



UNIVERSIDAD DE LOS ANDES
FACULTAD DE INGENIERÍA
ESCUELA DE INGENIERÍA ELÉCTRICA

RESTAURACIÓN DE IMÁGENES USANDO TÉCNICAS DE RECONSTRUCCIÓN DE SENSADO COMPRIMIDO

Br. Jerick Emilio Ordenes Sepúlveda

Mérida, Noviembre, 2012

UNIVERSIDAD DE LOS ANDES
FACULTAD DE INGENIERÍA
ESCUELA DE INGENIERÍA ELÉCTRICA

RESTAURACIÓN DE IMÁGENES USANDO TÉCNICAS DE
RECONSTRUCCIÓN DE SENSADO COMPRIMIDO

Trabajo de Grado presentado como requisito parcial para optar al título de Ingeniero
Electricista

Br. Jerick Emilio Ordenes Sepúlveda

Tutor: Prof. José Luis Paredes

Tutor: Prof. Juan Marcos Ramírez

Mérida, Noviembre, 2012

UNIVERSIDAD DE LOS ANDES
FACULTAD DE INGENIERÍA
ESCUELA DE INGENIERÍA ELÉCTRICA

RESTAURACIÓN DE IMÁGENES USANDO TÉCNICAS DE RECONSTRUCCIÓN DE SENSADO COMPRIMIDO

Br. Jerick Emilio Ordenes Sepúlveda

Trabajo de Grado presentado en cumplimiento parcial de los requisitos exigidos para optar al título de Ingeniero Electricista, aprobado en nombre de la Universidad de Los Andes por el siguiente Jurado.

Prof. José Luis Paredes

C.I: 10.107.184

Prof. Rubén Medina

C.I: 5.447.254

Prof. Emiro J. Ibarra

C.I: 17.027.152



ACTA DE TRABAJO DE GRADO

Los suscritos, Miembros del Jurado designado por el Consejo de Escuela de Ingeniería Eléctrica de la Universidad de Los Andes, para conocer y evaluar el Trabajo de Grado titulado: **“Restauración de imágenes usando técnicas de reconstrucción de sensado comprimido”** presentado por el(la) Bachiller: **Ordenes Sepúlveda Jerick Emilio**, titular de la Cédula de Identidad N° 23208290, cuyo Tutores fueron los **Prof. José Luis Paredes** y **Prof. Juan Marcos Ramírez**, titular de las Cédulas de Identidad N° 10107184 y 13967244, como requisito previo para optar al título de **Ingeniero Electricista**, de la Promoción correspondiente al **Semestre A-2012**, reunidos el día jueves, 08 de noviembre de 2012, después de la presentación oral y defensa pública en el **Salón de Reuniones de seminario de la Escuela de Ingeniería Eléctrica** de la Universidad de Los Andes y considerando que se han cumplido todos los requisitos exigidos en el Reglamento de Trabajo de Grado de la Escuela de Ingeniería Eléctrica para su elaboración y presentación, lo damos como **APROBADO** con una calificación de veinte **(20) Puntos. Mención Publicación.**

Prof. Rubén Medina
Jurado



Prof. José Luis Paredes
Tutor

Prof. Emiro Ibarra
Jurado

En Mérida, 07 de noviembre de 2012

Prof. Pablo Pérez Gamboa
Coordinador de Trabajo de Grado



ESCUELA DE INGENIERÍA ELÉCTRICA

Núcleo Universitario "Pedro Rincón Gutiérrez". La Hechicera.
Edificio "B" Facultad de Ingeniería. Nivel 4, ala Este (4E-03). Mérida 5101 A. Venezuela.
Teléfono: (+58) 274 240.2891. Fax: 240.2890.
www.ula.ve/ingenieria/electrica

Nuestro miedo más profundo no es que seamos inadecuados. Nuestro miedo más profundo es que somos inmensamente poderosos. Es nuestra luz, y no la oscuridad, lo que nos asusta. Nos preguntamos: ¿Quién soy yo para ser brillante, precioso, talentoso y fabuloso? En realidad, ¿Quién eres tú para no serlo? Eres hijo de Dios. Jugar a ser pequeño no sirve al mundo.

MARIANNE WILLIAMSON - A RETURN TO LOVE, 1992

Por ustedes, mamá, papá y hermanas

AGRADECIMIENTOS

Al Señor Dios Todopoderoso, sin Él nada es posible, por darme las virtudes más valiosas del planeta: el amor y la inteligencia.

A mi familia: madre, padre y hermanas, porque hemos tenido que luchar juntos y separados por nuestros sueños, pero siempre han estado allí demostrando su amor, preocupación y mejores deseos para que yo, quizá no el mejor hijo y hermano, cumpla los suyos, ¡les amo, gracias eternas!

A la ilustre Universidad de Los Andes, por permitirme unos años maravillosos de estudiante de la mano de mi muy especial Escuela de Ingeniería Eléctrica, cuyos profesores son testimonio de profesionalismo y humanidad, en cuyo personal técnico y administrativo se puede encontrar a los mejores amigos y sobre todo en cuyos estudiantes se tiene a los mejores compañeros y amigos que se pueda pedir, ¡gracias!

A mi gran amigo y tutor José Luis Paredes, por enseñarme mucho más que cómo ser un buen profesional, por enseñarme humildad y amor por la gente, y sobre todo por confiar en mí este tema de tesis el cual me ha beneficiado de una manera inconmensurable.

A Almer Becerra, porque su trabajo fue clave en la culminación de éste, por ser un gran amigo y compañero. También, a mis demás amigos, de la infancia, de la adolescencia y de esta adultez maravillosa llamada universidad, no les nombro porque son muchos, sepan que jamás hubiese sido una aventura valiosa sin su amistad.

A mis profesores Juan Marcos Ramírez, Pablo Pérez Gamboa, Rubén Medina y Emiro Ibarra, cuyos consejos han tenido un valor importantísimo para el desarrollo de este trabajo.

Al Consejo de Desarrollo Científico Humanístico, Tecnológico y de las Artes (CDCHTA) por el financiamiento parcial de este trabajo a través del proyecto I-1336-12-02-B.

Jerick E. Ordenes Sepúlveda

Jerick E. Ordenes Sepúlveda. Restauración de imágenes usando técnicas de reconstrucción de sensado comprimido. Universidad de Los Andes. Tutor(es): Prof. José L. Paredes, Prof. Juan M. Ramírez. Noviembre, 2012.

Resumen

En las últimas décadas el procesamiento digital de imágenes ha cobrado una especial importancia en distintas áreas de la ciencia, progresivamente se han venido desarrollando distintos métodos que aprovechan los avances computacionales para realizar procesamientos más robustos de distintas clases de datos. Este proyecto se ha centrado en el estudio de métodos de representación poco densa de señales aplicados a la restauración de imágenes y en esencia a la restauración frente a píxeles perdidos o a ruido de naturaleza impulsiva. Se propone el uso de un algoritmo robusto de reconstrucción propuesto recientemente en el área de *Sensado Comprimido* denominado *Weighted Median Regression* (WMR), con el que se abordaría el problema de restauración resolviendo un problema de regresión de desviación absoluta mínima regularizado por seudonorma ℓ_0 (ℓ_0 -LAD), tomando en cuenta tanto el conocimiento de la posición de los píxeles perdidos o del ruido impulsivo (dato habitual en restauración de imágenes) como el desconocimiento total de dicha información. Además, se estudia el comportamiento de este método de acuerdo a la utilización de diccionarios completos y redundantes de bases ortogonales en el problema de regresión, así como también la utilización de diccionarios entrenados a través del método *K times Singular Value Decomposition* (K-SVD), y se compara su desempeño con respecto a los métodos de filtrado convencional y el algoritmo de reconstrucción de señales poco densas *Orthogonal Matching Pursuit*.

Descriptor: restauración de imágenes, mediana ponderada regresiva, K-SVD, ruido impulsivo, píxeles perdidos.

ÍNDICE GENERAL

APROBACIÓN	ii
DEDICATORIA	iii
AGRADECIMIENTOS	iv
RESUMEN	v
INTRODUCCIÓN	1
Capítulo	pp.
1. MARCO TEÓRICO	4
1.1 REPRESENTACIÓN POCO DENSA DE SEÑALES	4
1.1.1 El Diccionario	7
1.1.1.1 <i>Diccionario DCT completo</i>	7
1.1.1.2 <i>Diccionario DCT sobrecompleto</i>	11
1.1.1.3 <i>Diccionarios entrenados</i>	13
1.1.1.4 <i>Diccionario holográfico</i>	13
1.2 K-SVD PARA ENTRENAMIENTO DE DICCIONARIOS	14
1.2.1 El algoritmo <i>K-means</i> y la Cuantización Vectorial	14
1.2.2 K-SVD como extensión de <i>K-means</i>	15
1.3 ALGORITMOS DE REPRESENTACIÓN POCO DENSA EN DICCIO- NARIOS REDUNDANTES	18
1.3.1 Algoritmo <i>Orthogonal Matching Pursuit</i> (OMP)	18
1.3.2 Algoritmo <i>Weighted Median Regression</i> (WMR)	20
2. RESTAURACIÓN DE IMÁGENES	22
2.1 FUNDAMENTOS DE LA RESTAURACIÓN DE IMÁGENES	22
2.2 RESTAURACIÓN DE IMÁGENES USANDO REPRESENTACIÓN PO- CO DENSA DE SEÑALES	24
2.2.1 Procesamiento de bloques solapados contra no solapados	26
2.2.2 Detección y remoción de ruido impulsivo	29
3. RESTAURACIÓN DE IMÁGENES BASADA EN WMR	32
3.1 WMR COMO MÉTODO ROBUSTO DE RESTAURACIÓN	32
3.1.1 Aproximación robusta por la norma ℓ_1	32
3.2 METODOLOGÍA APLICADA	40
3.2.1 Planteamiento matemático	40

4. SIMULACIÓN Y RESULTADOS	43
4.1 INTERPOLACIÓN DE PÍXELES PERDIDOS	43
4.1.1 Condiciones de procesamiento	44
4.1.2 Curvas de desempeño y simulaciones	46
4.2 SIMULACIONES PARA OTRAS CONTAMINACIONES	58
CONCLUSIONES	65
RECOMENDACIONES	67
APÉNDICE A. DEFINICIONES DE ERRORES	69
REFERENCIAS	70

ÍNDICE DE TABLAS

Tabla		pp.
1.1	Algoritmo K-SVD para entrenamiento de diccionarios	17
1.2	Algoritmo <i>Orthogonal Matching Pursuit</i>	19
1.3	Algoritmo <i>Weighted Median Regression</i>	21

ÍNDICE DE FIGURAS

Figura	pp.
1.1 Bases del diccionario DCT completo bidimensional	11
1.2 Bases del diccionario DCT sobrecompleto bidimensional	12
2.1 Proceso de degradación y restauración de imágenes	23
2.2 Pérdida de píxeles a través de la matriz \mathbf{M}	25
2.3 Esquema de restauración por bloques de una imagen	27
2.4 Restauración de imagen usando OMP para 30 % de píxeles perdidos	28
2.5 Esquema de restauración de imágenes contaminadas con ruido impulsivo	29
2.6 Imagen de entrenamiento y Diccionario entrenado para la imagen Peppers	30
2.7 Restauración de imagen usando OMP para 30 % de píxeles perdidos y usando diccionario K-SVD	31
3.1 Señal sintética y proyecciones	35
3.2 Reconstrucción de una señal sintética contaminada con impulsos y sin el uso del diccionario holográfico \mathbf{V}	36
3.3 Reconstrucción de una señal sintética contaminada con impulsos usando el diccionario holográfico \mathbf{V}	37
3.4 Reconstrucción de una señal sintética con pérdida de datos sin uso del diccionario holográfico \mathbf{V}	38
3.5 Reconstrucción de una señal sintética con pérdida de datos usando el diccionario holográfico \mathbf{V}	39
3.6 Distribución de $W_{i,j}$ para los píxeles en una imagen procesada con solapamiento completo de bloques	42
4.1 Imagen de prueba Lena e imagen de entrenamiento	45
4.2 Diccionarios seleccionados	45
4.3 Desempeño del algoritmo OMP con distintos diccionarios, sin solapar bloques durante el procesamiento	48
4.4 Desempeño del algoritmo OMP con distintos diccionarios, usando solapamiento completo de bloques en el procesamiento	49
4.5 Desempeño del algoritmo WMR con distintos diccionarios, sin solapar bloques durante el procesamiento	50
4.6 Desempeño del algoritmo WMR con distintos diccionarios, usando solapamiento completo de bloques en el procesamiento	51

4.7	Desempeño de los algoritmos WMR y OMP usando diccionario entrenado por K-SVD sin solapamiento de bloques, y desempeño de los métodos ACWMF y CWMF	53
4.8	Desempeño de los algoritmos WMR y OMP usando diccionario entrenado por K-SVD usando solapamiento completo de bloques, y desempeño de los métodos ACWMF y CWMF	54
4.9	Restauración completa de imagen para 20 % de píxeles perdidos	56
4.10	Restauración final de imagen para 20 % de píxeles perdidos	57
4.11	Máscara de texto aplicada en la prueba de desoclusión de imagen	58
4.12	Prueba de desoclusión para una imagen enmascarada con texto	59
4.13	Restauración completa de imagen para 20 % de píxeles perdidos sin uso del diccionario holográfico	61
4.14	Restauración de imagen con 10 % ruido impulsivo simétrico de tipo “sal y pimienta”	62
4.15	Restauración de imagen con ruido Gaussiano aditivo blanco con $SNR = 15\text{dB}$	64

INTRODUCCIÓN

Las técnicas de restauración de imágenes intentan mejorar la calidad de la imagen a partir de una versión degradada y contaminada de la misma. Esto es, dada una imagen observada \mathbf{y}_{obs} , la cual puede ser modelada de forma general como: $\mathbf{y}_{obs} = \mathbf{M}\mathbf{y} + \boldsymbol{\eta}$, donde \mathbf{y} es la imagen ideal, \mathbf{M} es la matriz de degradación y $\boldsymbol{\eta}$ es el ruido de contaminación propia del proceso de adquisición. En general se considera que \mathbf{y} tiene una representación poco densa en un dominio expandido por las bases de una transformada apropiada, por ejemplo: Fourier, DCT, *wavelet*, Hadamard. Así, $\mathbf{y} = \mathbf{A}\mathbf{x}$ donde \mathbf{A} es la matriz de transformación y \mathbf{x} los coeficientes que representan \mathbf{y} en el dominio expandido por las columnas de \mathbf{A} . Determinar \mathbf{y} a partir de \mathbf{y}_{obs} nos guía a resolver un problema inverso el cual es particularmente interesante debido a que \mathbf{A} es no invertible o mal condicionada a la presencia de la contaminación. Por lo que se recurre a técnicas de regularización en la cual un estimado de la señal ideal es determinado como $\hat{\mathbf{y}} = \mathbf{A}\hat{\mathbf{x}}$, donde $\hat{\mathbf{x}}$ es la solución al problema de optimización dado por:

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{M}\mathbf{A}\mathbf{x}\|_{\ell_p} + \tau \|\mathbf{x}\|_{\ell_0} \quad (1)$$

Con la introducción de las nuevas técnicas de sensado comprimido [1–3] se han desarrollado una gran variedad de herramientas que permiten resolver el problema de optimización descrito por (1), ver por ejemplo referencias [4–7].

En este proyecto de grado se usarán las técnicas de reconstrucción de sensado comprimido en el proceso de restauración de imágenes degradadas durante el proceso de captura. En principio se implementarán los métodos desarrollados por terceros investigadores y se propondrán variantes que contribuyan a mejorar su desempeño. Adicionalmente, se aplicarán los métodos de restauración implementados al procesamiento de imágenes deterioradas por la pérdida de píxeles.

Como objetivos específicos de este trabajo se tienen los siguientes:

- Familiarizarse con los conceptos de Sensado Comprimido.
- Estudiar los algoritmos de reconstrucción de señales poco densas a partir de su proyección aleatoria.
- Estudiar la transformada discreta del coseno (DCT), como herramienta de representación poco densa de imágenes.
- Estudiar la representación de imágenes en diccionarios sobrecompletos y en diccionarios entrenados a partir de un banco de imágenes.
- Aplicar el concepto de sensado comprimido a imágenes representadas en el dominio DCT, diccionarios sobrecompletos y diccionarios entrenados.
- Adaptar el algoritmo reconstrucción de sensado comprimido basado en mediana ponderada regresiva a la restauración de imágenes deterioradas por la no adquisición de algunos de sus píxeles.
- Comparar el desempeño del algoritmo implementado con algoritmos propuestos por terceros autores, que incluye, entre otros: filtrado no adaptativo, filtrado adaptativo y representación poco densa minimizando la norma ℓ_2 .

Este trabajo se organiza en cuatro capítulos. En el Capítulo 1 se introducen los conceptos básicos del problema de representación poco densa de señales e imágenes, se describen brevemente los algoritmos de reconstrucción de señales poco densas *Orthogonal Matching Pursuit* y *Weighted Median Regression* y el método de entrenamiento de diccionarios K-SVD. En el Capítulo 2 se describe el problema de restauración de imágenes usando técnicas de reconstrucción de representación poco densa de señales, en particular se describen métodos simples de procesamiento basados en *Orthogonal Matching Pursuit* y enfocados a la detección y remoción de ruido impulsivo. En el Capítulo 3 se presenta el algoritmo *Weighted Median Regression* como una alternativa robusta para la restauración de imágenes, se describe el problema de restauración en función de la naturaleza matemática de este método y se describe un procedimiento sencillo para realizar el procesamiento de

imágenes. Posteriormente, en el Capítulo 4 se presentan simulaciones y resultados sobre imágenes reales de los métodos estudiados y los métodos propuestos, utilizando diccionarios completos y redundantes de bases ortogonales y diccionarios entrenados. Por último, se presentan algunas conclusiones acerca del desempeño de los métodos propuestos y algunas recomendaciones para mejoras e investigaciones posteriores.

CAPÍTULO 1

MARCO TEÓRICO

En el área de Procesamiento de Señales e Imágenes se han estudiado de forma exhaustiva un amplio conjunto de procedimientos, tanto lineales como no-lineales, diseñados para realizar restauraciones sobre señales perturbadas basados en la teoría de Sensado Comprimido [1–3]. En esta área, el universo de conocimientos es muy amplio, por tanto, a lo largo de este capítulo se describirán brevemente la terminología y los conceptos relacionados con el problema abordado en el presente trabajo. Se iniciará con los conceptos de: *representación poco densa de señales* (del inglés *Sparse Signal Representation*) como enfoque de la teoría de Sensado Comprimido, se describirán los algoritmos de reconstrucción de señales poco densas más relevantes para este trabajo que han sido propuestos en la literatura, y la metodología y nociones básicas del entrenamiento de diccionarios basado en *K times Singular Value Decomposition* (K-SVD), como método seleccionado para construir diccionarios adaptados a las imágenes en proceso de restauración.

1.1 REPRESENTACIÓN POCO DENSA DE SEÑALES

Comúnmente las señales utilizadas en aplicaciones prácticas se pueden representar o aproximar, como una combinación lineal de un grupo reducido de formas de onda elementales. Esto, tiene como objetivo fundamental realzar o mostrar ciertas características de una señal de interés en un conjunto reducido de elementos no nulos.

Considere una señal \mathbf{y} como un vector de n componentes (i. e., $\mathbf{y} \in \mathbb{R}^n$), que puede ser escrita como la superposición de algunas señales elementales $\mathbf{a}_j \in \mathbb{R}^n$ parametrizadas por $j = \{1, \dots, K\}$ tal que

$$\mathbf{y} = \sum_{j=1}^K x_j \mathbf{a}_j \quad (1.1)$$

Donde $K \geq n$. El diccionario $\mathbf{A} = [\mathbf{a}_1 \mid \mathbf{a}_2 \mid \dots \mid \mathbf{a}_K]$ es la matriz $n \times K$ en cuyas columnas están dispuestas las señales elementales o átomos $\{\mathbf{a}_j\}_{j=1}^K$, que son asumidos normalizados con una norma ℓ_2 unitaria, y $\{x_j\}_{j=1}^K$ es el conjunto de coeficientes de representación de la señal \mathbf{y} que definen el vector $\mathbf{x} \in \mathbb{R}^K$. Una representación matricial de (1.1) es

$$\mathbf{y} = \mathbf{A}\mathbf{x}, \quad (1.2)$$

que describe un sistema de ecuaciones lineales con K incógnitas y n ecuaciones, donde \mathbf{y} y \mathbf{A} se asumen conocidos. Usualmente $n \ll K$, por esto el sistema de ecuaciones es subdeterminado por lo que múltiples soluciones son posibles.

En procesamiento de imágenes usualmente se considera una imagen $\mathbf{Y} \in \mathbb{R}^{\sqrt{n} \times \sqrt{n}}$, que ordenada lexicográficamente¹ puede ser representada como un vector columna $\mathbf{y} \in \mathbb{R}^n$. El objetivo principal en la representación poco densa de la imagen \mathbf{y} es encontrar en el diccionario \mathbf{A} el conjunto de coeficientes $\mathbf{x} = [x_1, x_2, \dots, x_K]^\dagger$, donde $\{\cdot\}^\dagger$ denota el operador *transposición*, que en lo posible tenga el menor número de coeficientes no nulos (poco denso²) y que reconstruya o aproxime lo mejor posible la señal \mathbf{y} como una combinación lineal de algunos pocos átomos del diccionario \mathbf{A} . Esto se traduce inevitablemente en encontrar una solución al problema inverso dado en la ecuación (1.2).

Una medida de poca densidad se puede cuantificar usando la seudonorma ℓ_0 del vector \mathbf{x} , denotada como $\|\mathbf{x}\|_{\ell_0}$, que consiste simplemente en contar la cantidad de elementos no nulos (distintos de cero) del vector en su argumento.

Observe que si el diccionario \mathbf{A} es una matriz ortogonal (base ortogonal) de $K = n$ elementos la representación es única y dada por $\mathbf{x} = \mathbf{A}^\dagger \mathbf{y}$. Sin embargo, la representación \mathbf{x} en un diccionario redundante ($K \gg n$) con el menor número de elementos distintos de cero posibles es ciertamente una representación más atractiva, como lo explica J. Tropp en el Capítulo 2 de la referencia [8], y es la solución al siguiente problema de minimización:

$$\arg \min_{\mathbf{x}} \|\mathbf{x}\|_{\ell_0} \quad \text{sujeto a} \quad \mathbf{y} = \mathbf{A}\mathbf{x} \quad (1.3)$$

¹Orden lexicográfico consiste en reordenar un arreglo bidimensional de datos (matriz) como un vector; a través de la concatenación de sus columnas.

²Del inglés *Sparsity* (escasez, poca densidad).

Una visión más realista de este problema es considerar a la señal de interés como una combinación lineal de señales elementales, que ha sido contaminada con ruido aditivo $\boldsymbol{\eta}$. Así, $\mathbf{y} = \mathbf{A}\mathbf{x} + \boldsymbol{\eta}$, para lo cual, la representación poco densa se reduce a un problema de aproximación, es decir, se desea encontrar \mathbf{x} sujeto a que tenga el menor número de elementos no nulos tal que $\mathbf{y} \approx \mathbf{A}\mathbf{x}$. Incorporando esto en (1.3), tenemos un problema de minimización más flexible descrito por

$$\arg \min_{\mathbf{x}} \|\mathbf{x}\|_{\ell_0} \quad \text{sujeto a} \quad \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_{\ell_p} \leq \varepsilon \quad (1.4)$$

Donde ε representa una desviación o error permisible, bajo la norma general ℓ_p , entre las muestras de la señal de interés y la aproximación a través de los coeficientes estimados. Usualmente las normas utilizadas para hacer el cálculo de la desviación son las normas ℓ_p para $p = 1, 2$ y ∞ , para mayores detalles se invita al lector a revisar el Capítulo 6 de la referencia [9].

