109, June 2018.

- [50] J. Martinez, S. Pistonesi, M. C. Maciel, and A. G. Flesia. Multi-scale fidelity measure for image fusion quality assessment. *Information Fusion*, 50:197 – 211, 2019.
- [51] D.E. Molina, M. Datcu, and D. Gleich. Cramer-rao bound-based evaluation of texture extraction from sar images. In *Systems, Signals and Image Processing, 2009. IWSSIP 2009. 16th International Conference on*, pages 1–4, 2009.
- [52] D.E. Molina, D. Gleich, and M. Datcu. Evaluation of bayesian despeckling and texture extraction methods based on gauss-markov and auto-binomial gibbs random fields: Application to terrasars-x data. *Geoscience and Remote Sensing, IEEE Transactions on*, 50(5):2001–2025, May 2012.
- [53] V. Petrović. Subjective tests for image fusion evaluation and objective metric validation. *Information Fusion*, 8(2):208–216, 2007.
- [54] G. Piella. *Adaptative Wavelets and their Applications to Image Fusion and Compression*. PhD thesis, Faculteit der Natuurwetenschappen, Wiskunde en Informatica. Universiteit van Amsterdam, 2003.
- [55] G. Piella. Image fusion for enhanced visualization: A variational approach. *International Journal of Computer Vision*, 83(1):1–11, 2009.
- [56] G. Piella and H. Heijmans. A new quality metric for image fusion. In *Proceedings of the International Conference on Image Processing*, volume 3, pages 173–176, 2003.
- [57] S. Pistonesi, J. Martinez, and S. M. Ojeda. A sensibility study of the autobinomial model estimation methods based on a feature similarity index. *Computational Statistics*, 31(4):1327–1357, December 2016.
- [58] S. Pistonesi, J. Martinez, S.M. Ojeda, and R. Vallejos. Structural Similarity Metrics for Quality Image Fusion Assessment: Algorithms. *Image Processing On Line*, 8:345–368, 2018.
- [59] V. Ponomarenko, V. Lukin, A. Zelensky, R. Egiazarian, M. Carli, and F. Battisti. Tid2008 - a database for evaluation of full-reference visual quality assessment metrics. *Advances of Modern Radioelectronics*, 10:30–45, 2009.
- [60] O. Rockinger. Image sequence fusion using a shift-invariant wavelet transform. In *Proceedings of the International Conference on Image Processing*, volume 3, pages 288–291, 1997.
- [61] O. Rockinger. Image fusion toolbox. <http://www.metapix.de/toolbox.htm/>, 1999.

- [62] V. Rulloni. *Texturas en imágenes binarias: síntesis, inpainting e imputación*. PhD thesis, Facultad de Ciencias Exactas, Físicas y Naturales. Universidad Nacional de Córdoba, 2014.
- [63] M. Schöder, H. Rehrauer, K. Seidel, and M. Datcu. Spatial information retrieval from remote sensing images: Part b. gibbs markov random fields. *IEEE Transaction on Geoscience and Remote Sensing*, 36:1446–1455, 1998.
- [64] M. Schröder, S. Seidel, and M. Datcu. Gibbs random field models for image content characterization. In *IEEE International Geoscience and Remote Sensing Symposium (IGARSS97)*, 1:258–260, pages 258–260, 1997.
- [65] H.R. Sheikh, M.F. Sabir, and A.C. Bovik. A statistical evaluation of recent full reference image quality assessment algorithms. *IEEE Transactions on Image Processing*, 15(11):3440–3451, 2006.
- [66] J. Silvestre-Blanes, T. Albero-Albero, I. Miralles, R. Pérez-Llorens, and J. Moreno. A public fabric database for defect detection methods and results. *Autex Research Journal*, 19, 06 2019.
- [67] T. Stathaki. *Image Fusion: Algorithms and Applications*. Academic Press, Inc., Orlando, FL, USA, 2008.
- [68] A. Toet and E.M. Franken. Perceptual evaluation of different image fusion schemes. *Displays*, 24(1):25–37, 2003.
- [69] R. O. Vallejos and G. García-Donato. Bayesian analysis of contaminated quarter plane moving average models. *Journal of Statistical Computation and Simulation*, 76(2):131–147, 2006.
- [70] M.J. Wainwright, O. Schwartz, and E. Simoncelli. *Natural image statistics and divisive normalization: Modeling nonlinearity and adaptation in cortical neurons*, pages 203–222. MIT Press, 2002.
- [71] B.A. Wandell. *Foundations of Vision*. Sinauer Associates, 1995.
- [72] P-W. Wang and B. Liu. A novel image fusion metric based on multi-scale analysis. In *Proceedings ICSP 2008*, pages 1398–1402, 2008.
- [73] Q. Wang, Y. Shen, and J. Jin. Performance evaluation of image fusion techniques. In *Image Fusion: Algorithms and Applications*, pages 469–492, New York, NY, USA, 2008. Academic.

- [74] Z. Wang and A.C. Bovik. A universal image quality index. *IEEE Signal Processing Letters*, 9(3):81–84, 2002.
- [75] Z. Wang and A.C. Bovik. Mean squared error: Love it or leave it? a new look at signal fidelity measures. *IEEE Signal Processing Magazine*, 26(1):98–117, 2009.
- [76] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004.
- [77] Z. Wang and Q. Li. Information content weighting for perceptual image quality assessment. *IEEE Transactions on Image Processing*, 20(5):1185–1198, 2011.
- [78] Z. Wang, E.P. Simoncelli, and A.C. Bovik. Multi-scale structural similarity for image quality assessment. In *Proceedings IEEE Asilomar Conference on Signals, Systems and Computers*, pages 1398–1402, 2003.
- [79] G. Winkler. *Image Analysis, Random Fields and Dynamic Monte Carlo Methods: A Mathematical Introduction*. Springer, 3rd edition, 1995.
- [80] H.R. Wu and K.R. Rao. *Digital Video Image Quality and Perceptual Coding*. Signal Processing and Communications. CRC Press, 2017.
- [81] C.S. Xydeas and V. Petrovic. Objective image fusion performance measure. *Electronics letters*, 36(4):308–309, 2000.
- [82] C. Yang, J. Zhang, X. Wang, and X. Liu. A novel similarity based quality metric for image fusion. *Information Fusion*, 9(2):156–160, April 2008.
- [83] Z. Zhang and R.S. Blum. A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application. In *Proceedings of the IEEE*, volume 87, pages 1315–1326, 1999.
- [84] Y. Zheng, E.A. Essock, B.C. Hansen, and A.M. Haun. A new metric based on extended spatial frequency and its application to dwt based fusion algorithms. *Inf. Fusion*, 8(2):177–192, April 2007.