La ecuación (1.4) muestra un problema de minimización restringido por una constante de error, usualmente denominado *error-constrained approximation* (aproximación restringida por error). Otro enfoque del problema de representación poco densa es intentar estimar la mejor representación de la señal utilizando una combinación de M átomos o menos del diccionario, usualmente denominada *M-term approximation* (aproximación de M términos) o *sparsity-constrained approximation* (aproximación restringida por poca densidad) [8]. Usando este enfoque el problema de representación se reformula como el siguiente problema de optimización:

$$\arg \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_{\ell_p} \quad \text{sujeto a} \quad \|\mathbf{x}\|_{\ell_0} \leq M \quad (1.5)$$

En ambos casos, la extracción de la representación más dispersa implica la solución de un problema de optimización no convexo debido al uso de la seudonorma ℓ_0 , lo que se reduce a un problema de complejidad no polinomial o NP-completo³, por consiguiente es necesario recurrir a algoritmos de optimización [5, 6] para obtener soluciones aproximadas a estos problemas. Estos algoritmos han sido extensivamente investigados y recopilados en

³Del inglés NP-*hard*, véase la referencia [4] para profundizar al respecto.

la literatura, más adelante se hace referencia a los utilizados en este proyecto.

1.1.1 El Diccionario

En todos los algoritmos que resuelven el problema de representación, se asume inicialmente conocido y fijo el diccionario \mathbf{A} . Por lo que la tarea de selección del diccionario que mejor represente la señal de interés es muy importante, ya sea que éste sea un diccionario preespecificado (formado posiblemente por una base ortogonal o combinación de bases ortogonales) o diseñado para ajustarse de la mejor manera posible al modelo de poca densidad establecido y definido en función de los datos que se representarán.

Como se mencionó anteriormente, un diccionario es una colección finita de señales elementales de norma unitaria. Si el diccionario se extiende por todo el espacio de la señal y $K \geq n$ se dice que el diccionario es *completo* o *total* [8], y en este caso una señal se puede aproximar con cero error usando una combinación lineal de átomos $\{\mathbf{a}_i\}$, para que el diccionario sea sobrecompleto o redundante se debe cumplir $K > n$, considerando que existe alguna dependencia lineal entre los átomos. En muchas situaciones prácticas los diccionarios redundantes son construidos uniendo uno o varios elementos bases de transformadas suficientemente incoherentes⁴ correspondientes a una base ortogonal, lo que es relativamente simple. Este trabajo se centra en la utilización de diccionarios provenientes de una transformada de bases ortogonales como la *Transformada Discreta del Coseno* (DCT) y diccionarios entrenados a partir de los datos que se desean representar.

1.1.1.1 Diccionario DCT completo

La transformada discreta del coseno tiene la propiedad de representar la información visualmente significativa de una imagen en un conjunto reducido de coeficientes; por esto, es frecuentemente utilizada en aplicaciones de compresión de imágenes y video (e. g., algoritmos estándar JPEG, MPEG).

Se puede obtener un diccionario completo DCT para señales unidimensionales como la transpuesta de la matriz de transformada discreta del coseno \mathbf{T} para una señal de entrada

⁴Se dice que dos bases son incoherentes cuando los elementos bases de la transformada no pueden ser representados como combinación lineal de algunos pocos elementos bases de la otra transformada.

$\mathbf{y} \in \mathbb{R}^N$, definida como

$$T_{i,j} = \begin{cases} \frac{1}{\sqrt{N}} & i = 0, \quad 0 \leq j \leq N - 1 \\ \sqrt{\frac{2}{N}} \cos\left(\frac{\pi(2j+1)i}{2N}\right) & 1 \leq i \leq N - 1, \quad 0 \leq j \leq N - 1 \end{cases} \quad (1.6)$$

Siendo \mathbf{T} ortogonal y normalizada en sus filas a una norma ℓ_2 unitaria.

Así, la transformada discreta del coseno unidimensional \mathbf{x} de \mathbf{y} está definida por

$$\mathbf{x} = \mathbf{T}\mathbf{y} \quad (1.7)$$

Debido a que \mathbf{T} es ortogonal, lo que implica que su inversa es igual a su transpuesta $\mathbf{T}^{-1} = \mathbf{T}^\dagger$, se puede determinar la transformada inversa fácilmente como

$$\mathbf{y} = \mathbf{T}^\dagger \mathbf{x} \quad (1.8)$$

Siendo esta última la ecuación de interés, ya que $\mathbf{x} \in \mathbb{R}^N$ contiene los coeficientes de representación correspondientes a \mathbf{y} , que usualmente se busca sea lo más poco denso posible. Ambas representaciones, \mathbf{y} y \mathbf{x} son equivalentes pero en dominios distintos.

Así, podemos definir un diccionario usando las bases de la transformada DCT para señales unidimensionales: $\mathbf{A}_1 = \mathbf{T}^\dagger$. Por tanto, un diccionario de bases DCT es una matriz cuyas columnas (átomos) representan señales coseno muestreadas a distintas frecuencias. Para el diccionario completo se mantiene la propiedad de ortogonalidad entre los átomos, es decir, no existe una dependencia lineal entre ellos.

Ahora bien, en este proyecto las señales en cuestión son imágenes, que son señales con variaciones tanto en la dirección vertical como en la horizontal para las que un diccionario como \mathbf{A}_1 no es apropiado. Esto es, al plantear el problema inverso de la DCT en (1.8) con la señal de entrada como una imagen $\mathbf{Y} \in \mathbb{R}^{N \times M}$ donde cada columna es tratada como una señal unidimensional independiente, se tendría ahora una matriz de coeficientes $\mathbf{X} \in \mathbb{R}^{N \times M}$ cuyas columnas corresponden a la representación de la columnas de \mathbf{Y} en el dominio DCT unidimensional, explotando solo la correlación vertical en cada columna de la imagen.

Se deben emplear entonces diccionarios que consideren las variaciones horizontales y verticales simultáneamente de la imagen en cuestión. Para esto se parte de las definiciones de la transformada DCT bidimensional, entre la cuales se puede encontrar una definición dependiente de la propiedad de separabilidad de la transformada, como describe S. A. Khayam en [10], de la transformada unidimensional \mathbf{T} definida en (1.6).

Así, la transformada bidimensional se puede obtener como

$$\mathbf{X} = \mathbf{T}\mathbf{Y}\mathbf{T}^\dagger, \quad (1.9)$$

y la transformada inversa tal como

$$\mathbf{Y} = \mathbf{T}^\dagger\mathbf{X}\mathbf{T} \quad (1.10)$$

Sin embargo, no se puede obtener directamente de (1.10) un diccionario DCT bidimensional \mathbf{A} . Por esto, y haciendo uso de ordenamiento lexicográfico de la imagen \mathbf{Y} , se puede demostrar que un diccionario DCT bidimensional $\mathbf{A} \in \mathbb{R}^{n \times n}$ puede ser obtenido a partir del *producto Kronecker* entre un diccionario DCT unidimensional \mathbf{A}_1 de tamaño $\sqrt{n} \times \sqrt{n}$ y sí mismo, esto es $\mathbf{A}_1 \otimes \mathbf{A}_1$, donde \otimes denota dicho producto definido como:

Definición 1.1. *Producto Kronecker.* Sea \mathbb{R} el conjunto de los números reales. El producto Kronecker de una matriz $\mathbf{A} = [a_{i,j}] \in \mathbb{R}^{m \times n}$ y una matriz $\mathbf{B} = [b_{i,j}] \in \mathbb{R}^{p \times q}$ está definido como una matriz bloque $\mathbf{C} \in \mathbb{R}^{mp \times nq}$, tal que

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{1,1}\mathbf{B} & \dots & a_{1,n}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{m,1}\mathbf{B} & \dots & a_{m,n}\mathbf{B} \end{bmatrix} = \mathbf{C} \quad (1.11)$$

Cada bloque de \mathbf{C} es una submatriz formada por el producto entre cada elemento escalar $a_{i,j}$ (de la matriz \mathbf{A}) y la matriz \mathbf{B} .

Por ejemplo, para una matriz \mathbf{A} de 2×2 y \mathbf{B} de 2×3 se tiene la matriz \mathbf{C} de dimensión

4×6 , dada por:

$$\begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix} \otimes \begin{bmatrix} b_{1,1} & b_{1,2} & b_{1,3} \\ b_{2,1} & b_{2,2} & b_{2,3} \end{bmatrix} = \begin{bmatrix} a_{1,1}b_{1,1} & a_{1,1}b_{1,2} & a_{1,1}b_{1,3} & a_{1,2}b_{1,1} & a_{1,2}b_{1,2} & a_{1,2}b_{1,3} \\ a_{1,1}b_{2,1} & a_{1,1}b_{2,2} & a_{1,1}b_{2,3} & a_{1,2}b_{2,1} & a_{1,2}b_{2,2} & a_{1,2}b_{2,3} \\ a_{2,1}b_{1,1} & a_{2,1}b_{1,2} & a_{2,1}b_{1,3} & a_{2,2}b_{1,1} & a_{2,2}b_{1,2} & a_{2,2}b_{1,3} \\ a_{2,1}b_{2,1} & a_{2,1}b_{2,2} & a_{2,1}b_{2,3} & a_{2,2}b_{2,1} & a_{2,2}b_{2,2} & a_{2,2}b_{2,3} \end{bmatrix}$$

Ahora bien, teniendo en cuenta la siguiente propiedad del producto Kronecker [11]:

Propiedad 1.1. Para $\mathbf{PRQ} = \mathbf{C}$, donde \mathbf{P} , \mathbf{R} y \mathbf{Q} son matrices conocidas y \mathbf{C} es la incógnita

$$\text{vec}(\mathbf{PRQ}) = (\mathbf{Q}^\dagger \otimes \mathbf{P})\text{vec}(\mathbf{R}) = \text{vec}(\mathbf{C}) \quad (1.12)$$

Donde $\text{vec}(\cdot)$ es el ordenamiento lexicográfico de la matriz en el argumento.

Siendo $N = \sqrt{n}$ para la ecuación (1.6), y relacionando (1.12) con (1.10), teniendo $\mathbf{P} = \mathbf{T}^\dagger$, $\mathbf{R} = \mathbf{X}$, $\mathbf{Q} = \mathbf{T}$ y $\mathbf{C} = \mathbf{Y}$, se tiene

$$\text{vec}(\mathbf{Y}) = (\mathbf{T}^\dagger \otimes \mathbf{T}^\dagger)\text{vec}(\mathbf{X}). \quad (1.13)$$

Ya que se cumple que $\mathbf{A}_1 = \mathbf{T}^\dagger$ se redefine la ecuación anterior como

$$\text{vec}(\mathbf{Y}) = (\mathbf{A}_1 \otimes \mathbf{A}_1)\text{vec}(\mathbf{X}), \quad (1.14)$$

dando paso a la definición del diccionario DCT bidimensional como

$$\mathbf{A} = \mathbf{A}_1 \otimes \mathbf{A}_1. \quad (1.15)$$

Las dimensiones deben ser consistentes, por lo tanto, la imagen \mathbf{Y} debe ser cuadrada de dimensión $\sqrt{n} \times \sqrt{n}$ para poder ser representada por los átomos del diccionario. Habitualmente se aplican los métodos de procesamiento sobre secciones de una imagen en estudio, comúnmente estas secciones son bloques de 8×8 o 16×16 píxeles, con el objeto de no producir diccionarios tan grandes para representar una imagen ordenada lexicográficamente. Típicamente una imagen de prueba estandarizada en el área de procesamiento

de señales es de 256×256 o 512×512 píxeles (múltiplos de 8 píxeles). Así, se puede construir un diccionario DCT bidimensional \mathbf{A} , que permita representar una imagen (bloque) $\mathbf{Y} \in \mathbb{R}^{8 \times 8}$ que ordenada lexicográficamente es un vector $\mathbf{y} \in \mathbb{R}^{64}$, a partir de un diccionario $\mathbf{A}_1 \in \mathbb{R}^{8 \times 8}$. Las 64 funciones base (átomos) del diccionario completo resultante $\mathbf{A} \in \mathbb{R}^{64 \times 64}$ están representadas en la figura 1.1, reordenadas de tal forma para apreciar la consideración de la variación vertical y horizontal de píxeles en cada una.

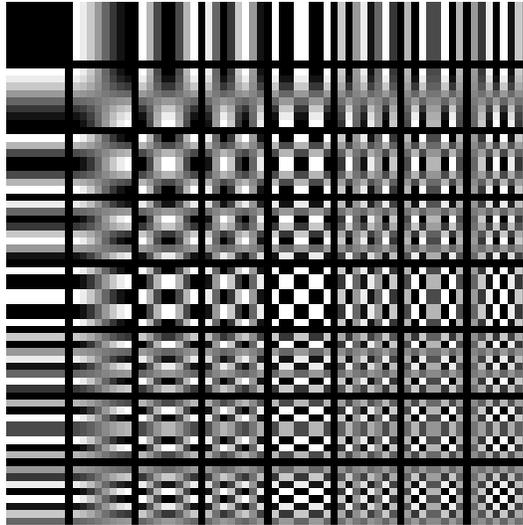


Figura 1.1. Bases DCT bidimensionales para matrices de tamaño 8×8 .

1.1.1.2 *Diccionario DCT sobrecompleto*

Los diccionarios sobrecompletos son utilizados en distintas áreas de investigación porque su uso puede derivar en aproximaciones más estables, más robustas y en descomposiciones más compactas, que usando diccionarios completos de bases ortogonales. Un diccionario DCT sobrecompleto, al igual que un diccionario completo, está compuesto por señales coseno a distintas frecuencias con un paso menor en frecuencia entre un átomo y otro, así se dispone de señales que permiten representar variaciones existentes entre dos átomos de un diccionario completo.

De forma análoga a un diccionario DCT completo bidimensional construido según (1.15), se puede utilizar un diccionario DCT sobrecompleto unidimensional $\mathbf{A}_1 \in \mathbb{R}^{\sqrt{n} \times \sqrt{K}}$ para construir un diccionario DCT sobrecompleto bidimensional separable $\mathbf{A} \in \mathbb{R}^{n \times K}$. La definición del diccionario DCT sobrecompleto unidimensional \mathbf{A}_1 de tamaño $N \times M$, con

$M > N$, está dada por la siguiente ecuación

$$A_{1,i,j} = \begin{cases} \frac{1}{\sqrt{N}} & 0 \leq i \leq N - 1, j = 0 \\ \cos\left(\frac{\pi i j}{M}\right) & 0 \leq i \leq N - 1, \text{ para } j = 1, 2, \dots, M - 1 \end{cases} \quad (1.16)$$

Esta ecuación para \mathbf{A}_1 es más simple que la ecuación (1.6) para la matriz de transformada completa DCT unidimensional, puesto que acá se debe extraer el valor medio y normalizar los átomos para $1 \leq j \leq M - 1$. Así, se puede construir un diccionario sobrecompleto bidimensional usando (1.15), tal que $\mathbf{A} = \mathbf{A}_1 \otimes \mathbf{A}_1$ con $N = \sqrt{n}$ y $M = \sqrt{K}$ en (1.16).

Como el diccionario es sobrecompleto se debe escoger el valor de K de acuerdo a las necesidades del problema. Valores usuales del número de átomos para representar una matriz (bloque de imagen) de 8×8 píxeles pueden ser $K = 256$ como lo usa M. Elad en las pruebas del Capítulo 15 en la referencia [12] o $K = 441$ en comparación con un diccionario Haar⁵ sobrecompleto como lo muestra M. Aharon y col. en [13]. En la figura 1.2 se puede apreciar el conjunto de átomos de un diccionario DCT bidimensional sobrecompleto y separable de 256 átomos para representar matrices de 8×8 píxeles.

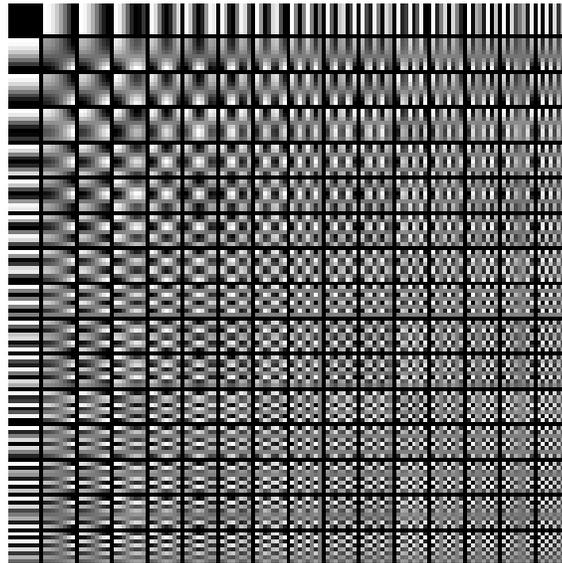


Figura 1.2. Bases del diccionario DCT sobrecompleto bidimensional, 256 átomos para matrices de tamaño 8×8 .

⁵Diccionario cuyos átomos provienen de las funciones *wavelet* Haar.

1.1.1.3 *Diccionarios entrenados*

La tarea de selección del diccionario que permita encontrar la representación de menor densidad de una señal de interés es fundamental. Numerosos estudios se han concentrado, en lugar de la selección, en el diseño de diccionarios a partir de conjuntos finitos de datos, es decir, que los átomos del diccionario se adapten a los datos que representarán.

El razonamiento de los distintos métodos de entrenamiento parte de tener un conjunto de señales⁶ $\mathbf{Y} = \{\mathbf{y}_i\}_{i=1}^N$ para las que existe un diccionario \mathbf{A} que permite representar dicho conjunto a través de combinaciones lineales de algunos de sus átomos, por ello se les conoce como *diccionarios basados en datos* (del inglés *data-driven dictionaries*). Sin embargo, la cantidad de posibles diccionarios que pueden representar \mathbf{Y} con efectividad es elevada, por esto, los distintos métodos de entrenamiento tratan de encontrar el diccionario \mathbf{A} que proporcione la mejor solución al problema (1.3) para cada vector \mathbf{y}_i .

En la literatura se han propuesto recientemente algoritmos de entrenamiento basados en métodos de *Cuantización Vectorial* (CV) o más específicamente en la generalización del algoritmo *K-means* para el diseño de *libro de códigos* (del inglés *codebook*) en el problema de CV, como MOD [14] y K-SVD [13], en este último se centra la atención en el presente trabajo y será descrito con cierto detalle en la Sección 1.2. Se pueden encontrar también métodos con enfoques probabilísticos como *Estimación por Máxima Verosimilitud* [15] o *Máximo A-Posteriori* [16] en el que se desarrollan funciones de probabilidad condicionada dependientes de \mathbf{A} y \mathbf{Y} , o métodos de unión de bases ortogonales [17]. El proceso de entrenamiento, partiendo de la relación (1.3), involucra como mínimo dos etapas, una de representación poco densa de cada señal de entrenamiento y otra de actualización de los átomos del diccionario hasta alcanzar cierta convergencia en las relaciones dependientes de la calidad de la representación de \mathbf{Y} .

1.1.1.4 *Diccionario holográfico*

El diccionario holográfico es un concepto íntimamente relacionado con la teoría de Sensado Comprimido (*Compressive Sensing*) [1–3], diseñada para recuperar señales de cierta dimensión a partir de señales de menor dimensión a través de la resolución de

⁶Abusando de la notación \mathbf{Y} se refiere aquí a un conjunto de señales, aunque en notaciones posteriores hace referencia a una matriz cuyos vectores columnas son las señales \mathbf{y}_i para $i = 1, 2, \dots, N$.

problemas de representación poco densa de señales, como los estudiados anteriormente.

Un diccionario holográfico \mathbf{V} es simplemente una proyección del diccionario \mathbf{A} sobre una matriz de medición \mathbf{M} , así

$$\mathbf{V} = \mathbf{M}\mathbf{A} = \left[\mathbf{v}_1 \mid \mathbf{v}_2 \mid \dots \mid \mathbf{v}_K \right]. \quad (1.17)$$

Donde los elementos $\{\mathbf{v}_i\}_{i=1}^K$ son el conjunto de átomos del diccionario holográfico, cada uno de dimensión $L \ll K$ elementos. La matriz de medición \mathbf{M} , en pocas palabras, es aquella que permite reducir una señal de dimensión elevada (proveniente, por ejemplo, de un fenómeno físico) en una señal de menor dimensión, que contiene la información y características necesarias para poder reconstruir con gran precisión la señal original resolviendo el problema inverso dado por (1.2).

1.2 K-SVD PARA ENTRENAMIENTO DE DICCIONARIOS

El algoritmo K-SVD, por *K times Singular Value Decomposition* (descomposición en valores singulares K veces), es un método recientemente propuesto para el entrenamiento o diseño de diccionarios sobrecompletos, desarrollado por M. Aharon y col. en [13]. Es una generalización del algoritmo *K-means* para entrenamiento de libro de códigos en CV, que introduce la solución de un problema de descomposición en valores singulares como proceso de actualización de los átomos del diccionario.

1.2.1 El algoritmo *K-means* y la Cuantización Vectorial

Partiendo de CV se tiene un libro de códigos $\mathbf{C} = \left[\mathbf{c}_1 \mid \mathbf{c}_2 \mid \dots \mid \mathbf{c}_K \right]$ de K palabras código, que puede ser usado para representar un conjunto de vectores denotados como $\mathbf{Y} = \{\mathbf{y}_i\}_{i=1}^N$ con $N \gg K$. Cada señal es aproximadamente representada como la palabra código del diccionario o libro de códigos más cercana en distancia euclidiana. Así, se puede escribir $\mathbf{y}'_i = \mathbf{C}\mathbf{x}_i$ como la aproximación a la señal \mathbf{y}_i , donde $\mathbf{x}_i = \mathbf{e}_j$ es un vector con todas las componentes iguales a cero excepto la componente en la j -ésima posición. Este problema es observado como un claro ejemplo del problema de representación poco densa de señales bajo ciertas condiciones extremas, esto es, que solo un átomo tiene participación en la

señal y su coeficiente es forzado a ser igual a 1.

Como se ha mencionado, el problema de entrenamiento se reduce a la búsqueda del diccionario (libro de códigos) que mejor represente el conjunto de señales dado. Así, en CV el entrenamiento se reduce a la búsqueda de un \mathbf{C} que minimize cierta función de error sujeta a la estructura de \mathbf{X} , que está conformado por bases triviales. Este problema de minimización está descrito por

$$\arg \min_{\mathbf{C}, \mathbf{x}} \|\mathbf{Y} - \mathbf{C}\mathbf{X}\|_F^2, \quad \text{sujeto a } \forall i \mathbf{x}_i = \mathbf{e}_k \text{ para algún } k, \quad (1.18)$$

donde $\|\cdot\|_F$ denota la *norma Frobenius* o *norma Hilbert-Schmidt* definida como

$$\|\mathbf{A}\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |A_{i,j}|^2}, \quad (1.19)$$

que es una extensión del concepto de norma vectorial ℓ_2 para el cálculo de la norma matricial utilizando el ordenamiento lexicográfico de la matriz $\mathbf{A} \in \mathbb{R}^{m \times n}$.

El algoritmo *K-means* o *K-times Mean computation* (K veces cómputo de media), algoritmo usado ampliamente en el área de análisis de agrupamiento de datos, aborda este problema de forma iterativa en dos etapas por iteración, una de codificación poco densa de señales y otra de actualización del libro de códigos. En la etapa de codificación se asume conocido y fijo el libro de códigos \mathbf{C} o $\mathbf{C}^{(J-1)}$, siendo J el índice de control de la iteración, y se resuelve el problema de minimización en (1.18) para cada vector columna de \mathbf{X} . En la siguiente etapa se asume fijo \mathbf{X} y se busca una actualización de \mathbf{C} que también resuelva (1.18). Este método está extensivamente estudiado y dispuesto en la literatura del área, se recomienda revisar el Capítulo 8 en [8], Capítulo 4 en [18] o el Capítulo 10 en la referencia [19] para información más detallada.

1.2.2 K-SVD como extensión de *K-means*

Se puede ver una relación explícita entre el problema de representación poco densa de señales y CV. En representación poco densa de señales una señal de entrada puede ser representada por la combinación de varias palabras código denominadas átomos o elementos del diccionario, además se permite a cada átomo tener más de un elemento

distinto de cero y que estos posean valores apropiados a fin de que se minimice en lo posible el error en la representación. De esta manera se puede reescribir el problema en (1.18) con un diccionario \mathbf{A} para un conjunto de señales de entrenamiento \mathbf{Y} , tal que

$$\arg \min_{\mathbf{A}, \mathbf{X}} \|\mathbf{Y} - \mathbf{A}\mathbf{X}\|_F^2, \quad \text{sujeto a } \forall i, \|\mathbf{x}_i\|_{\ell_0} \leq T_0 \quad (1.20)$$

teniendo en cuenta la estructura poco densa de cada vector \mathbf{x}_i . O el problema de optimización menos restrictivo restringido por un error residual objetivo

$$\arg \min_{\mathbf{A}, \mathbf{X}} \sum_i \|\mathbf{x}_i\|_{\ell_0}, \quad \text{sujeto a } \|\mathbf{Y} - \mathbf{A}\mathbf{X}\|_F^2 \leq \varepsilon \quad (1.21)$$

Siguiendo el esquema de *K-means*, este algoritmo aborda el problema en dos etapas, la etapa de representación poco densa que resuelve, a través de cualquier algoritmo de búsqueda (e. g., OMP [5], Batch OMP [20], MP [21], entre otros), N problemas de la forma

$$\arg \min_{\mathbf{x}_i} \|\mathbf{y}_i - \mathbf{A}\mathbf{x}_i\|_{\ell_2}^2, \quad \text{sujeto a } \|\mathbf{x}_i\|_{\ell_0} \leq T_0, \quad \text{para } i = 1, 2, \dots, N. \quad (1.22)$$

Y la etapa de actualización del diccionario, en la que incorpora una solución basada en *descomposición en valores singulares* de una matriz de error. Se recomienda al lector revisar la referencia [13] y sus referencias para la comprensión completa de este método. La Tabla 1.1 describe el algoritmo en cuestión para la solución de (1.20). En la etapa de inicialización del algoritmo se puede fijar el diccionario $\mathbf{A}^{(0)}$ como un diccionario preespecificado sobrecompleto, o escoger de forma aleatoria componentes del conjunto \mathbf{Y} como átomos iniciales. La notación \mathbf{x}_T^j denota la j -ésima fila de la matriz de coeficientes \mathbf{X} . Las rutinas propias de los desarrolladores de este método y utilizadas en este trabajo están disponibles en: <http://www.cs.technion.ac.il/~elad/software/> sitio web académico de M. Elad y <http://www.cs.technion.ac.il/~ronrubin/software.html> sitio web académico de R. Rubinstein.

Este algoritmo tiene un costo computacional que depende de la cantidad de componentes en \mathbf{Y} , la dimensión de la componente \mathbf{y}_i y la cantidad K de átomos del diccionario. R. Rubinstein y col. han propuesto en [20] una versión computacionalmente eficiente del algoritmo K-SVD que utiliza como algoritmo de optimización al algoritmo *Batch-OMP*;

Cuadro 1.1. Algoritmo K-SVD para entrenamiento de diccionarios.

ENTRADA	<ul style="list-style-type: none"> • Conjunto de señales de entrenamiento \mathbf{Y}. • Nivel de poca densidad T_0
INICIALIZACIÓN	<ul style="list-style-type: none"> • Fijar el diccionario $\mathbf{A}^{(0)} \in \mathbb{R}^{n \times K}$ con las columnas normalizadas bajo la norma ℓ_2 y el contador de iteración $J = 1$.
ITERACIÓN	<p>Repetir hasta alcanzar convergencia</p> <ul style="list-style-type: none"> • Etapa de codificación poco densa: Use cualquier algoritmo de búsqueda para calcular los vectores de representación \mathbf{x}_i para cada ejemplar \mathbf{y}_i, por aproximación de la solución de $\arg \min_{\mathbf{x}_i} \ \mathbf{y}_i - \mathbf{A}\mathbf{x}_i\ _{\ell_2}^2, \quad \text{sujeto a } \ \mathbf{x}_i\ _{\ell_0} \leq T_0, \text{ para } i = 1, 2, \dots, N.$ • Etapa de actualización del diccionario: Para cada columna $k = 1, 2, \dots, K$ en $\mathbf{A}^{(J-1)}$, actualizar por <ul style="list-style-type: none"> – Definir el grupo de ejemplares en \mathbf{Y} que usan este átomo, $\omega_k = \{i 1 \leq i \leq N, \mathbf{x}_T^k(i) \neq 0\}$. – Calcular la matriz de error de la representación general, \mathbf{E}_k, por $\mathbf{E}_k = \mathbf{Y} - \sum_{j \neq k} \mathbf{a}_j \mathbf{x}_T^j$ – Restringir \mathbf{E}_k escogiendo solo las columnas correspondientes a ω_k, y obtener \mathbf{E}_k^R. – Aplicar descomposición SVD $\mathbf{E}_k^R = \mathbf{U}\Delta\mathbf{V}^T$. Elegir la columna actualizada del diccionario \mathbf{a}_k como la primera columna de \mathbf{U}. Actualizar el vector de coeficientes \mathbf{x}_R^k como la primera columna de \mathbf{V} multiplicado por $\Delta(1,1)$. • Establezca $J = J + 1$. • Chequear criterio de parada.

Adaptado de *The K-SVD algorithm*, figura 2 en la referencia [13].

una versión de OMP diseñada para codificación poco densa de señales de gran dimensión.

La utilización de un diccionario entrenado a partir de un método como K-SVD en un proceso de restauración de imagen, introduce énfasis en la recuperación de características (correspondientes a la naturaleza de los datos) que han sido perdidas en los procesos de medición, como detalles, bordes, textura, curvas, etc. Además y de forma general, los métodos de entrenamiento pueden ser vistos como métodos de reducción de bases de datos de imágenes o señales, con el objeto de reducir cantidades significativas de información a cambio de costo computacional, dependiendo de las aplicaciones donde tenga utilidad significativa. En el siguiente capítulo se usa esta técnica de entrenamiento de diccionario

a la aplicación de restauración de imágenes.

1.3 ALGORITMOS DE REPRESENTACIÓN POCO DENSA EN DICCIONARIOS REDUNDANTES

En esta sección se describirán brevemente dos algoritmos diseñados para dar solución al problema de optimización relacionado con el problema de representación poco densa de señales, específicamente los problemas en (1.3), (1.4) y (1.5). El algoritmo *Orthogonal Matching Pursuit* [5], que es un método estándar de reconstrucción en el área de procesamiento de señales, y el algoritmo *Weighted Median Regression*, propuesto recientemente en la literatura [6].

1.3.1 Algoritmo *Orthogonal Matching Pursuit* (OMP)

El algoritmo OMP [4, 5] es un algoritmo de optimización iterativa de la clase búsqueda voraz (del inglés *greedy-pursuit*), que al igual que la versión no-ortogonal (algoritmo *Matching Pursuit*, denotado por MP) de la que se deriva, tiene como objetivo encontrar el átomo \mathbf{a}_j del diccionario \mathbf{A} que tenga la mayor correlación con la señal de entrada \mathbf{y} , remueve la contribución de dicho átomo en la señal definiendo una señal residuo \mathbf{r} , se busca luego el vector \mathbf{a}_j que mejor se ajuste a este residuo. Se procede de forma iterativa hasta alcanzar un criterio de parada o un número máximo de iteraciones preespecificado. La Tabla 1.2 resume el algoritmo OMP.

R. Rubinstein y col. explican en [20] que este algoritmo puede abordar específicamente dos problemas de optimización, el problema de aproximación restringida por poca densidad dado por:

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_{\ell_2}^2 \quad \text{sujeto a} \quad \|\mathbf{x}\|_{\ell_0} \leq T_0, \quad (1.23)$$

y el problema de aproximación restringida por error

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{x}\|_{\ell_0} \quad \text{sujeto a} \quad \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_{\ell_2}^2 \leq \varepsilon. \quad (1.24)$$

El uso de $\|\cdot\|_{\ell_2}^2$, como función objetivo, introduce un problema de mínimos cuadrados cuya solución permite encontrar la estimación de los coeficientes \mathbf{x} .

Tabla 1.2. Algoritmo *Orthogonal Matching Pursuit*.

ENTRADA	<ul style="list-style-type: none"> • Diccionario $\mathbf{A} \in \mathbb{R}^{n \times K}$. • Señal $\mathbf{y} \in \mathbb{R}^n$. • Número de elementos no nulos T_0
INICIALIZACIÓN	<ul style="list-style-type: none"> • Inicializar el residuo $\mathbf{r}^{(0)} = \mathbf{y}$, el conjunto de índices $\Lambda^{(0)} = \emptyset$ y el contador de iteración $t = 1$.
PROCEDIMIENTO	<ul style="list-style-type: none"> • Encuentra el índice $\lambda^{(t)}$ que resuelve el problema de optimización: $\lambda^{(t)} = \arg \max_{j=1, \dots, K} \langle \mathbf{r}^{(t-1)}, \mathbf{a}_j \rangle.$ Si el máximo ocurre para varios índices rompa el empate determinísticamente. • Incrementar el conjunto de índices y la matriz de átomos escogidos: $\Lambda^{(t)} = \Lambda^{(t-1)} \cup \{\lambda^{(t)}\}$ y $\mathbf{A}^{(t)} = [\mathbf{A}^{(t-1)} \mathbf{a}_{\lambda^{(t)}}]$. Se usa la convención de que $\mathbf{A}^{(0)}$ es una matriz vacía. • Resolver el problema de mínimos cuadrados para obtener un nuevo estimado de la señal: $\mathbf{x}^{(t)} = \arg \min_{\mathbf{x}} \ \mathbf{y} - \mathbf{A}^{(t)} \mathbf{x}\ _{\ell_2} = \left(\mathbf{A}^{(t)\dagger} \mathbf{A}^{(t)} \right)^{-1} \mathbf{A}^{(t)\dagger} \mathbf{y}$ • Calcular el nuevo residuo: $\mathbf{r}^{(t)} = \mathbf{y} - \mathbf{A}^{(t)} \mathbf{x}^{(t)}$ • Incrementar $t = t + 1$ y repetir todo si $t < T_0$. • La estimación $\hat{\mathbf{x}}$ para la señal ideal tiene índices distintos de cero en las componentes en Λ_{T_0}. El valor de la señal estimado $\hat{\mathbf{x}}$ en la componente λ_j es igual a la j-ésima componente de $\mathbf{x}^{(t)}$.
SALIDA	<ul style="list-style-type: none"> • La representación estimada $\hat{\mathbf{x}} \in \mathbb{R}^K$. • El conjunto Λ_{T_0} con T_0 elementos para $\{1, \dots, K\}$.

Adaptado de *OMP for Signal Recovery*, Algoritmo 3 en la referencia [5].

El algoritmo OMP es tal vez el algoritmo mayormente utilizado para representación poco densa del cual se derivan una gran variedad de algoritmos de reconstrucción, ya que proporciona errores pequeños en la reconstrucción de señales poco densas manteniendo tiempos de ejecución bajos.

1.3.2 Algoritmo *Weighted Median Regression* (WMR)

Recientemente Paredes y col. han propuesto en [6] el algoritmo *Weighted Median Regression* como un algoritmo de reconstrucción de señales poco densas para la teoría de Sensado Comprimido. Éste es un algoritmo iterativo de búsqueda voraz que emplea un esquema basado en *descenso de coordenadas* y el operador de *mediana ponderada* para abordar el problema de minimización descrito por

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_{\ell_1} + \tau \|\mathbf{x}\|_{\ell_0} \quad (1.25)$$

A este problema se le conoce como ℓ_0 -LAD o ℓ_0 -regularized LAD *regression problem* (problema de regresión de desviación absoluta mínima regularizado por la seudonorma ℓ_0), en el que se introduce como función objetivo la norma ℓ_1 , denotada como $\|\cdot\|_{\ell_1}$, para introducir robustez frente a ruido de naturaleza impulsiva o valores atípicos (de corta duración y alta energía, del inglés *outliers*) en el proceso de estimación de los coeficientes en \mathbf{x} .

Siguiendo el esquema de *descenso de coordenadas*, se divide un problema multidimensional en N problemas escalares, que se resuelven uno a la vez para estimar el valor de cada componente de \mathbf{x} , haciendo uso del operador *mediana ponderada* y haciendo una verificación de la relevancia de la contribución de los átomos a través de un proceso de umbralización fuerte, que se encarga de minimizar el término $\|\mathbf{x}\|_{\ell_0}$. Dichos problemas de optimización escalar asociados a cada componente del vector \mathbf{x} se resuelven iterativamente, manteniendo los valores de los coeficientes obtenidos en iteraciones anteriores, hasta alcanzar un criterio de convergencia. En (1.25) τ es un parámetro de regularización que controla la inducción de la poca densidad en \mathbf{x} y controla la minimización de la norma ℓ_1 , éste se convierte en una variable de umbralización que decrece conforme aumenta el número de iteraciones. Para mayores detalles revisar la referencia [6]. La Tabla 1.3 resume este método.

El valor de \mathbf{r}_n en la Tabla 1.3 está dado por $\mathbf{r}_n = \mathbf{y} - \sum_{j=1, j \neq n}^N \mathbf{A}_j x_j$. Este algoritmo presenta ventajas considerables frente a la presencia de ruido impulsivo en la señal de entrada con respecto a OMP. Sin embargo, su costo computacional es elevado debido a la aplicación del operador de *mediana ponderada* lo cual lo convierte en un algoritmo lento

Tabla 1.3. Algoritmo *Weighted Median Regression*.

ENTRADA	<ul style="list-style-type: none"> • Señal $\mathbf{y} \in \mathbb{R}^M$. • Diccionario $\mathbf{A} \in \mathbb{R}^{M \times N}$. • Número de iteraciones K_0 y energía residual objetivo ϵ.
INICIALIZACIÓN	<ul style="list-style-type: none"> • Parámetro de umbralización dura inicial: $T_h = T_{h_i} = \ \mathbf{A}^T \mathbf{y}\ _{\ell_\infty}$. • Contador de iteración $k = 1$. • Estimación en $k = 1$: $\hat{\mathbf{x}}^{(1)} = \mathbf{0}_N$.
ITERACIÓN	<p>Para la n-ésima componente de $\hat{\mathbf{x}}$, $n = 1, 2, \dots, N$, calcular</p>
ETAPA A	$\tilde{x} = \text{MEDIANA} \left(\left A_{in} \diamond \frac{y_i - \sum_{j=1, j \neq n}^N A_{ij} \hat{x}_j^{(k)}}{A_{in}} \right _{i=1}^M \right)$ $\hat{x}_n^{(k)} = \begin{cases} \tilde{x} & \text{si } \ \mathbf{r}_n\ _{\ell_1} - \ \mathbf{r}_n - \tilde{x} \mathbf{A}_n\ _{\ell_1} > T_h \\ 0 & \text{Otros} \end{cases}$
ETAPA B	<p>Actualizar el parámetro de umbralización dura y la estimación de \mathbf{x}</p> $T_h = T_{h_i} \beta^k$ $\hat{\mathbf{x}}^{(k+1)} = \hat{\mathbf{x}}^{(k)}$
ETAPA C	<p>Comprobar el criterio de parada</p> <p>Si $\left(\frac{\ \mathbf{y} - \mathbf{A} \hat{\mathbf{x}}^{(k)}\ _{\ell_2}^2}{\ \mathbf{y}\ _{\ell_2}^2} > \epsilon \right)$ y $k \leq K_0$ entonces aumenta $k = k + 1$ y repite otra iteración; de lo contrario, finaliza.</p>
SALIDA	Señal poco densa recuperada $\hat{\mathbf{x}}$.

Adaptado de *Iterative WM Regression Algorithm*, Tabla 1 en la referencia [6].

para el procesamiento de imágenes, por esto, en este proyecto se ha utilizado una versión acelerada de este algoritmo propuesta por A. Becerra y col. en la referencia [22], que en un principio estima el soporte de \mathbf{x} , es decir, el conjunto de índices de los coeficientes que son distintos de cero para evitar el cálculo de medianas sobre los coeficientes nulos en cada iteración del algoritmo.

CAPÍTULO 2

RESTAURACIÓN DE IMÁGENES

La restauración de imágenes es el conjunto de técnicas a través de las cuales se busca reconstruir una imagen que ha sido degradada a través de cierto proceso del cual se tiene algún conocimiento previo. Dentro del conjunto de técnicas de restauración de imágenes se encuentra la tarea de *inpainting* o *retoque de imagen* (término de uso no habitual), que aborda específicamente la restauración de información perdida o enmascarada sobre secciones previamente identificadas en la imagen. Para esta tarea, existen distintos métodos diseñados que abordan la restauración, métodos variacionales [23], métodos basados en problemas de ecuaciones de derivadas parciales (PDE) [24], otros basados en filtrado convencional [25] y recientemente algunos métodos basados en algoritmos de reconstrucción de señales poco densas [26, 27]. Este trabajo se ha enfocado principalmente en abordar la estimación de píxeles perdidos (*inpainting*) en imágenes usando técnicas de representación poco densa de señales.

2.1 FUNDAMENTOS DE LA RESTAURACIÓN DE IMÁGENES

En general, la restauración digital de imágenes intenta reconstruir o recuperar una imagen que ha sido degradada en alguna forma, a través de técnicas de procesamiento que dependen del conocimiento previo que se tenga del fenómeno de degradación al que la imagen fue sometida. Dichas técnicas están construidas para modelar la aplicación del procedimiento inverso a la correspondiente degradación, con el fin de recuperar la imagen original (idealmente), es decir, la imagen sin degradación. La figura 2.1 ilustra el proceso de degradación y restauración de una imagen de entrada \mathbf{Y} .

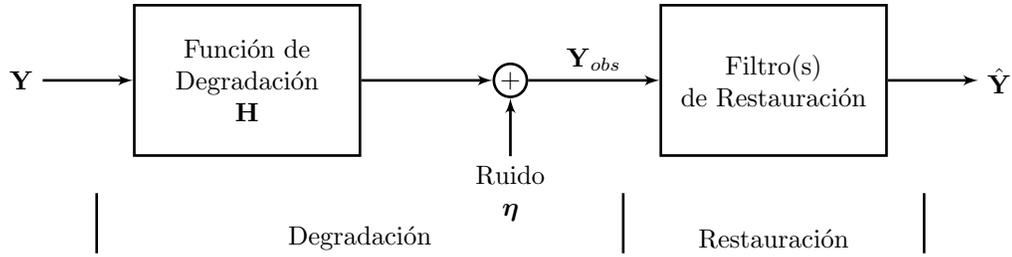


Figura 2.1. Proceso de degradación y restauración de imágenes, adaptado de la figura 5.1 de la referencia [28].

En la figura 2.1 se puede apreciar que el proceso de degradación es modelado por una función \mathbf{H} que actúa sobre la imagen de entrada \mathbf{Y} y el término de ruido aditivo dado por la función $\boldsymbol{\eta}$, así la imagen degradada observada \mathbf{Y}_{obs} está dada por:

$$\mathbf{Y}_{obs} = \mathbf{H}[\mathbf{Y}] + \boldsymbol{\eta}. \quad (2.1)$$

Si \mathbf{H} es un proceso lineal e invariante en el espacio, la función de degradación se puede escribir, en el dominio espacial, como: $\mathbf{H}[\mathbf{Y}] = \mathbf{H} \otimes \mathbf{Y}$, donde \otimes denota la operación “convolución”. Además, la función de ruido $\boldsymbol{\eta}$ (habitualmente aplicada, como se muestra, de forma aditiva) puede ser de alguna distribución como: Gaussiana, uniforme, binomial, log-normal, Laplaciana, entre otras.

El objetivo de la restauración es encontrar una estimación $\hat{\mathbf{Y}}$ lo más cercana posible a la imagen original, teniendo en cuenta que mientras mayor información se tenga del proceso de degradación mejor será la estimación. En la figura 2.1 el proceso de restauración está descrito por filtros. Sin embargo, la restauración puede depender de metodologías más complejas que algún método de filtrado convencional.

Por otra parte, y dentro de esta área de estudio, la tarea de *inpainting* digital de imágenes está íntimamente relacionada con la restauración manual de cuadros artísticos, donde se conoce *a priori* la posición de las secciones deterioradas de la obra y luego, el artista, considerando la información disponible del cuadro, puede reconstruir dichas secciones con gran precisión.

Tradicionalmente, en la interpretación digital de una restauración manual se asume Ω como una sección de la imagen que va a ser restaurada y $\partial\Omega$ como su frontera. Luego, los bordes circundantes a $\partial\Omega$ son continuados hacia dentro de la región manteniendo su

ángulo de incidencia, creando así sub-regiones dentro Ω que luego serán rellenadas con el color correspondiente a las regiones de la frontera, posteriormente se acentúan detalles para garantizar la uniformidad de la imagen.

Inicialmente, trabajos importantes fueron los de S. Masnou en [23], quien propuso un método variacional para realizar restauración de partes ocluidas de una imagen, y M. Bertalmio y col. en [24] que propusieron un enfoque matemático para abordar el problema de restauración basado en la resolución de un problema de ecuaciones de derivadas parciales (PDE), ambos problemas intentan emular un proceso de restauración manual siguiendo los principios presentados anteriormente. Sin embargo, procesos como estos no consideran la restauración de características, como por ejemplo texturas, de la imagen en el lugar donde existe la obstrucción. Otros enfoques de mayor interés para este proyecto son aquellos relacionados con la detección y remoción de ruido de naturaleza impulsiva como el propuesto por R. Chan y col. en [25], que hace uso del *filtro adaptativo de mediana ponderada central* y el *método variacional para preservación de detalles*. Más recientemente se ha propuesto métodos iterativos que explotan el concepto de representación poco densa de imágenes en diccionarios redundantes sobre el problema de restauración, partiendo de la necesidad de realizar restauraciones más adecuadas a la naturaleza de las imágenes, lo cual está íntimamente relacionado con la selección de los diccionarios más adecuados para estimar la mejor representación poco densa.

2.2 RESTAURACIÓN DE IMÁGENES USANDO REPRESENTACIÓN POCO DENSA DE SEÑALES

Suponga que $\mathbf{y} \in \mathbb{R}^n$, una imagen ordenada lexicográficamente, es una señal que tiene una representación poco densa, entonces se cumple que $\mathbf{y} = \mathbf{A}\mathbf{x}$, donde $\mathbf{x} \in \mathbb{R}^K$ es el vector de coeficientes de representación con solo T_0 elementos no nulos, es decir, $\|\mathbf{x}\|_{\ell_0} = T_0$ con $T_0 \ll K$. Ahora, suponga que se observa una versión deteriorada de \mathbf{y} tal que $\mathbf{y}_{obs} = \mathbf{M}\mathbf{y}$, donde $\mathbf{M} \in \mathbb{R}^{n \times n}$ es una matriz diagonal binaria cuyos elementos en la diagonal están

definidos por

$$m_{i,i} = \begin{cases} 0 & \text{si } y_i \text{ es un píxel perdido} \\ 1 & \text{otros.} \end{cases} \quad (2.2)$$

Así, los píxeles perdidos toman valor cero dentro de la imagen \mathbf{y}_{obs} , la figura 2.2 ilustra la pérdida de píxeles, donde los píxeles con valor cero (representados por el nivel de gris negro) corresponden a los píxeles perdidos.

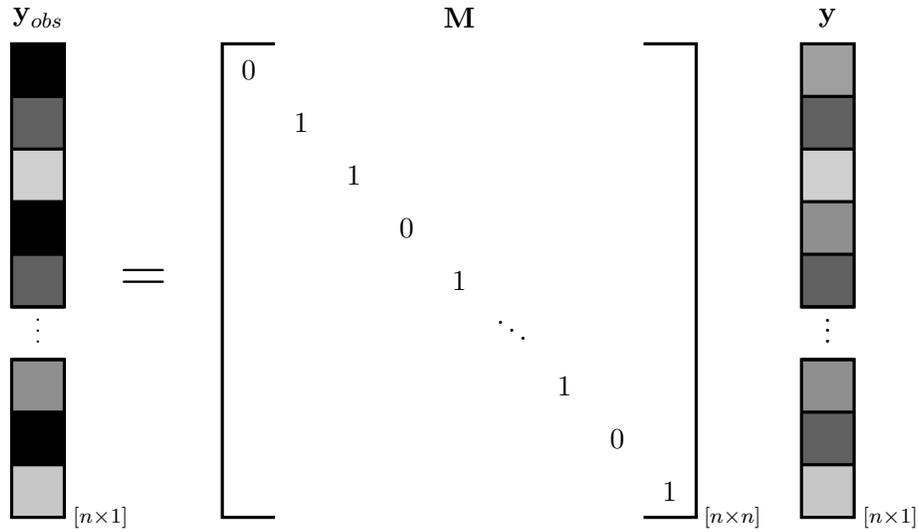


Figura 2.2. Pérdida de píxeles a través de la matriz \mathbf{M} .

La tarea de restauración intenta recuperar el vector de coeficientes $\hat{\mathbf{x}}$ a partir de las muestras observadas \mathbf{y}_{obs} , para en función de éste estimar la imagen $\mathbf{y} = \mathbf{A}\hat{\mathbf{x}}$. Considerando esto como un problema inverso similar al planteado en la ecuación (1.3), el problema de restauración se puede formular como la solución al problema de optimización

$$\arg \min_{\mathbf{x}} \|\mathbf{x}\|_{\ell_0} \quad \text{sujeto a} \quad \mathbf{y}_{obs} = \mathbf{M}\mathbf{A}\mathbf{x}. \quad (2.3)$$

Al encontrar la solución $\hat{\mathbf{x}}$ a este problema de optimización, la imagen restaurada viene dada como $\mathbf{A}\hat{\mathbf{x}}$, debido a que se conoce que \mathbf{x} es la solución con menor densidad al sistema de ecuaciones $\mathbf{y}_{obs} = \mathbf{M}\mathbf{A}\mathbf{x}$, siendo así $\mathbf{x} = \hat{\mathbf{x}}$.

El uso de la matriz \mathbf{M} en (2.3) implica el uso de un diccionario holográfico $\mathbf{V} = \mathbf{M}\mathbf{A}$, que simplemente es el diccionario \mathbf{A} con elementos iguales a cero en las filas que tienen contribución en los elementos de \mathbf{y}_{obs} correspondientes a píxeles perdidos.

M. Elad en el Capítulo 15 de la referencia [12] y M. Aharon y col. en [13], proponen esquemas simples de restauración de imágenes deterioradas con píxeles perdidos y ruido gaussiano aditivo, que resuelven el problema de optimización anterior utilizando OMP, realizando un procesamiento *bloque por bloque* de la imagen en cuestión a fin de minimizar el costo computacional y usar diccionarios de tamaño relativamente pequeños que puedan ser almacenados en memoria de acceso rápido. Este procesamiento puede realizarse tomándose bloques solapados o no solapados como se describe a continuación.

2.2.1 Procesamiento de bloques solapados contra no solapados

El procesamiento *bloque por bloque* usualmente se puede realizar de dos maneras, una utilizando los bloques de dimensión $\sqrt{n} \times \sqrt{n}$ que conforman a la imagen, sin solaparse unos con otros. Mientras un segundo esquema utiliza los bloques completamente solapados.

Suponga una imagen \mathbf{Y} de $N \times M$ píxeles, formada por bloques de $\sqrt{n} \times \sqrt{n}$ píxeles, siendo $N = \alpha\sqrt{n}$ y $M = \beta\sqrt{n}$ con $\alpha, \beta \in \mathbb{N}$. Un bloque de la imagen se denota como $\mathcal{P}_{i,j}\{\mathbf{Y}\}$, donde i, j indica la posición del primer píxel (píxel superior izquierdo) del bloque, los índices de los píxeles que forman el bloque $\mathcal{P}_{i,j}\{\mathbf{Y}\}$ varían desde i, j hasta $i + \sqrt{n} - 1, j + \sqrt{n} - 1$. La imagen \mathbf{Y} está formada por $N M/n$ bloques contiguos no solapados, $\mathcal{P}_{i,j}\{\mathbf{Y}\}$ para los índices $i = \{1, 1 + \sqrt{n}, 1 + 2\sqrt{n}, \dots, N - \sqrt{n} + 1\}$ y $j = \{1, 1 + \sqrt{n}, 1 + 2\sqrt{n}, \dots, M - \sqrt{n} + 1\}$. Sin embargo, si se permite que los bloques se solapen la imagen se puede procesar como $(N - \sqrt{n} + 1)(M - \sqrt{n} + 1)$ bloques solapados completamente, producto de variar los índices del operador \mathcal{P} en una unidad, tal que $i = \{1, 2, \dots, N - \sqrt{n} + 1\}$ y $j = \{1, 2, \dots, M - \sqrt{n} + 1\}$. Obsérvese que también pudiera darse el caso de bloques parcialmente solapados donde la variación de los índices de \mathcal{P} es mayor a la unidad.

Para una imagen contaminada, cada bloque (ordenado lexicográficamente) procesado con OMP devuelve su correspondiente vector de representación $\hat{\mathbf{x}}$, y éste devuelve un bloque reconstruido de la imagen a través de $\mathbf{A}\hat{\mathbf{x}}$. Cuando se hace el procesamiento sin solapamiento la totalidad de los bloques procesados conforman la imagen reconstruida, por otra parte, cuando se utilizan bloques completamente solapados se tiene que recurrir al promedio (media aritmética) entre los píxeles que se solapan para luego formar la imagen reconstruida. La figura 2.3 ilustra el proceso de restauración usando OMP y los

dos enfoques; con y sin solapamiento.

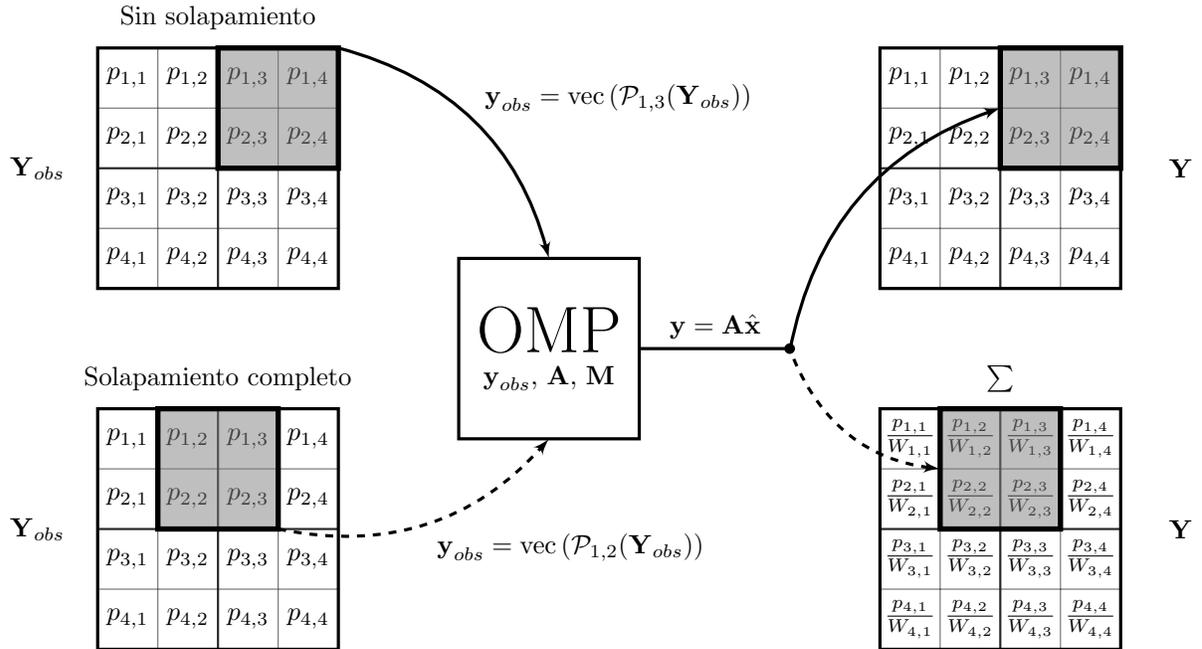


Figura 2.3. Esquema de reconstrucción para una imagen \mathbf{Y}_{obs} de 4×4 píxeles, con bloques de 2×2 píxeles. Línea continua muestra proceso sin solapamiento de bloques. Línea punteada muestra proceso con solapamiento completo de bloques. El factor $w_{i,j}$ es el número total de veces que el píxel $p_{i,j}$ será estimado.

Ambos enfoques tiene sus ventajas y desventajas, usar solapamiento completo de bloques lleva a una mejor reconstrucción de la imagen, pero requiere de mayor tiempo computacional, pues deben procesarse $(N - \sqrt{n} + 1)(M - \sqrt{n} + 1)$ bloques en lugar de $N M/n$ bloques para el caso de bloques no solapados.

A fin de ilustrar los efectos de una restauración de este tipo se ha preparado una prueba de la imagen estándar de procesamiento **Peppers**, cuyas dimensiones son: 256×256 píxeles, con el 30% de los píxeles perdidos localizados aleatoriamente. Con el uso de un diccionario DCT bidimensional sobrecompleto de dimensión 64×256 , para atender un procesamiento de bloques de 8×8 píxeles. Se ha utilizado OMP resolviendo el problema de aproximación restringido por poca densidad de la ecuación (1.23) en el Capítulo 1, para $T_0 = 4$, es decir, logrando una poca densidad de a lo sumo 4 elementos distintos de cero entre los coeficientes de representación correspondientes a cada bloque de la imagen. Para este problema se debe resolver el problema inverso dado por (1.23) para $256 \cdot 256/64 = 1024$ bloques para el caso de no solapamiento y $(256 - 8 + 1)(256 - 8 + 1) = 62001$ bloques

para el caso con solapamiento. En la figura 2.4 se muestran los resultados de esta prueba, incluyendo valores de medidas de desempeño de la reconstrucción: *relación pico de la señal a ruido* (PSNR), *error medio absoluto* (MAE) y la *raíz del error medio cuadrático* (RMSE), definidas en (A.2), (A.3) y (A.4) respectivamente en el Apéndice A.

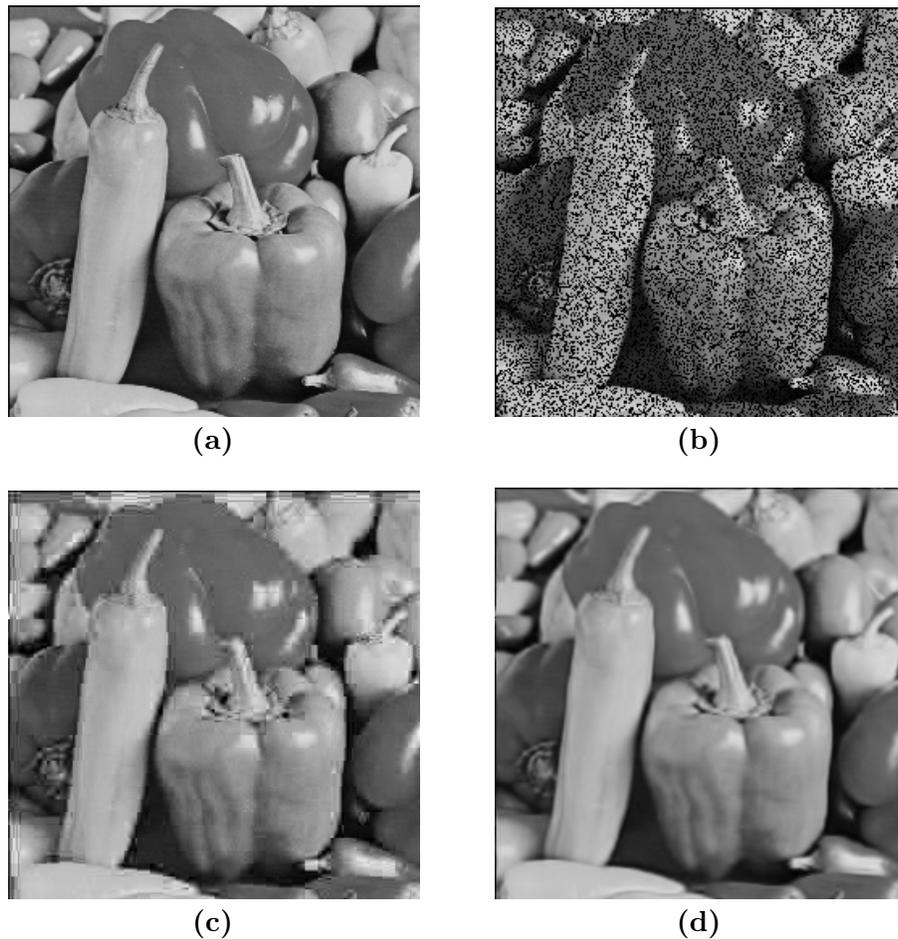


Figura 2.4. Prueba de restauración usando OMP. (a) Imagen original. (b) Imagen con 30% píxeles perdidos. (c) Imagen reconstruida sin solapamiento de bloques, $t_r = 4,51$ s. $PSNR = 26,21$ dB, $MAE = 6,91$, $RMSE = 12,42$. (d) Imagen reconstruida solapando completamente los bloques, $t_r = 260,35$ s. $PSNR = 29,82$ dB, $MAE = 4,42$, $RMSE = 8,20$.

Al procesar sin solapamiento de bloques se puede observar un *efecto de bloque* en la imagen reconstruida, en caso contrario, al solapar completamente se reduce el *efecto de bloque* dando una reconstrucción más suave producto de la media aritmética entre los píxeles que se solapan. Claro está que el costo computacional que se paga por este

desempeño es alto¹, pues se requiere 260,35 s para reconstruir la imagen usando el enfoque de bloques solapados, con apenas 4,51 s con bloques no solapados.

2.2.2 Detección y remoción de ruido impulsivo

B. Deka y col. en la referencia [27] propone un completo esquema de detección y remoción de ruido impulsivo de valores aleatorios, que involucra una etapa de restauración como se ha descrito previamente, este esquema se muestra en la figura 2.5. El ruido impulsivo considerado en el esquema de la figura 2.5 puede tomar valores aleatorios dentro del rango dinámico (valor mínimo y máximo) de la imagen \mathbf{y} , y está contenido en la imagen observada \mathbf{y}_{obs} . El esquema en cuestión consta de tres etapas, una de detección y filtrado, una etapa de diseño (aprendizaje) de diccionario, y una tercera etapa de restauración de imagen a través de procesamiento con OMP.

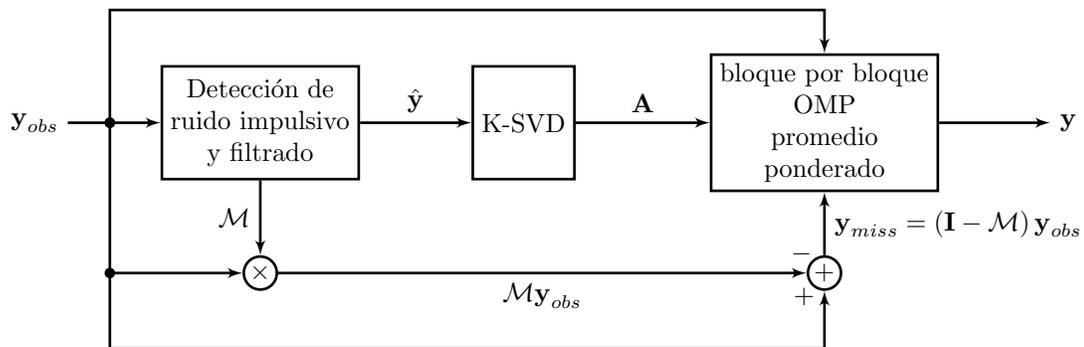


Figura 2.5. Esquema de restauración de imágenes contaminadas con ruido impulsivo.

En la primera etapa del esquema se procesa \mathbf{y}_{obs} con el método de filtrado basado en el *filtro adaptativo de mediana ponderada central* y el *método variacional para preservación de detalles* [25]. De acá se obtiene \mathcal{M} , el operador de contaminación sobre \mathbf{y} , y además una estimación preliminar de la imagen denominada $\hat{\mathbf{y}}$, es decir, se determina qué píxel ha sido contaminado y se sustituye este valor contaminado por un valor aproximado definido por los píxeles vecinos al contaminado. La imagen estimada luego del filtrado, $\hat{\mathbf{y}}$, es utilizada para entrenar un diccionario \mathbf{A} a través del algoritmo K-SVD; etapa que no es imprescindible debido a que se puede sustituir con el uso de un diccionario preespecificado en la etapa siguiente. El operador \mathcal{M} es la matriz de pérdida \mathbf{M} , utilizada para

¹Esta prueba se realizó en una PC Windows XP x86 AMD Phenom II X3 720 2,8 GHz

remover los valores de los píxeles presuntamente contaminados con ruido impulsivo tal como se muestra en la figura 2.2. La tercera etapa, denominada *Sparse Reconstruction stage* (etapa de reconstrucción poco densa), resuelve el problema en (2.3) *bloque por bloque* (considerando solapamiento completo entre bloques) usando OMP como técnica de resolución del problema de optimización (2.3), teniendo como parámetros de entrada la imagen \mathbf{y}_{obs} , el diccionario \mathbf{A} y el vector \mathbf{y}_{miss} cuyas componentes son cero excepto en las posiciones donde se han perdido píxeles; estos últimos parámetros se pueden resumir con la entrada del diccionario holográfico $\mathbf{V} = \mathbf{AM}$. Obsérvese la gran similitud entre este esquema de limpieza de señal ante contaminaciones impulsivas y el esquema de restauración de imágenes descrito anteriormente.

El uso de un diccionario entrenado por K-SVD infiere en una restauración más precisa, debido a que los coeficientes de representación estimados tendrán más capacidad para restaurar la imagen con una mayor tendencia hacia su naturaleza original, esto es, mayor preservación de bordes, texturas, curvas, etc. Para ilustrar esta afirmación se ha preparado una prueba con las mismas características de la prueba en la figura 2.4, utilizando un diccionario entrenado por los 62001 bloques solapados de una versión degradada de la imagen **Peppers**. El método K-SVD se ha inicializado con un diccionario DCT sobre-completo bidimensional de dimensión 64×256 . La figura 2.6 muestra la imagen de entrenamiento y el diccionario resultante, y la figura 2.7 muestra los resultados de la prueba.

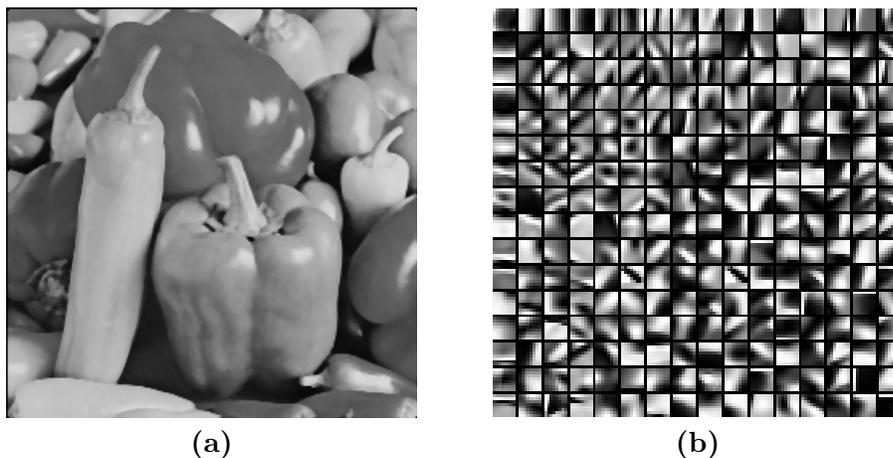


Figura 2.6. Imagen de entrenamiento y Diccionario entrenado para la imagen Peppers. (a) Imagen de entrenamiento, versión degradada de la imagen original por un filtro de mediana de ventana 3×3 píxeles. $PSNR = 32, 20$ dB, $MAE = 2, 69$, $RMSE = 6, 24$. (b) Bases del diccionario entrenado por K-SVD de dimensión 64×256 .

La degradación de la imagen original, para construir la imagen de entrenamiento de la figura 2.6(a), se efectuó a través de un filtro convencional de mediana de ventana 3×3 píxeles, esto para ser más realistas y emular rápidamente un proceso de detección en el que la imagen se ha degradado.

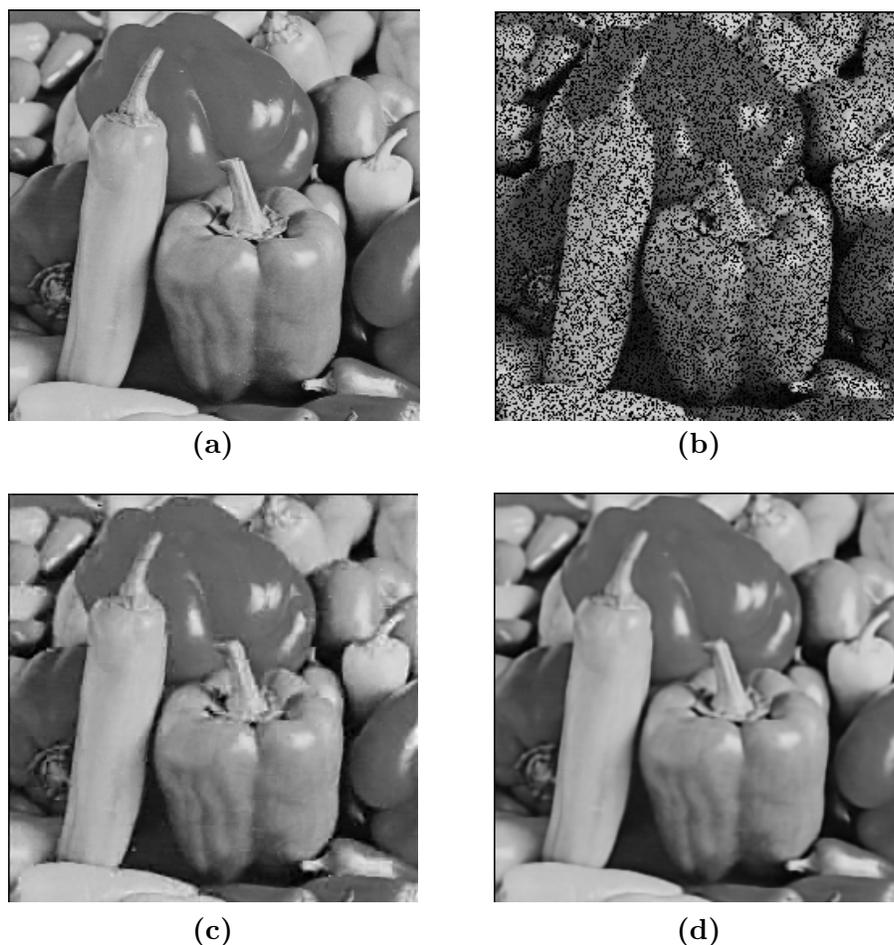


Figura 2.7. Prueba de restauración usando OMP y diccionario entrenado por K-SVD. (a) Imagen original. (b) Imagen con 30 % píxeles perdidos. (c) Imagen reconstruida sin solapamiento de bloques. $PSNR = 29,68$ dB, $MAE = 4,79$, $RMSE = 8,33$. (d) Imagen reconstruida solapando completamente los bloques. $PSNR = 32,46$ dB, $MAE = 3,45$, $RMSE = 6,05$.

Se obtiene una mejora de 3,47 dB en la medida de PSNR respecto a la prueba en la figura 2.4 no solapando los bloques durante el procesamiento, se obtuvo además una mejora de 2,64 dB usando bloques completamente solapados, ambas significativas ya que un incremento de 0,5 dB puede ser visualmente apreciado en una imagen restaurada.

CAPÍTULO 3

RESTAURACIÓN DE IMÁGENES BASADA EN WMR

En este capítulo se propone el uso del algoritmo de reconstrucción *Weighted Median Regression* como una alternativa robusta a la resolución del problema de restauración o interpolación de píxeles perdidos y remoción de ruido impulsivo utilizando representación poco densa de señales, abordando el problema presentado en (2.3).

3.1 WMR COMO MÉTODO ROBUSTO DE RESTAURACIÓN

El algoritmo *Weighted Median Regression* es presentado en [6] como un método robusto de reconstrucción en problemas de sensado comprimido. Su naturaleza matemática permite afirmar desde un principio que este método está adecuado al tratamiento de contaminaciones impulsivas sobre señales, y además intenta promover la estimación exhaustiva de señales efectivamente poco densas.

3.1.1 Aproximación robusta por la norma ℓ_1

El problema de minimización abordado por el algoritmo WMR en (1.25) es un problema de aproximación regularizada, en el que se presenta como función objetivo al término de ajuste de datos $\|\mathbf{y} - \mathbf{Ax}\|_{\ell_1}$ con una regularización dada por la seudonorma ℓ_0 , $(\tau\|\mathbf{x}\|_{\ell_0})$ que intenta introducir pocos elementos no nulos en la aproximación. Dicha función objetivo *en principio* plantea un problema de aproximación por norma ℓ_1 de la forma:

$$\arg \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{Ax}\|_{\ell_1}. \quad (3.1)$$

Que es un problema de aproximación robusto debido a la naturaleza de la norma en cuestión. S. Boyd y col. en el Capítulo 7 de [9] explican que partiendo del hecho de que cualquier problema de aproximación por norma ℓ_p puede ser interpretado como un problema de *estimación por máxima verosimilitud* para cierta densidad de ruido donde la señal observada es de la forma $\mathbf{y} = \mathbf{Ax} + \boldsymbol{\eta}$ y se intenta minimizar la norma ℓ_p del ruido, i.e. $\|\boldsymbol{\eta}\| = \|\mathbf{y} - \mathbf{Ax}\|_{\ell_p}$. Entonces, la robustez de la norma ℓ_1 puede ser entendida al considerar el problema de estimación por máxima verosimilitud con una densidad de ruido Laplaciana, en cuyo caso el método en cuestión está a la espera de valores residuales más grandes para controlar su efecto sobre el proceso de aproximación. Para mayor detalle se recomienda al lector revisar el Capítulo 6 y 7 de la referencia [9].

Por otra parte, el problema de minimización que resuelve el algoritmo OMP en (1.23) es *en principio* un problema de aproximación centrado en la minimización del término $\|\mathbf{y} - \mathbf{Ax}\|_{\ell_2}^2$, que es esencialmente un problema de aproximación por norma ℓ_2 definido por

$$\arg \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{Ax}\|_{\ell_2}. \quad (3.2)$$

Que también puede ser interpretado como un problema de estimación por máxima verosimilitud pero considerando una densidad de ruido de tipo Gaussiana.

Estas afirmaciones básicas hacen surgir la idea de que el algoritmo WMR es más adecuado para la restauración de imágenes naturales con píxeles perdidos, debido a que la pérdida de píxeles puede ser considerada como valores atípicos en una imagen de cierta naturaleza, o ruido que puede ser estadísticamente modelado por una distribución de colas pesadas como la Laplaciana. En contraposición al algoritmo OMP que se asume adecuado para la restauración de imágenes contaminadas con ruido Gaussiano.

Adicionalmente es bien conocido en la literatura que los detalles y bordes en una imagen son mejor preservados cuando la estrategia de procesamiento se basa en la norma ℓ_1 en lugar de la norma ℓ_2 . Se esperaría, que así como el filtrado lineal es óptimo bajo el principio de máxima verosimilitud cuando la contaminación de fondo sigue una distribución Gaussiana, y los filtros basados en mediana ponderada son apropiados para distribuciones de ruido Laplacianas, el proceso de restauración basado en minimización de la norma ℓ_1 ofrezca las mismas ventajas que se consiguen en procesos de filtrado donde la mediana

ponderada (la cual tiene sus raíces en el uso de la norma ℓ_1) es usada como el operador fundamental en lugar del operador promedio.

La calidad de una reconstrucción depende también de la información que se tenga de la posición de los impulsos (o píxeles perdidos) que contaminan la imagen, es decir, si se conoce o no *a priori* la matriz de pérdida o contaminación \mathbf{M} , tal que $\mathbf{y}_{obs} = \mathbf{M}\mathbf{y}$, para una imagen \mathbf{Y} y su versión observada contaminada \mathbf{Y}_{obs} ordenadas lexicográficamente.

Si \mathbf{M} es conocido, para el problema de restauración descrito por:

$$\arg \min_{\mathbf{x}} \|\mathbf{x}\|_{\ell_0} \quad \text{sujeto a} \quad \mathbf{y}_{obs} = \mathbf{M}\mathbf{A}\mathbf{x}, \quad (3.3)$$

se dispondrá de la información necesaria, debido al diccionario holográfico $\mathbf{V} = \mathbf{M}\mathbf{A}$, para que a través de algoritmos como OMP o WMR se pueda estimar una solución $\hat{\mathbf{x}}$, despreciando la información que incorpore algún impulso o algún píxel perdido en la estimación de los coeficientes. De lo contrario, si no se conoce la matriz de contaminación \mathbf{M} los algoritmos tienden a ser altamente vulnerables a los efectos del ruido o la pérdida de datos, al abordar el problema en (3.3) descartando el uso del diccionario holográfico \mathbf{V} y aplicando los métodos de reconstrucción solo tomando en cuenta el diccionario \mathbf{A} .

Para ilustrar la robustez del algoritmo WMR se presenta un grupo de pruebas de reconstrucción sobre una señal sintética poco densa \mathbf{x} de 64 elementos y 4 elementos no nulos, que proyectada sobre un diccionario \mathbf{A} completo unidimensional de 64 átomos formado por las bases ortogonales DCT da origen a una señal \mathbf{y} de 64 elementos, como se muestra en la figura 3.1. La señal \mathbf{y} es sometida a una contaminación impulsiva para una primera prueba, y para una segunda prueba se ha simulado una pérdida de algunas de sus muestras, formado así una señal ruidosa observada \mathbf{y}_{obs} en ambos casos. El objetivo de ambas pruebas es recuperar el vector de coeficientes \mathbf{x} a través de los algoritmos de optimización estudiados, abordando el problema en (3.3) utilizando el diccionario holográfico cuando \mathbf{M} es conocida y descartando su uso en el caso contrario, buscando una reconstrucción con el menor error posible a partir de las muestras \mathbf{y}_{obs} . La figura 3.2 ilustra una reconstrucción de la señal \mathbf{y} contaminada con impulsos y asumiendo desconocida la posición de estos, es decir, sin uso del diccionario holográfico \mathbf{V} , mientras la figura 3.3 muestra los resultados de la reconstrucción usando \mathbf{V} , análogamente para la simulación de pérdida de datos las

figuras 3.4 y 3.5 ilustran los resultados de las pruebas, descartando y usando el diccionario holográfico \mathbf{V} respectivamente.

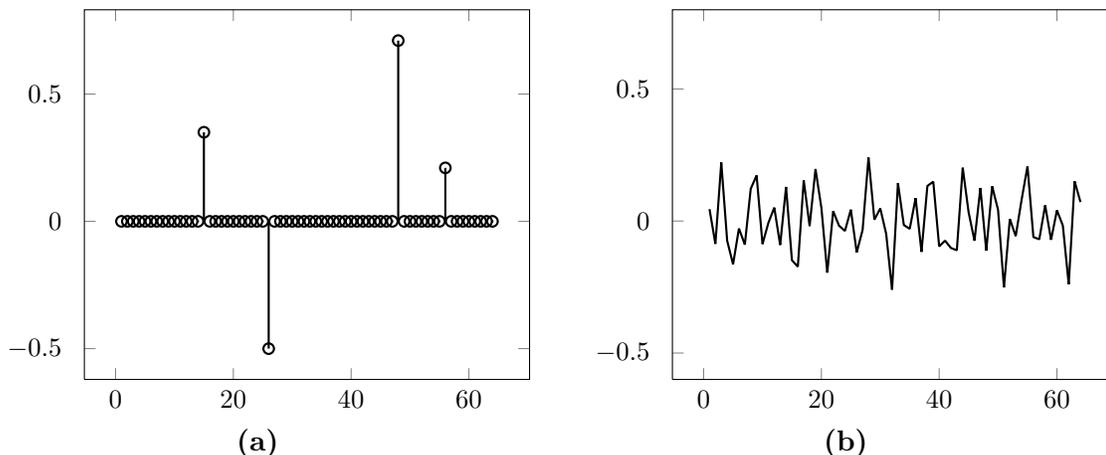


Figura 3.1. Señal sintética y proyecciones. (a) Señal poco densa original. (b) Proyecciones en el dominio DCT.

A partir de estas pruebas es evidente que ambos algoritmos tienen un rendimiento adecuado cuando se conoce la posición de los impulsos o los píxeles perdidos. Sin embargo, estas posiciones habitualmente son desconocidas porque son aleatorias, de forma que se requiere el uso de métodos que las estimen, por esto, algunos autores describen métodos completos de detección y restauración de ruido impulsivo como se ha descrito en la Sección 2.2.2. Del mismo modo, se puede apreciar que el comportamiento del algoritmo WMR en ambos casos de contaminación es el esperado cuando se descarta el uso del diccionario holográfico, es decir, el efecto de la contaminación es despreciado notablemente debido a la robustez del algoritmo como se ha descrito anteriormente.

Este trabajo se centra en presentar el algoritmo WMR como alternativa robusta a la restauración de imágenes, y específicamente se centra en abordar la interpolación de píxeles perdidos usando representación poco densa de señales, o en su defecto abordar la etapa de reconstrucción poco densa del método descrito en la figura 2.5. En la siguiente sección se describe el procedimiento efectuado para realizar restauración de imágenes basada en representación poco densa de señales usando *Weighted Median Regression*.

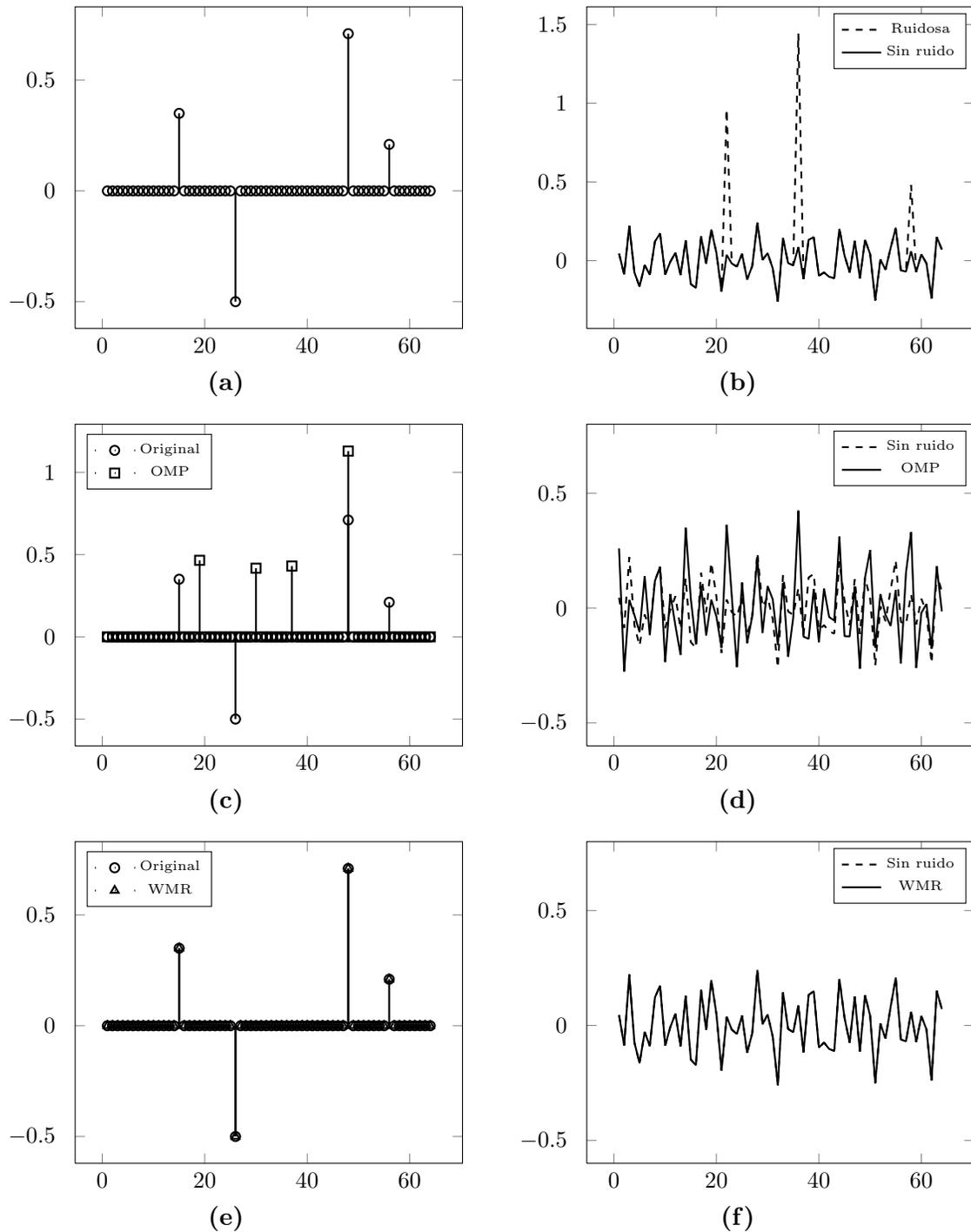


Figura 3.2. Reconstrucción de una señal sintética contaminada con impulsos y sin el uso del diccionario holográfico V . (a) Señal sintética poco densa x . (b) Proyecciones originales y junto a proyecciones ruidosas y_{obs} de la señal x . (c) Señal estimada \hat{x} usando OMP, y señal original, $MAE = 0,04$, $RMSE = 0,14$. (d) Proyecciones de la señal estimada \hat{x} con OMP, y proyecciones sin ruido originales. (e) Señal estimada \hat{x} usando WMR, y señal original, $MAE = 2,17 \cdot 10^{-18}$, $RMSE = 1,04 \cdot 10^{-17}$. (d) Proyecciones de la señal estimada \hat{x} con WMR, y proyecciones sin ruido originales.

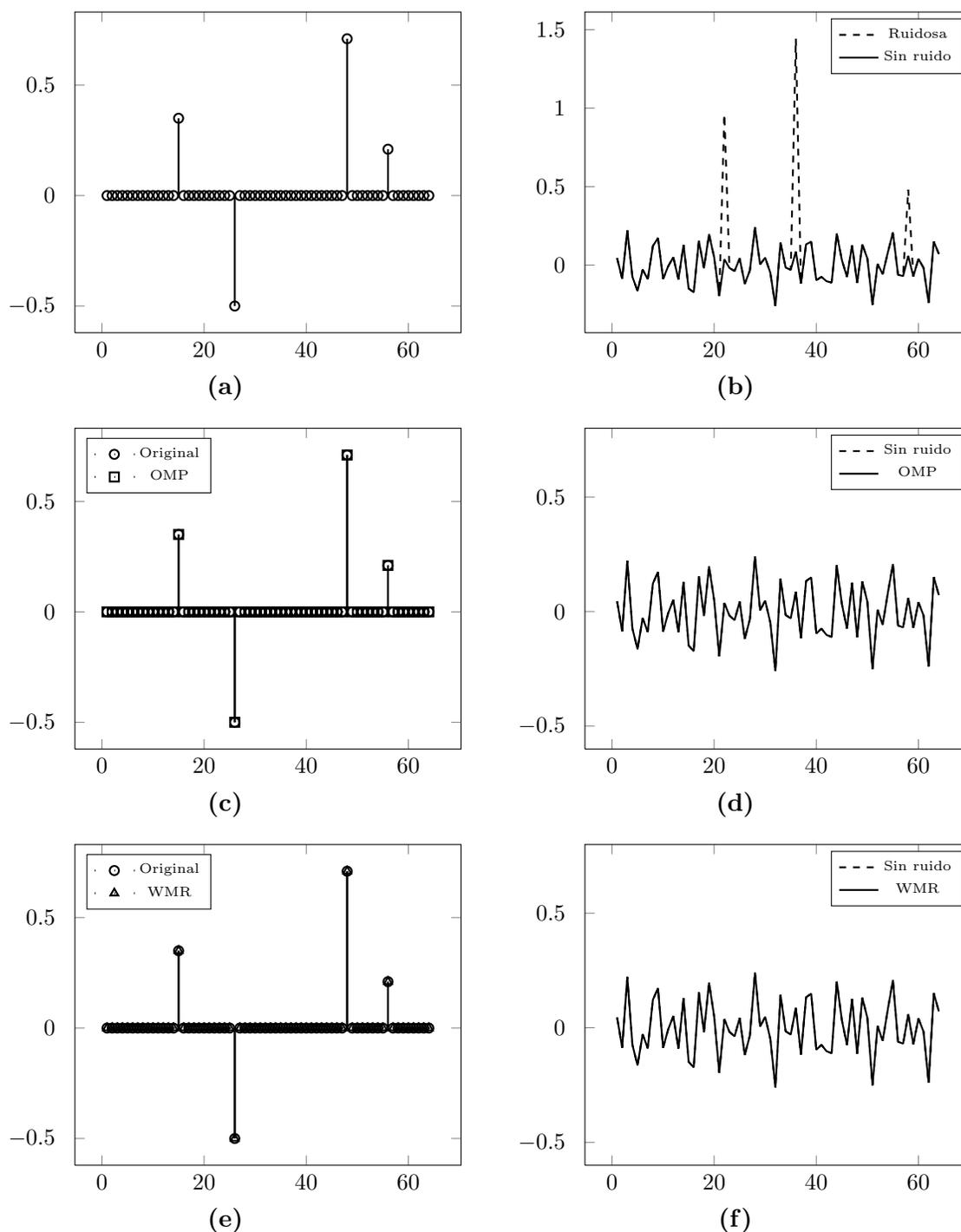


Figura 3.3. Reconstrucción de una señal sintética contaminada con impulsos usando diccionario holográfico V . (a) Señal sintética poco densa x . (b) Proyecciones originales y junto a proyecciones ruidosas y_{obs} de la señal x . (c) Señal estimada \hat{x} usando OMP, y señal original, $MAE = 1,26 \cdot 10^{-17}$, $RMSE = 6,48 \cdot 10^{-17}$. (d) Proyecciones de la señal estimada \hat{x} con OMP, y proyecciones sin ruido originales. (e) Señal estimada \hat{x} usando WMR, y señal original, $MAE = 7,15 \cdot 10^{-13}$, $RMSE = 3,29 \cdot 10^{-12}$. (f) Proyecciones de la señal estimada \hat{x} con WMR, y proyecciones sin ruido originales.

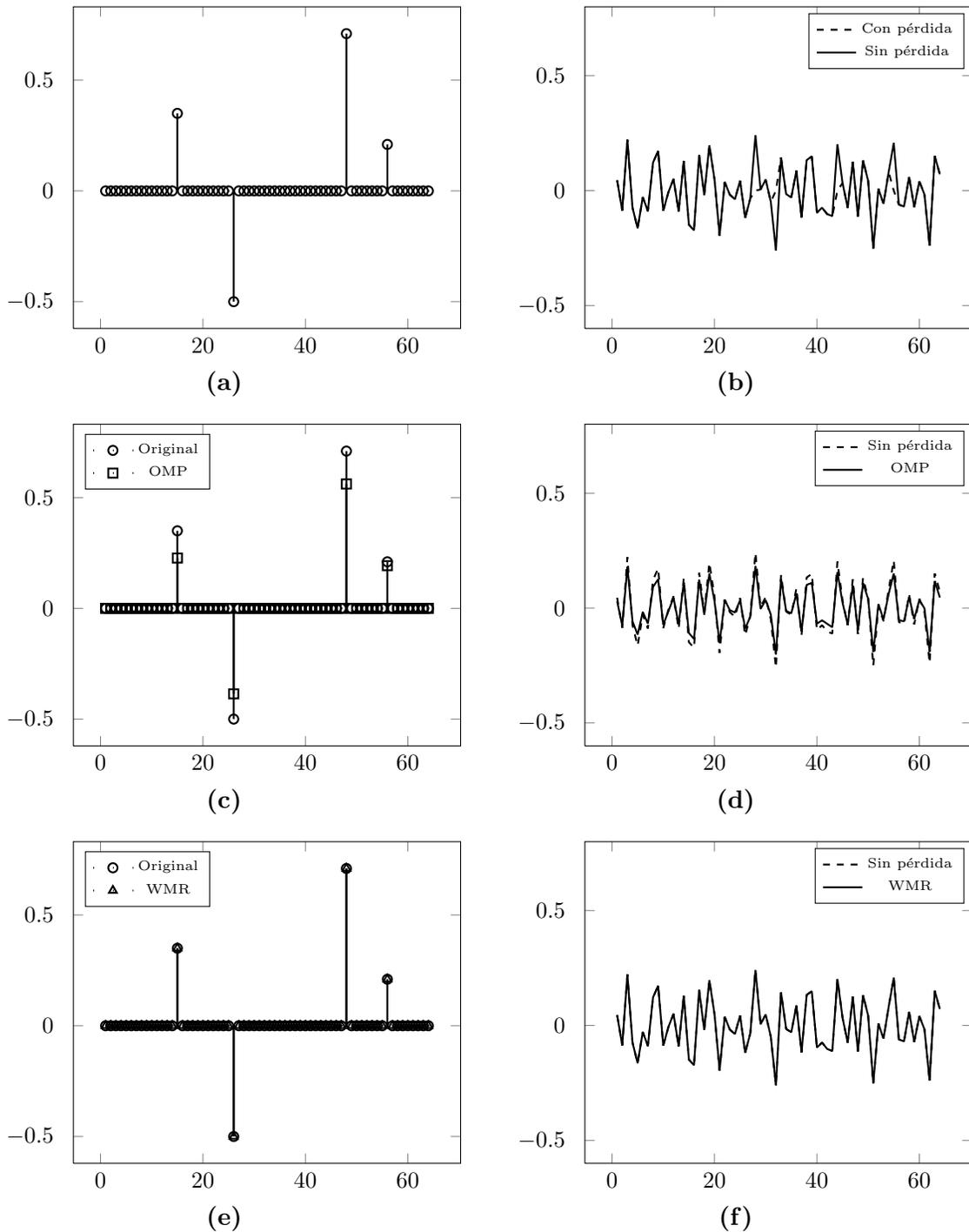


Figura 3.4. Reconstrucción de una señal sintética con pérdida de datos sin uso del diccionario holográfico V . (a) Señal sintética poco densa x . (b) Proyecciones originales y junto a proyecciones con pérdida y_{obs} de la señal x . (c) Señal estimada \hat{x} usando OMP, y señal original, $MAE = 0,01$, $RMSE = 0,03$. (d) Proyecciones de la señal estimada \hat{x} con OMP, y proyecciones sin pérdida originales. (e) Señal estimada \hat{x} usando WMR, y señal original, $MAE = 6,07 \cdot 10^{-18}$, $RMSE = 3,17 \cdot 10^{-17}$. (f) Proyecciones de la señal estimada \hat{x} con WMR, y proyecciones sin pérdida originales.

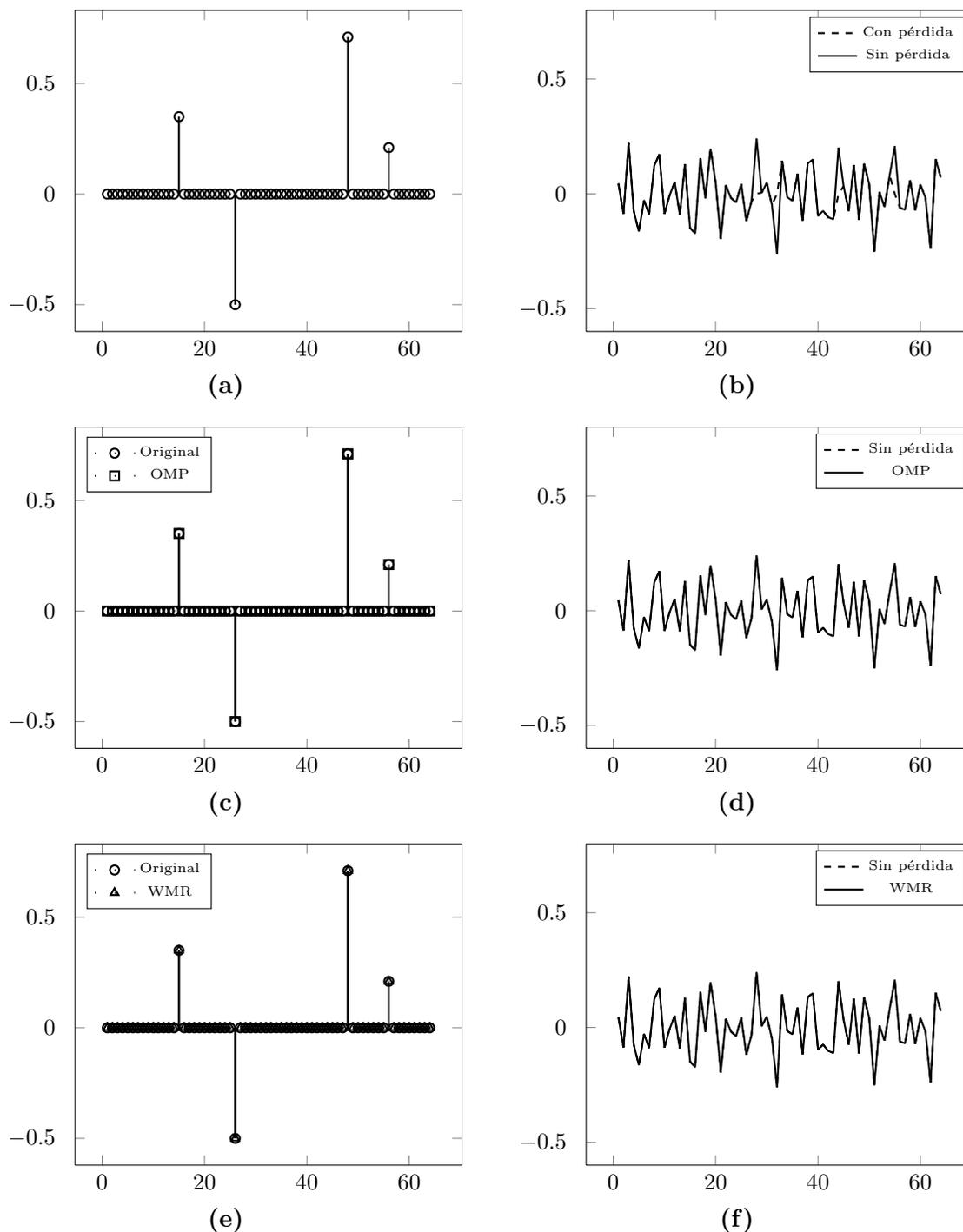


Figura 3.5. Reconstrucción de una señal sintética con pérdida de datos usando el diccionario holográfico V . (a) Señal sintética poco densa x . (b) Proyecciones originales y junto a proyecciones con pérdida y_{obs} de la señal x . (c) Señal estimada \hat{x} usando OMP, y señal original, $MAE = 8,24 \cdot 10^{-18}$, $RMSE = 4 \cdot 10^{-17}$. (d) Proyecciones de la señal estimada \hat{x} con OMP, y proyecciones sin pérdida originales. (e) Señal estimada \hat{x} usando WMR, y señal original, $MAE = 1,51 \cdot 10^{-12}$, $RMSE = 8,48 \cdot 10^{-12}$. (d) Proyecciones de la señal estimada \hat{x} con WMR, y proyecciones sin pérdida originales.

3.2 METODOLOGÍA APLICADA

En breve, el proceso de restauración de píxeles perdidos de una imagen es realizado usando procesamiento *bloque por bloque*. Partiendo inicialmente en la extracción de un bloque en particular de la imagen, seguido de la construcción del diccionario holográfico que define su contaminación y finaliza en la aplicación de algún método de reconstrucción poco densa de señales, que devuelve el bloque restaurado a través de la proyección de los coeficientes estimados sobre el diccionario “no holográfico” seleccionado.

3.2.1 Planteamiento matemático

Inicialmente se parte de una versión contaminada \mathbf{Y}_{obs} de la imagen $\mathbf{Y} \in \mathbb{R}^{N \times M}$, tal que $\mathbf{Y}_{obs} = \mathcal{M}\{\mathbf{Y}\}$, donde \mathcal{M} es el operador de pérdida de píxeles sobre todo el espacio de \mathbf{Y} . Se tiene un bloque $\mathcal{P}_{i,j}\{\mathbf{Y}_{obs}\}$ de tamaño $\sqrt{n} \times \sqrt{n}$ de la imagen contaminada como parte fundamental de procesamiento, donde los índices i, j pueden variar dependiendo del enfoque de procesamiento utilizado, ya sea sin solapamiento o con solapamiento completo (o parcial; no utilizado en este trabajo) de todos los bloques. Siendo $i = \{1, 1 + \sqrt{n}, 1 + 2\sqrt{n}, \dots, N - \sqrt{n} + 1\}$ y $j = \{1, 1 + \sqrt{n}, 1 + 2\sqrt{n}, \dots, M - \sqrt{n} + 1\}$ para la identificación de los bloques sin solapamiento, e $i = \{1, 2, \dots, N - \sqrt{n} + 1\}$ y $j = \{1, 2, \dots, M - \sqrt{n} + 1\}$ para solapamiento completo.

Cada bloque $\mathcal{P}_{i,j}\{Y\}$ de la imagen original puede ser representado como una combinación lineal de ondas elementales, así, el problema de representación objetivo en términos del bloque ordenado lexicográficamente, tal que $\mathbf{y}_{\mathcal{P}_{i,j}} = \text{vec}(\mathcal{P}_{i,j}\{\mathbf{Y}\})$, es:

$$\mathbf{y}_{\mathcal{P}_{i,j}} = \mathbf{A}\mathbf{x}_{\mathcal{P}_{i,j}}. \quad (3.4)$$

Donde $\mathbf{x}_{\mathcal{P}_{i,j}}$ es el vector de coeficientes de representación poco densa correspondiente al bloque $\mathcal{P}_{i,j}\{\mathbf{Y}\}$, y \mathbf{A} algún diccionario apropiadamente definido o diseñado de acuerdo a la estrategia descrita en la Sección 1.2. Se puede reescribir este problema de representación para cada bloque $\mathcal{P}_{i,j}\{\mathbf{Y}_{obs}\}$ de la imagen contaminada como

$$\mathbf{y}_{obs\mathcal{P}_{i,j}} = \mathbf{M}_{\mathcal{P}_{i,j}}\mathbf{A}\mathbf{x}_{\mathcal{P}_{i,j}}. \quad (3.5)$$

Donde $\mathbf{M}_{\mathcal{P}_{i,j}}$ es una matriz diagonal de tamaño $\sqrt{n} \times \sqrt{n}$ que introduce la pérdida de píxeles en el bloque $\mathcal{P}_{i,j}\{\mathbf{Y}\}$, dando origen al diccionario holográfico $\mathbf{V}_{\mathcal{P}_{i,j}} = \mathbf{M}_{\mathcal{P}_{i,j}}\mathbf{A}$, y $\mathbf{y}_{obs\mathcal{P}_{i,j}}$ es el bloque contaminado ordenado lexicográficamente, tal que $\mathbf{y}_{obs\mathcal{P}_{i,j}} = \text{vec}(\mathcal{P}_{i,j}\{\mathbf{Y}_{obs}\})$.

El problema de minimización robusto y planteado para cada bloque contaminado está definido por

$$\hat{\mathbf{x}}_{\mathcal{P}_{i,j}} = \arg \min_{\mathbf{x}_{\mathcal{P}_{i,j}}} \|\mathbf{y}_{obs\mathcal{P}_{i,j}} - \mathbf{V}_{\mathcal{P}_{i,j}}\mathbf{x}_{\mathcal{P}_{i,j}}\|_{\ell_1} + \tau \|\mathbf{x}_{\mathcal{P}_{i,j}}\|_{\ell_0}, \quad (3.6)$$

donde $\hat{\mathbf{x}}_{\mathcal{P}_{i,j}}$ es la solución óptima al problema inverso de (3.5), abordado por el algoritmo *Weighted Median Regression*.

La imagen reconstruida $\hat{\mathbf{Y}}$ está formada por bloques $\mathcal{P}_{i,j}\{\hat{\mathbf{Y}}\}$ definidos por

$$\mathcal{P}_{i,j}\{\hat{\mathbf{Y}}\} = \text{mat}(\mathbf{A}\mathbf{x}_{\mathcal{P}_{i,j}}), \quad (3.7)$$

donde $\text{mat}(\cdot)$ es el ordenamiento matricial del vector en su argumento, es decir, operación inversa al ordenamiento lexicográfico definido tal que $\mathbf{X} = \text{mat}(\text{vec}(\mathbf{X}))$. Sin embargo, hay que tomar en cuenta que en un proceso de restauración donde se permite un solapamiento entre bloques la reconstrucción de la imagen $\hat{\mathbf{Y}}$ debe considerar la media entre los píxeles solapados devueltos en cada estimación de $\mathcal{P}_{i,j}\{\hat{\mathbf{Y}}\}$. Cada píxel $\hat{\mathbf{Y}}_{i,j}$ en cada bloque será estimado $W_{i,j}$ veces, éste parámetro podría tomar algún valor entre 1 y como máximo n , es decir, $1 \leq W_{i,j} \leq n$ con $W_{i,j} \in \mathbb{N}$. El número de veces que se pueda estimar cada píxel depende del tamaño de la imagen, la posición del píxel, el tamaño del bloque y el índice de solapamiento, que en este trabajo siempre será asumido como 1 (solapamiento completo). Entonces, cada píxel de $\hat{\mathbf{Y}}$ está definido como

$$\hat{\mathbf{Y}}_{i,j} = \frac{1}{W_{i,j}} \sum_{w=1}^{W_{i,j}} \hat{\mathbf{Y}}_{i,j_w}. \quad (3.8)$$

Donde $\hat{\mathbf{Y}}_{i,j_w}$ es el valor del píxel $\hat{\mathbf{Y}}_{i,j}$ obtenido en la w -ésima estimación. La figura 3.6 indica el valor de $W_{i,j}$ para cada posición de los píxeles en una sección de la parte superior izquierda de una imagen $\mathbf{Y} \in \mathbb{R}^{N \times M}$, para un procesamiento de bloques de tamaño 8×8 completamente solapados, con $N, M \gg 8$.

Es necesario acotar que en este trabajo hasta ahora no se ha considerado que las

1	2	3	4	5	6	7	8	8	8
2	4	6	8	10	12	14	16	16	16
3	6	9	12	15	18	21	24	24	24
4	8	12	16	20	24	28	32	32	32
5	10	15	20	25	30	35	40	40	40
6	12	18	24	30	36	42	48	48	48
7	14	21	28	35	42	49	56	56	56
8	16	24	32	40	48	56	64	64	64
8	16	24	32	40	48	56	64	64	64
8	16	24	32	40	48	56	64	64	64

Figura 3.6. Distribución de $W_{i,j}$ para los píxeles en una imagen procesada con solapamiento completo de bloques de tamaño 8×8 píxeles.

imágenes tengan alguna contaminación adicional a la pérdida de píxeles, por lo tanto, la imagen restaurada final $\hat{\mathbf{Y}}_f$ será la resultante de cambiar los píxeles presuntamente perdido de la imagen observada \mathbf{Y}_{obs} por los píxeles estimados de $\hat{\mathbf{Y}}$ que les correspondan. Así, una estimación directa del bloque restaurado final es

$$\mathcal{P}_{i,j}\{\hat{\mathbf{Y}}_f\} = \text{mat} \left(\mathbf{y}_{obs\mathcal{P}_{i,j}} + (\mathbf{I} - \mathbf{M}_{\mathcal{P}_{i,j}})\hat{\mathbf{y}}_{\mathcal{P}_{i,j}} \right). \quad (3.9)$$

Donde \mathbf{I} es la matriz identidad y $\hat{\mathbf{y}}_{\mathcal{P}_{i,j}}$ es el bloque ordenado lexicográficamente correspondiente de la imagen estimada preliminar $\hat{\mathbf{Y}}$.

CAPÍTULO 4

SIMULACIÓN Y RESULTADOS

En este capítulo se describe el comportamiento del algoritmo *Weighted Median Regression* en el problema de restauración de píxeles perdidos, en comparación con el método de filtrado por mediana ponderada central (denotado como CWMF, del inglés *Center Weighted Median Filter*) [29,30], el método adaptativo de filtrado por mediana ponderada central (denotado como ACWMF, del inglés *Adaptive Center Weighted Median Filter*) [31] y la restauración basada en representación poco densa de señales a través de *Orthogonal Matching Pursuit*, usando diccionarios completos y sobrecompletos tanto preespecificados como entrenados. También, se presentan pruebas donde la imagen en estudio además de tener píxeles perdidos se encuentra contaminada con ruido modelado por otros tipos de distribución.

4.1 INTERPOLACIÓN DE PÍXELES PERDIDOS

En los capítulos anteriores se ha afirmado que el tema central de este trabajo es la interpolación o restauración de píxeles perdidos a través de técnicas de reconstrucción de señales poco densas. En esta sección se presentan simulaciones y curvas de errores o medidas de desempeño que ilustran el comportamiento de la metodología descrita en la Sección 3.2 del Capítulo 3. Inicialmente, no se considera la presencia de ruido aditivo u otro tipo de ruido en la imagen además de la contaminación por pérdida de información, por esto las imágenes completamente restauradas son iguales a las imágenes contaminadas excepto en los píxeles que han sido perdidos, los cuales son reemplazados por el valor correspondiente producto del proceso de estimación. Sin embargo, también se presentan resultados de la estimación total de las imágenes (todos sus píxeles) arrojadas por los métodos aplicados, para hacer una comparación más adecuada del desempeño de cada uno y cuantificar sus

capacidades de reconstrucción a partir de una cantidad de información disponible explícitamente menor. También, se asume como dato fundamental (como en la literatura del área de restauración de imágenes) la posición de los píxeles perdidos.

4.1.1 Condiciones de procesamiento

Los bloques de procesamiento para los algoritmos WMR y OMP son de tamaño 8×8 píxeles, lo que se traduce en una señal de procesamiento de $n = 64$ componentes luego del ordenamiento lexicográfico. El tamaño de los diccionarios seleccionados debe corresponder al tamaño del bloque en cuestión, así, los diccionarios utilizados deben ser de tamaño $64 \times K$, donde K es el número de átomos o columnas del diccionario. Principalmente se han seleccionado diccionarios redundantes: el diccionario DCT sobrecompleto bidimensional (en adelante referido como ODCT) y un diccionario entrenado por K-SVD de dimensiones 64×256 .

Aunque el problema inverso de (3.5) tiene una solución directa dada por $\hat{\mathbf{x}}_{\mathcal{P}_{i,j}} = \mathbf{A}^\dagger \mathbf{y}_{obs\mathcal{P}_{i,j}}$ cuando el diccionario \mathbf{A} es completo y ortogonal, se hace uso de un diccionario DCT completo bidimensional (en adelante referido como DCT) de tamaño 64×64 usando algoritmos de optimización debido a que una solución directa obtenida por inversión de la matriz (diccionario) tiende a producir errores elevados, por lo que se plantea estimar la señal con la menor cantidad de elementos no nulos en el dominio DCT bidimensional a partir de la cual se pueda obtener una imagen sin contaminación.

El comportamiento de los algoritmos de reconstrucción está íntimamente relacionado con la naturaleza de las imágenes que se deseen procesar, más aun cuando se utilizan diccionarios provenientes de algún método de entrenamiento, por esto los resultados y medidas de desempeño presentados en este capítulo son en función de la imagen estándar de procesamiento *Lena* redimensionada a 256×256 píxeles por razones de rapidez de procesamiento, seleccionada por ser una imagen muy completa en cuanto a detalles, texturas y zonas homogéneas. El diccionario entrenado por K-SVD utilizado ha sido construido a partir de los 62001 bloques (completamente solapados) que conforman una versión degradada por un filtro de mediana de ventana 3×3 de la imagen *Lena*, el método es inicializado por un diccionario ODCT y se aborda el problema de aproximación restringido por poca

densidad en (1.20) con $T_0 = 4$.

La figura 4.1 muestra la imagen original Lena de 256×256 píxeles y la versión ligeramente degradada usada para el entrenamiento. Seguido, la figura 4.2 muestra las bases reordenadas como bloques de los diccionarios utilizados.



Figura 4.1. (a) Imagen original de prueba Lena 256×256 píxeles. (b) Imagen de entrenamiento, Lena degradada por filtro de mediana de ventana 3×3 .

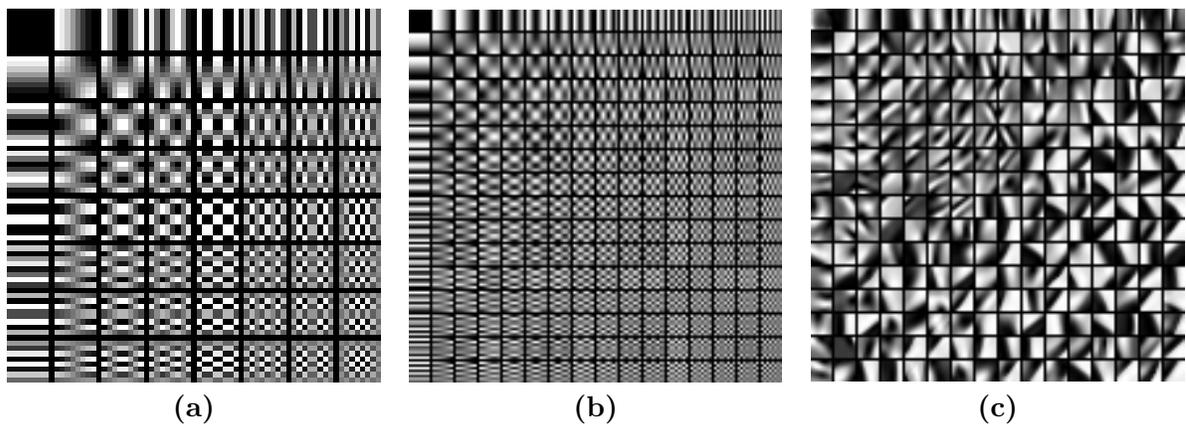


Figura 4.2. Diccionarios Seleccionados. (a) Bases del diccionario DCT completo 64×64 (DCT). (b) Bases del diccionario DCT sobre-completo 64×256 (ODCT). (c) Bases del diccionario entrenado por K-SVD 64×256 .

En comparación con los algoritmos que usan representación poco densa se presentan métodos convencionales de procesamiento de imágenes que han sido reportados como adecuados para tratamiento de contaminaciones de naturaleza impulsiva. En particular, el método de filtrado CWMF, el cual se aplica en una ventana de 5×5 con un peso de píxel

central $W_c = 5$ tomado como ejemplo de la referencia [30]. El método ACWMF, el cual se aplica en una ventana de 3×3 con parámetros: $s = 0.6$ y $[\delta_0; \delta_1; \delta_2; \delta_3] = [55; 40; 25; 15]$, como recomendaciones en la referencia [31].

La metodología seguida al aplicar OMP es la misma que al aplicar WMR con la distinción que el problema de minimización abordado esta dado por

$$\hat{\mathbf{x}}_{\mathcal{P}_{i,j}} = \arg \min_{\mathbf{x}_{\mathcal{P}_{i,j}}} \|\mathbf{y}_{obs\mathcal{P}_{i,j}} - \mathbf{V}_{\mathcal{P}_{i,j}} \mathbf{x}_{\mathcal{P}_{i,j}}\|_{\ell_2}^2 \quad \text{sujeto a} \quad \|\mathbf{x}_{\mathcal{P}_{i,j}}\|_{\ell_0} \leq T_0, \quad (4.1)$$

con $T_0 = 4$.

Además, se utilizó una versión acelerada del algoritmo WMR propuesta por A. Becerra en [22], donde se hace una estimación del soporte de la señal usando el subespacio. Los parámetros de entrada utilizados para el algoritmo son: $\beta = 0,95$, $\varepsilon = 1 \cdot 10^{-5}$, $\tau_i = \|\mathbf{A}^\dagger \mathbf{y}_{obs\mathcal{P}_{i,j}}\|_{\ell_\infty}$ y $K_0 = 140$. Se debe considerar que los métodos de aceleración [22] pueden incorporar algún decremento en el desempeño del algoritmo WMR en términos del error de reconstrucción [22], puesto que al acelerar el algoritmo se sacrifica un poco la robustez del mismo. El algoritmo utilizado tiene un parámetro adicional que indica el nivel de poca densidad necesitado, acá $T_0 = 4$. Como medida de desempeño se usa el PSNR (relación pico de la señal a ruido), MAE (error medio absoluto) y RMSE (raíz del error medio cuadrático), definidas en el Apéndice A.

4.1.2 Curvas de desempeño y simulaciones

Es importante hacer una comparación bastante rigurosa entre al algoritmo OMP y el algoritmo WMR, por esto se han construido curvas de desempeño de ambos métodos utilizando los distintos diccionarios seleccionados, para ilustrar en primera instancia qué diccionario es más apropiado para abordar el problema de restauración para ambos algoritmos y luego para establecer una comparación entre la mejor opción de cada algoritmo en contraposición a los desempeños presentados por los métodos más convencionales de filtrado ACWMF y CWMF.

En la figura 4.3 se muestran las curvas de desempeño (PSNR, MAE, RMSE) del algoritmo OMP haciendo un procesamiento de bloques no solapados para un porcentaje de pérdida de píxeles variable entre 5% y 70% con un paso de 5% entre cada punto de las

curvas. Las curvas de la izquierda corresponden a la realización de la reconstrucción total o completa de la imagen, es decir, la estimación total de los píxeles de la imagen a partir de la aplicación del método, las curvas de la derecha corresponden a la reconstrucción final de la imagen donde solo se reemplazan los píxeles perdidos por los correspondientes píxeles estimados. Cada punto de las curvas presentadas está estimado en base a 25 realizaciones aleatorias de pérdida de píxeles para el porcentaje de pérdida correspondiente. Se puede observar la mejoría del método cuando se utilizan diccionarios redundantes, tanto la utilización del diccionario entrenado por K-SVD como el diccionario ODCT propicia un mejor desempeño que al hacer uso de un diccionario completo DCT. Del mismo modo, la figura 4.4 muestra las distintas curvas de desempeño del algoritmo OMP para un procesamiento de bloques solapados completamente.

Bajo las mismas condiciones, en la figura 4.5 se muestran las distintas curvas de desempeño del algoritmo WMR para un procesamiento de bloques no solapados y análogamente la figura 4.6 muestra el desempeño al hacer uso de solapamiento completo de bloques. Se puede apreciar en ambas figuras que el desempeño de WMR usando un diccionario entrenado por K-SVD es el esperado, es decir, superior al desempeño correspondiente a la utilización de diccionarios preespecificados como el ODCT y el DCT. Sin embargo, este algoritmo tiene un comportamiento singular, y es que en el uso de los diccionarios provenientes de la transformada DCT no se cumple de manera notable la afirmación de que un diccionario sobrecompleto siempre llevará a mejores representaciones que un diccionario completo, por el contrario el desempeño para un diccionario DCT es ligeramente superior al desempeño para un diccionario ODCT, esto hace pensar que el algoritmo WMR es más sensible a los efectos del ruido en una imagen cuando se utilizan diccionarios redundantes en términos de la misma transformada.

Debido al costo computacional del algoritmo WMR y el procesamiento de bloques completamente solapados se optó por presentar en la figura 4.6 las curvas correspondientes a una versión optimizada de la versión acelerada de WMR utilizada, por esto se recomienda interpretar dicha figura en el contexto del comportamiento del algoritmo usando distintos diccionarios para el procesamiento, no en comparación con OMP.

Dado que de las figuras descritas anteriormente se puede apreciar que el comportamiento de ambos algoritmos es mejor en la utilización del diccionario entrenado por K-SVD,

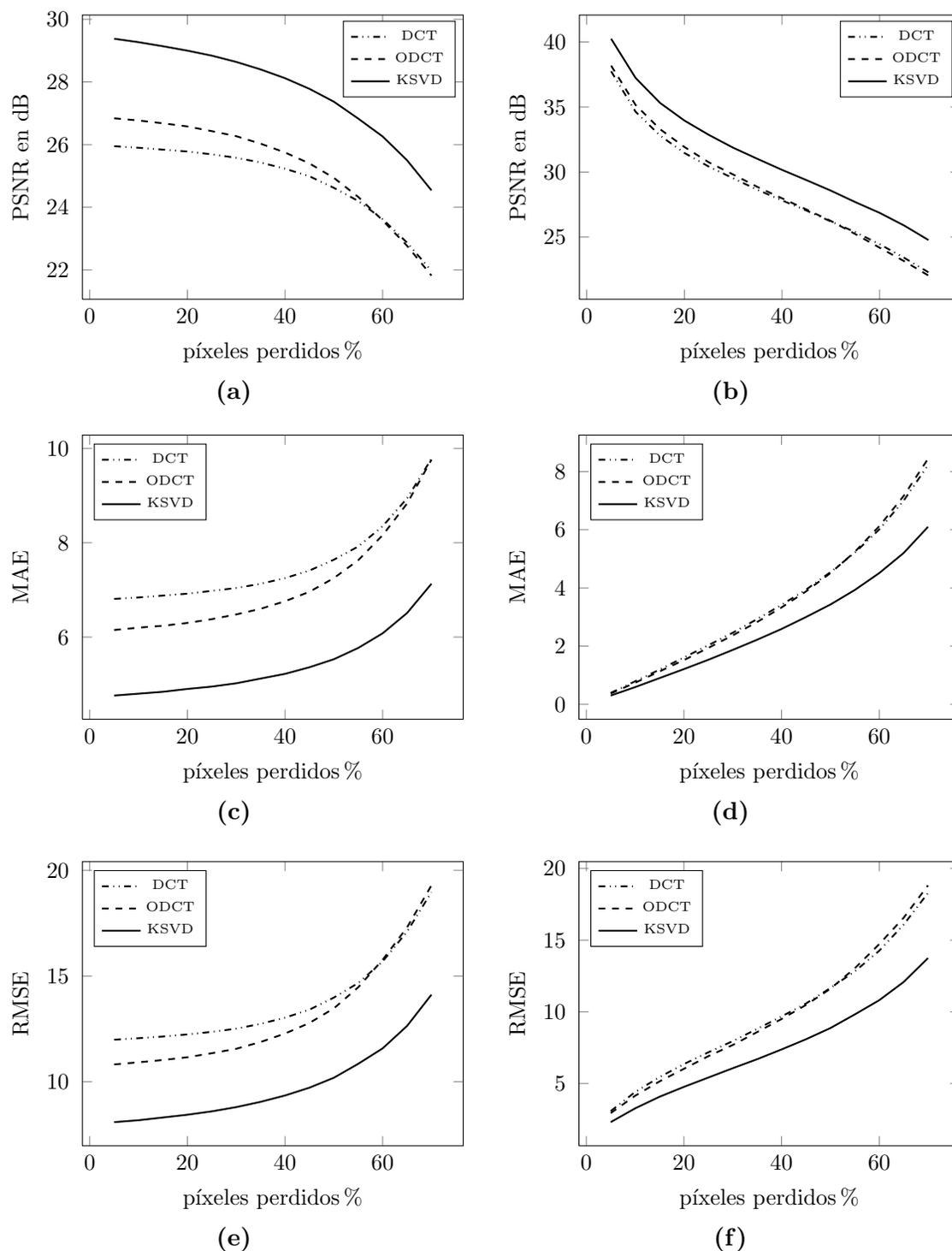


Figura 4.3. Desempeño del algoritmo OMP con distintos diccionarios, sin solapar bloques durante el procesamiento, para una variación de pérdida de píxeles entre 5% y 70%. (a) PSNR de la reconstrucción total de imagen (estimación de todos los píxeles de la imagen). (b) PSNR de la reconstrucción final (cambiando solo los píxeles perdidos por los estimados). (c) MAE de la reconstrucción total. (d) MAE de la reconstrucción final. (e) RMSE de la reconstrucción total. (f) RMSE de la reconstrucción final.

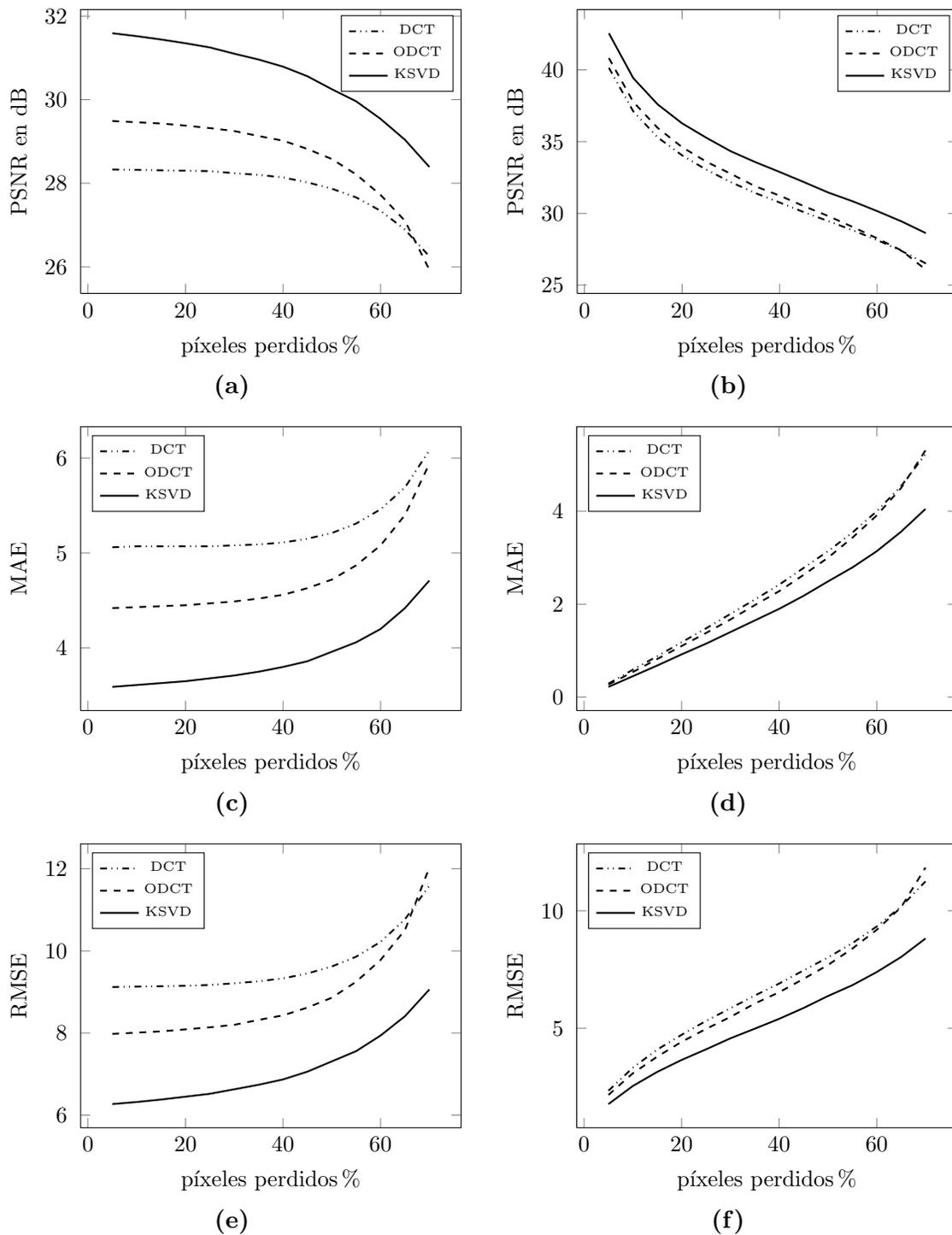


Figura 4.4. Desempeño del algoritmo OMP con distintos diccionarios, usando solapamiento completo de bloques en el procesamiento, para una variación de pérdida de píxeles entre 5% y 70%. (a) PSNR de la reconstrucción total de imagen (estimación de todos los píxeles de la imagen). (b) PSNR de la reconstrucción final (cambiando solo los píxeles perdidos por los estimados). (c) MAE de la reconstrucción total. (d) MAE de la reconstrucción final. (e) RMSE de la reconstrucción total. (f) RMSE de la reconstrucción final.

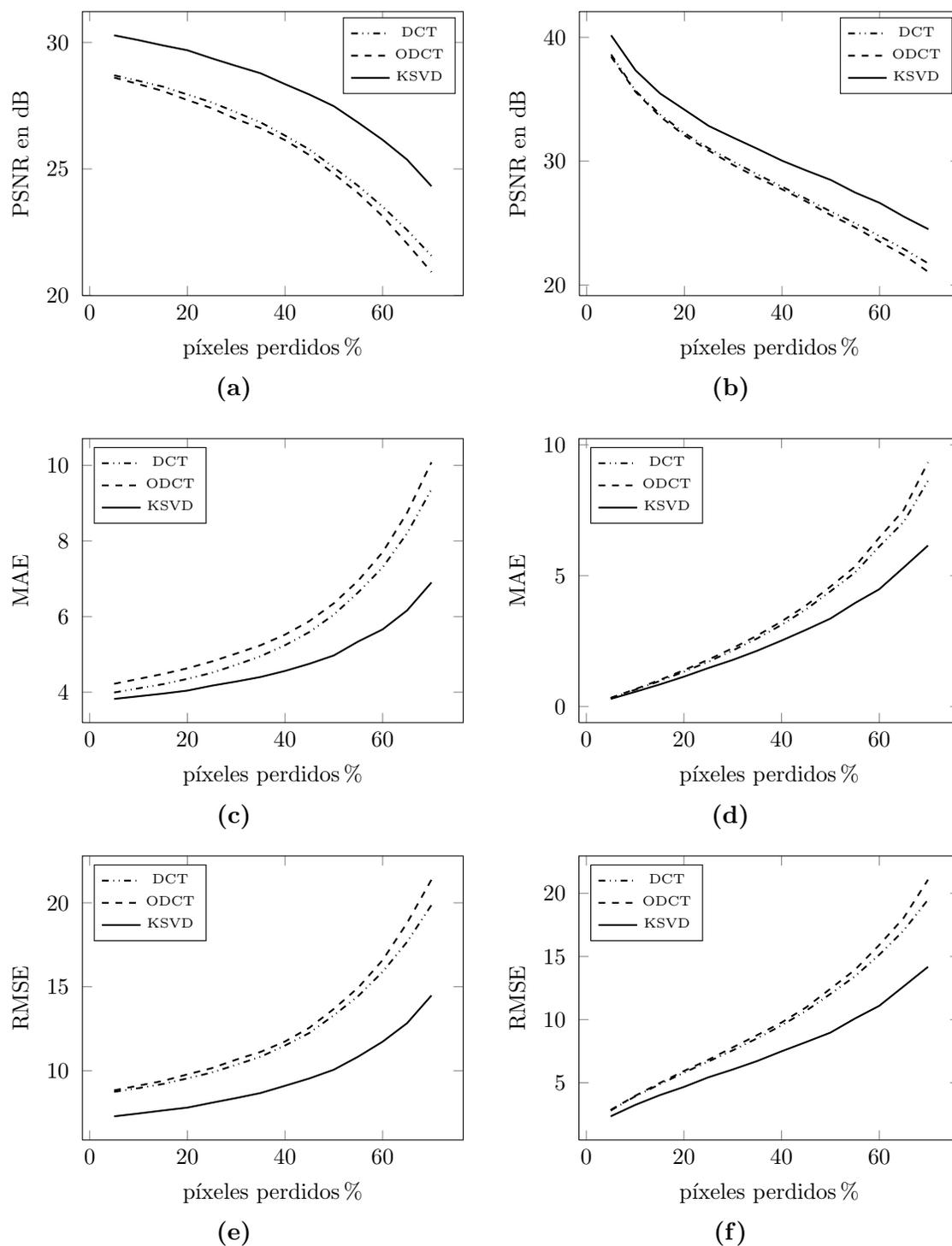


Figura 4.5. Desempeño del algoritmo WMR con distintos diccionarios, sin solapar bloques durante el procesamiento, para una variación de pérdida de píxeles entre 5% y 70%. (a) PSNR de la reconstrucción total de imagen (estimación de todos los píxeles de la imagen). (b) PSNR de la reconstrucción final (cambiando solo los píxeles perdidos por los estimados). (c) MAE de la reconstrucción total. (d) MAE de la reconstrucción final. (e) RMSE de la reconstrucción total. (f) RMSE de la reconstrucción final.

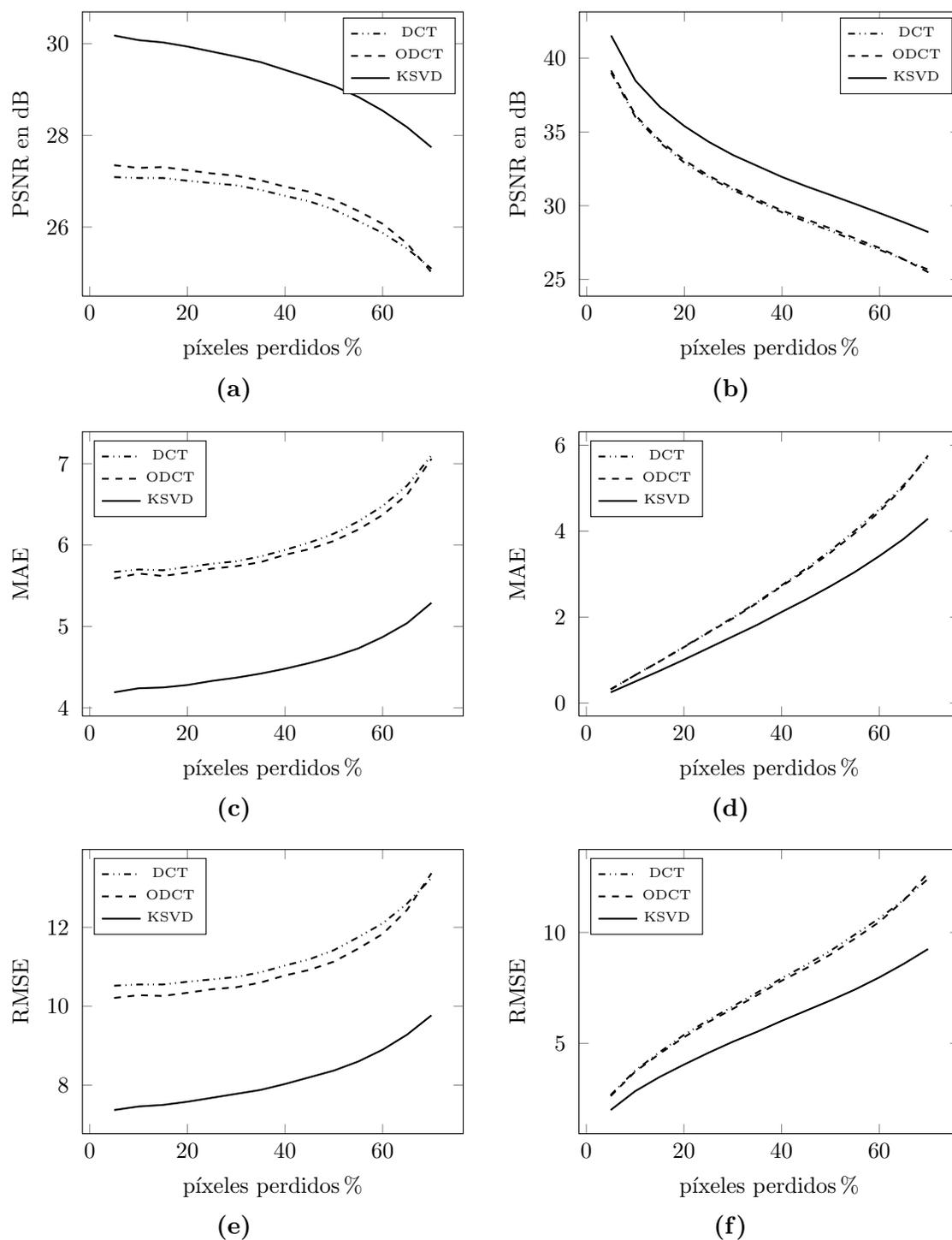


Figura 4.6. Desempeño del algoritmo WMR (versión acelerada optimizada) con distintos diccionarios, usando solapamiento completo de bloques en el procesamiento, para una variación de pérdida de píxeles entre 5% y 70%. (a) PSNR de la reconstrucción total de imagen (estimación de todos los píxeles de la imagen). (b) PSNR de la reconstrucción final (cambiando solo los píxeles perdidos por los estimados). (c) MAE de la reconstrucción total. (d) MAE de la reconstrucción final. (e) RMSE de la reconstrucción total. (f) RMSE de la reconstrucción final.

se ha preparado en la figura 4.7 para procesamiento sin solapamiento de bloques y en la figura 4.8 para solapamiento completo, una comparación del desempeño del algoritmo OMP y el algoritmo WMR (versión acelerada) usando el diccionario entrenado por K-SVD, se compara además con el desempeño de los métodos ACWMF y CWMF aplicados con parámetros recomendados en la literatura.

En estas figuras se puede apreciar que los métodos de restauración de imágenes basados en el problema de representación poco densa de señales son significativamente superiores a métodos convencionales de filtrado como ACWMF o CWMF, sobre todo cuando la contaminación o en este caso la pérdida de píxeles es elevada. La razón fundamental es precisamente la resolución del problema de representación poco densa de señales, puesto que el vector de coeficientes estimados tenderá siempre con alta probabilidad al vector de coeficientes óptimo para realizar la restauración, además y en el contexto de una restauración finalizada la imagen resultante no necesariamente será la mejor estimación realizada a partir de las proyecciones de los coeficientes obtenidos sobre el diccionario seleccionado, sino que se puede refinar aún más la restauración si se aplica el proceso de restauración ahora sobre ella, y así hasta alcanzar una convergencia en la restauración. A diferencia de los métodos de filtrado convencional habituales como CWMF dado que en cada estimación la imagen tendería a deteriorarse, ya que por la aplicación del operador de mediana directamente sobre la imagen se perderían progresivamente por ejemplo los detalles, las texturas y los bordes, o para un método adaptativo como ACWMF, que es además recursivo, la primera estimación es posiblemente la mejor porque un segundo procesamiento sobre la primera imagen reconstruida no cumpliría con lo necesario para que el método detecte los píxeles ruidosos (que en una segunda estimación ya tendrían un valor estimado previamente) por lo que la imagen a la salida sería igual a la de entrada.

En las figuras 4.7 y 4.8, los comportamientos de los algoritmos de optimización se aprecian similares a causa de las curvas correspondientes a los métodos de filtrado, si bien es cierto que el comportamiento es similar, en las curvas de la izquierda el algoritmo WMR presenta una mejora sobre el algoritmo OMP para contaminaciones entre 5% y 55%. Por otra parte, en las curvas de la derecha el comportamiento de los algoritmos de optimización es muy similar, y esto es esperado debido a que una gran cantidad de píxeles (píxeles no perdidos) entre las reconstrucciones de ambos métodos tienen el mismo valor, puesto que

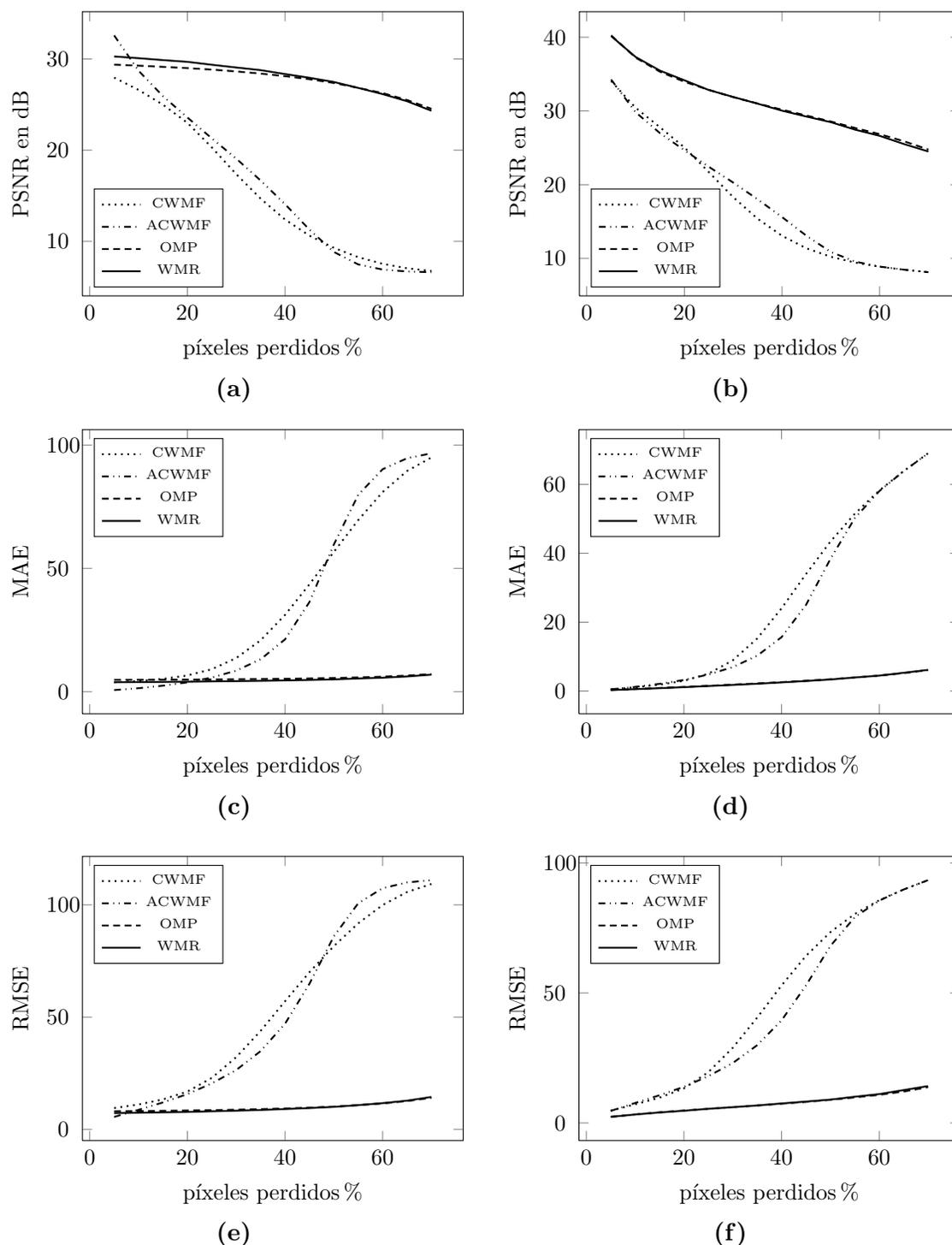


Figura 4.7. Desempeño de los algoritmos WMR y OMP usando diccionario entrenado por K-SVD sin solapamiento de bloques, y desempeño de los métodos ACWMMF y CWMMF, para una variación de pérdida de píxeles entre 5% y 70%. (a) PSNR de la reconstrucción total de imagen (estimación de todos los píxeles de la imagen). (b) PSNR de la reconstrucción final (cambiando solo los píxeles perdidos por los estimados). (c) MAE de la reconstrucción total. (d) MAE de la reconstrucción final. (e) RMSE de la reconstrucción total. (f) RMSE de la reconstrucción final.

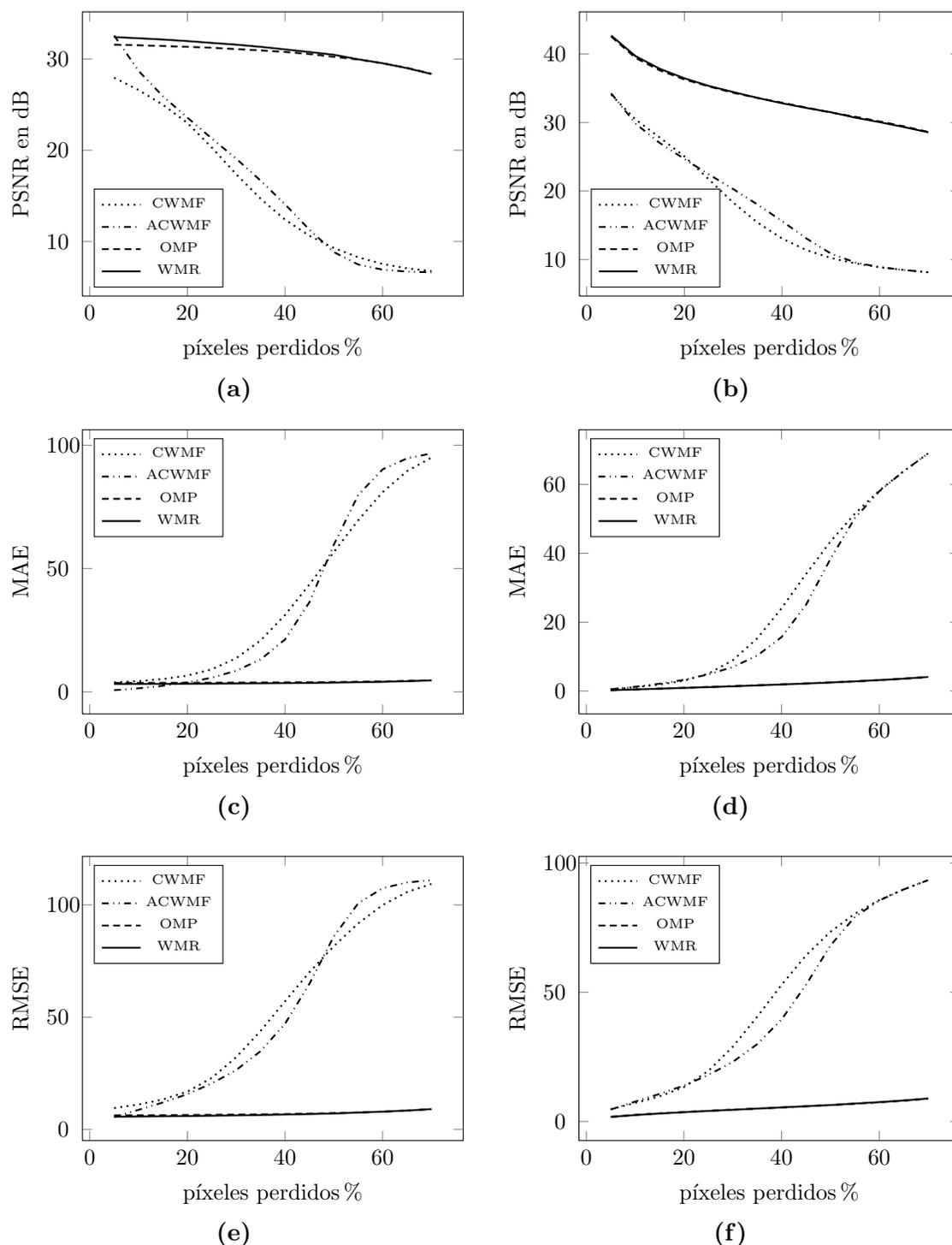


Figura 4.8. Desempeño de los algoritmos WMR y OMP usando diccionario entrenado por K-SVD usando solapamiento completo de bloques, y desempeño de los métodos ACWWMF y CWMF, para una variación de pérdida de píxeles entre 5% y 70%. (a) PSNR de la reconstrucción total de imagen (estimación de todos los píxeles de la imagen). (b) PSNR de la reconstrucción final (cambiando solo los píxeles perdidos por los estimados). (c) MAE de la reconstrucción total. (d) MAE de la reconstrucción final. (e) RMSE de la reconstrucción total. (f) RMSE de la reconstrucción final.

mantienen el valor original que no es cambiado en la reconstrucción final.

El costo computacional del algoritmo WMR es muy elevado, lo que se traduce en una restauración por bloques completamente solapados que tarda unos 14700 s en realizarse para una imagen de 256×256 píxeles sobre un hardware PC Windows 7 x64 Intel Core i3 2100 3,1 GHz, en contraposición al algoritmo OMP para el cual una realización de bloques completamente solapados tarda unos 110 s. Las reconstrucciones realizadas con las versiones aceleradas de WMR aún tienen un costo elevado pero en tiempo significativamente menor que el algoritmo WMR original, a la versión acelerada de WMR utilizada en esta sección le toma para una restauración con bloques completamente solapados unos 4900 s sobre el mismo hardware.

En la figura 4.9 se presenta una simulación de la restauración completa, es decir, todos los píxeles de la imagen *Lena* para una realización de pérdida de píxeles aleatoria correspondiente al 20 % del total de píxeles de la imagen, utilizando el diccionario entrenado por K-SVD de la figura 4.2(c) y un procesamiento *bloque por bloque* de solapamiento completo para los algoritmos WMR y OMP en comparación con los métodos de filtrado ACWMF y CWMF, donde se puede apreciar el comportamiento de la metodología planteada y la superioridad de métodos de restauración de imágenes basados en representación poco densa de señales. La figura 4.10 ilustra los resultados para una restauración final donde los píxeles perdidos son los únicos cambiados por sus correspondientes estimados en todos los métodos. En estas pruebas se puede apreciar el comportamiento de los métodos bajo los mejores enfoques de procesamiento.

El punto clave en el desempeño de la metodología propuesta está en el pleno conocimiento de la matriz de pérdida de píxeles, que al ser desconocida puede ser estimada utilizando métodos como el ACWMF. Es esta información lo que brinda principalmente una superioridad en el desempeño de los algoritmos de reconstrucción frente a los métodos de filtrado convencional, conocer \mathbf{M} para cada bloque procesado implica un mejor aprovechamiento de la información para formar una imagen restaurada. En la siguiente sección se presentan algunas simulaciones donde se desconoce por completo \mathbf{M} para cada bloque de procesamiento. También, se presenta una prueba de desoclusión de imágenes donde sí se conoce la máscara que ocasiona la oclusión de secciones que serán restauradas.



Figura 4.9. Restauración completa (todos los píxeles) de imagen para 20 % de píxeles perdidos. (a) Imagen Original. (b) Imagen con 20 % de píxeles perdidos. (c) Imagen restaurada con ACWMF. $PSNR = 23,52$ dB, $MAE = 3,83$, $RMSE = 15,87$. (d) Imagen restaurada con CWMF. $PSNR = 23,01$ dB, $MAE = 6,60$, $RMSE = 16,83$. (e) Imagen restaurada con OMP. $PSNR = 31,32$ dB, $MAE = 3,65$, $RMSE = 6,46$. (f) Imagen restaurada con WMR. $PSNR = 31,94$ dB, $MAE = 3,26$, $RMSE = 6,02$.



Figura 4.10. Restauración final (cambiando sólo píxeles perdidos) de imagen para 20% de píxeles perdidos. (a) Imagen Original. (b) Imagen con 20% de píxeles perdidos. (c) Imagen restaurada con ACWMF. $PSNR = 24,60$ dB, $MAE = 3,27$, $RMSE = 14,02$. (d) Imagen restaurada con CWMF. $PSNR = 25,10$ dB, $MAE = 2,99$, $RMSE = 13,22$. (e) Imagen restaurada con OMP. $PSNR = 36,19$ dB, $MAE = 0,92$, $RMSE = 3,69$. (f) Imagen restaurada con WMR. $PSNR = 36,27$ dB, $MAE = 0,90$, $RMSE = 3,65$.

4.2 SIMULACIONES PARA OTRAS CONTAMINACIONES

En esta sección se presentan simulaciones específicas para ciertos tipos de contaminaciones como: oclusión de sectores de la imagen a través de una máscara, ruido sal y pimienta, ruido impulsivo asimétrico como píxeles perdidos de posiciones desconocidas. Para las simulaciones con ruido se descarta el uso del diccionario holográfico \mathbf{V} en la metodología de reconstrucción.

En la figura 4.12 se muestra una prueba de restauración donde la imagen original ha sido enmascarada con la porción de texto visible en la figura 4.11(a), donde la parte de color negro es de valor 0, a la que le corresponde la máscara de contaminación de la figura 4.11(b) a partir de la cual se extrae la matriz de contaminación \mathbf{M} para cada bloque, ésta tiene valor 0 en la posiciones de la máscara donde existe texto, considerando las variaciones en la escala de color presente en los bordes del mismo, y valor 1 en el resto de la matriz. Como se muestra en la figura 4.12 los resultados de la prueba indican que la restauración utilizando los algoritmos de reconstrucción OMP y WMR es relativamente similar, con una mejora en la restauración usando WMR de 0,35 dB de PSNR respecto a OMP cuando se estiman todos los píxeles de la imagen. Sin embargo, y para este caso la estimación de los píxeles dentro de la máscara parece no ser la mejor, ya que cuando se cambian los píxeles enmascarados por sus correspondientes estimados a través de WMR la reconstrucción tiene un desempeño menor comparada con OMP de unos 0,20 dB.

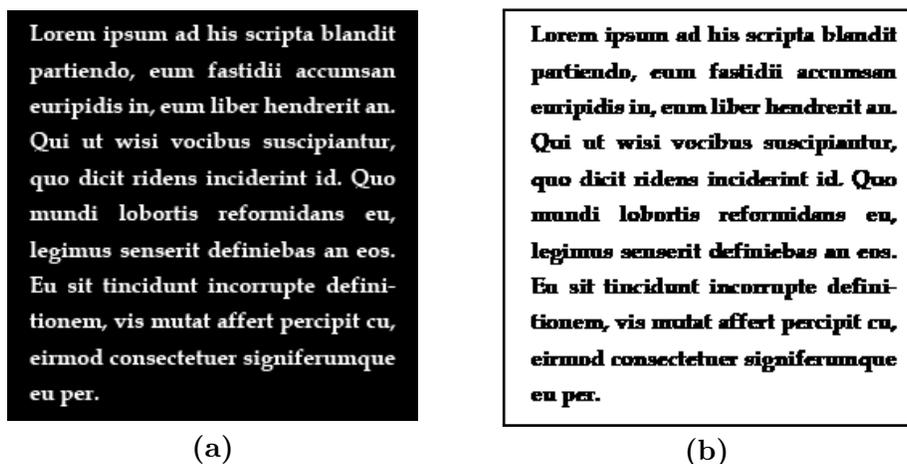


Figura 4.11. Máscara de texto aplicada en la prueba de desocclusión de imagen. (a) Máscara aplicada, el texto es colocado sobre la imagen. (b) Máscara de contaminación.



Figura 4.12. Prueba de desoclusión para una imagen enmascarada con texto. (a) Imagen Original. (b) Imagen enmascarada. (c) Reconstrucción total (todos sus píxeles) de la imagen usando OMP. $PSNR = 30,32$ dB, $MAE = 3,95$, $RMSE = 7,25$. (d) Reconstrucción total de la imagen usando WMR. $PSNR = 30,67$ dB, $MAE = 3,57$, $RMSE = 6,98$. (e) Reconstrucción final (cambiado solo píxeles enmascarados) de la imagen usando OMP. $PSNR = 33,82$ dB, $MAE = 1,11$, $RMSE = 4,85$. (f) Reconstrucción final de la imagen usando WMR. $PSNR = 33,62$ dB, $MAE = 1,11$, $RMSE = 4,96$.

En la figura 4.13 se muestra una restauración de la imagen *Lena* bajo las misma pérdida de píxeles de la prueba presentada en las figuras 4.9 y 4.10, esta vez desconociendo la matriz de pérdida que actúa sobre cada bloque. Los resultados obtenidos expresan las ventajas del algoritmo WMR para reconstrucción de la pérdida de píxeles considerada como contaminación por ruido impulsivo asimétrico, sobre el algoritmo OMP y los métodos de filtrado ACWMF y CWMF. Obsérvese que la medida de MAE con menor valor, que indica la desviación entre los valores absolutos de los píxeles estimados y los originales, le corresponde al método ACWMF, esto es porque este método preserva píxeles que no asume contaminados para formar la imagen restaurada. Sin embargo, las otras medidas de desempeño, PSNR y RMSE, indican que la mejor estimación general correspondiente al espacio de la imagen es obtenida por el algoritmo WMR. La mejora respecto a OMP es bastante significativa, de 7,04 dB, esto es por la alta vulnerabilidad de este método frente a los efectos del ruido. La mejora respecto a ACWMF es de 1,78 dB, que aunque no es una estimación que conserva mejor los detalles de la imagen es una estimación más uniforme, y respecto a CWMF de 2.29 dB. Nótese también que la restauración con WMR de la figura 4.13(f) tiene un desempeño menor significativo que las restauraciones en las figuras 4.9(f) y 4.10(f), producto de asumir desconocida la matriz de pérdida de píxeles por tanto la presencia en el procesamiento del diccionario holográfico.

En la figura 4.14 se presenta una prueba de restauración de la imagen *Lena* contaminada con 10% de ruido impulsivo simétrico de tipo “sal y pimienta”. En esta simulación los métodos convencionales son superiores a los algoritmos de reconstrucción WMR y OMP, recuerde que se descarta el uso del diccionario holográfico \mathbf{V} ya que se asumen desconocidas la posiciones de los impulsos en la imagen. Sin embargo, note la superioridad del algoritmo WMR (versión acelerada) en la restauración frente al algoritmo OMP cuya restauración es evidentemente más vulnerable a la presencia de los impulsos.

En la figura 4.15 se somete la imagen a una contaminación de ruido Gaussiano aditivo blanco con una relación señal a ruido $SNR = 15$ dB y a una pérdida de píxeles del 20%. En dicha figura se presentan restauraciones resultantes del procesamiento de bloques completamente solapados teniendo en cuenta el uso de la matriz de pérdida de píxeles en cada bloque, que define el diccionario holográfico \mathbf{V} , y descartándola completamente. Es evidente la superioridad de la metodología cuando se considera conocida la matriz de



Figura 4.13. Restauración completa (todos los píxeles) de imagen para 20% de píxeles perdidos sin uso del diccionario holográfico. (a) Imagen Original. (b) Imagen con 20% de píxeles perdidos. (c) Imagen restaurada con ACWMF. $PSNR = 23,52$ dB, $MAE = 3,83$, $RMSE = 15,87$. (d) Imagen restaurada con CWMF. $PSNR = 23,01$ dB, $MAE = 6,60$, $RMSE = 16,83$. (e) Imagen restaurada con OMP. $PSNR = 18,26$ dB, $MAE = 21,15$, $RMSE = 29,08$. (f) Imagen restaurada con WMR. $PSNR = 25,30$ dB, $MAE = 7,85$, $RMSE = 12,92$.



Figura 4.14. Restauración de imagen con 10% ruido impulsivo simétrico de tipo “sal y pimienta”. (a) Imagen Original. (b) Imagen contaminada con 10% de ruido “sal y pimienta”. (c) Imagen restaurada con ACWMF. $PSNR = 32,05$ dB, $MAE = 0,90$, $RMSE = 5,94$. (d) Imagen restaurada con CWMF. $PSNR = 27,92$ dB, $MAE = 3,86$, $RMSE = 9,56$. (e) Imagen restaurada con OMP. $PSNR = 22,01$ dB, $MAE = 12,85$, $RMSE = 18,88$. (f) Imagen restaurada con WMR. $PSNR = 27,85$ dB, $MAE = 5,87$, $RMSE = 9,64$.

pérdida incluso cuando los píxeles disponibles para la estimación están contaminados, aunque la reconstrucción no es del todo limpia, para esto se necesita aplicar los métodos con ciertas modificaciones (dependientes de la varianza del ruido) para que minimicen los efectos del ruido aditivo en la imagen. Además con esta prueba se enfatiza en el hecho de que el algoritmo WMR es más robusto en la restauración de píxeles perdidos considerados como ruido impulsivo asimétrico en un procedimiento donde se descarta el conocimiento de la matriz de contaminación correspondiente a cada bloque.



Figura 4.15. Restauración de imagen con ruido Gaussiano aditivo blanco con $SNR = 15$ dB. (a) Imagen Original. (b) Imagen contaminada con ruido Gaussiano aditivo blanco con $SNR = 15$ dB. (c) Imagen restaurada con OMP usando el diccionario holográfico. $PSNR = 27,15$ dB, $MAE = 7,86$, $RMSE = 10,45$. (d) Imagen restaurada con WMR usando el diccionario holográfico. $PSNR = 26,58$ dB, $MAE = 8,55$, $RMSE = 11,15$. (e) Imagen restaurada con OMP descartando el uso del diccionario holográfico. $PSNR = 18,10$ dB, $MAE = 21,98$, $RMSE = 29,60$. (f) Imagen restaurada con WMR descartando el uso del diccionario holográfico. $PSNR = 22,50$ dB, $MAE = 13,08$, $RMSE = 17,86$.

CONCLUSIONES

- El algoritmo *Weighted Median Regression* es una alternativa muy atractiva para abordar procedimientos de restauración de imágenes. Bajo ciertas condiciones presenta incluso un desempeño superior que el algoritmo *Orthogonal Matching Pursuit*. Sin embargo, es un algoritmo de alto costo computacional, cuya ejecución y la de cualquiera de sus variantes depende de distintos parámetros que deben ser sintonizados de acuerdo a los requerimientos del problema para poder garantizar la mejor de las restauraciones a la salida del método, estos parámetros son: control de descenso en la umbralización fuerte (β), número máximo de iteraciones (K_0) y energía residual objetivo (ϵ). Por otra parte, el costo computacional del algoritmo asociado a la aplicación del operador de mediana ponderada no es una razón de peso para hacerlo descartable en el planteamiento de algún procedimiento, debido a los avances en computación que permiten cada vez más procedimientos de alta complejidad casi sin medir las capacidades que el hardware tenga para responder a ellos. Es su robustez lo que le hace un método elegible y altamente considerable en cualquier aplicación que involucre el problema de representación poco densa de señales.
- El uso de diccionarios entrenados por K-SVD agrega más costo computacional a cualquier metodología de reconstrucción, pero esto es una desventaja casi irrelevante, puesto que las restauraciones obtenidas son altamente superiores en comparación con el uso de diccionarios provenientes de bases de la transformada discreta del coseno (las de uso más común). La utilización de un diccionario entrenado no depende necesariamente de una estimación previa de la imagen contaminada como se ha presentado en este trabajo, sino que en principio depende de datos que comparan la naturaleza de la imagen contaminada, por esto es habitual la utilización de este método de entrenamiento en conjuntos o bases de datos de imágenes de cierta naturaleza bien definida, por ejemplo: rostros, huellas, pictogramas, etc.
- La metodología seguida para realizar restauración de píxeles perdidos, o las distintas

restauraciones estudiadas en este trabajo, no es única y se puede efectuar de distintas maneras, incorporar variantes y permitir el paso a una metodología más compleja de procesamiento. La metodología propuesta es un enfoque sencillo con el que se intenta promover una igualdad de condiciones para ambos métodos (OMP y WMR) con el fin de hacer una comparación correcta y mostrar las bondades del algoritmo WMR. Se puede incluso proponer recursividad en el procesamiento de bloques, garantizando un aumento en la información disponible (en principio más confiable) para la estimación de bloques subsiguientes, o se puede proponer iterar el procedimiento completo de restauración sobre la imagen resultante para refinar los valores obtenidos.

- Como se mostró en este trabajo, el uso de representación poco densa de señales como parte fundamental de la restauración de imágenes dio un mejor desempeño con respecto a las técnicas de filtrado convencional, las ventajas son significativas, y su aplicación en conjunto con diccionarios más adecuados es en efecto una gran propuesta en el área de procesamiento de imágenes médicas, arte, cine y televisión. Sin embargo, es importante señalar que la aplicabilidad de estos conceptos va de la mano con el desarrollo de métodos de detección de ruido y perturbaciones cada vez más complejos, puesto que el conocimiento previo de la naturaleza de estas llevará con mayor probabilidad a la mejor estimación posible para construir una imagen restaurada.

RECOMENDACIONES

El algoritmo *Weighed Median Regression* se propone aquí como una alternativa para abordar el problema de restauración de imágenes basada en representación poco densa de señales, este algoritmo fue propuesto inicialmente para la teoría de sensado comprimido aplicada a señales unidimensionales. En el caso particular de imágenes, aún conociendo las ventajas del algoritmo, es necesario atender ciertas características con el fin de hacerlo un método completamente adecuado para trabajar con grandes conjuntos de datos, también es atractivo el planteamiento de una metodología de procesamiento más compleja para brindar mayor robustez a la estimación, por esto se hacen las siguientes recomendaciones para trabajo futuro:

- Estudiar la paralelización del algoritmo WMR. Desarrollar rutinas de procesamiento que aprovechen los avances en programación paralela, que permitan una ejecución rápida y la explotación de éste método en su totalidad, utilizando herramientas de hardware para procesamiento gráfico como por ejemplo el lenguaje CUDA de NVIDIA, o la implementación de las rutinas en lenguajes compilados que proporcionen mayor velocidad de procesamiento.
- Aprovechar enfoques recursivos en la metodología de procesamiento para garantizar en primera instancia una mejor estimación que realizando un procesamiento con la metodología propuesta. Se prevé que considerar recursividad en el procesamiento *bloque por bloque* cuando se hace uso de algún solapamiento entre bloques la estimación de cada uno de ellos dependerá en algún índice (dependiendo del índice de solapamiento) del bloque estimado anterior, garantizando así, si la estimación previa es correcta, que en la estimación actual esté involucrada mayor información veraz y menor ruido.
- Se propone el uso del operador de *orden estadístico ponderado* (denotado como WOS, del inglés *Weighted Order Statistic*) [32] que es un operador más robusto para tratamiento de ruido cuyas distribuciones corresponden a colas pesadas asimétricas, como

una pérdida de píxeles, frente al operador de mediana ponderada propio del algoritmo WMR. Con este operador más robusto se podría realizar una mejor restauración cuando no se conoce la matriz de pérdida de píxeles, debido a que se encuentra equipado con un parámetro adicional que controla el rango de la salida del operador sobre el conjunto de muestras replicadas de acuerdo a las ponderaciones correspondientes.

APÉNDICE A

DEFINICIONES DE ERRORES

En este apéndice se describe la ecuación matemática que define cada una de las distintas medidas de desempeño utilizadas en este trabajo aplicadas a los procesos de restauración. Todas las medidas se definen en función de las imágenes en cuestión: la imagen original \mathbf{Y} de tamaño $M \times N$ y la imagen restaurada $\hat{\mathbf{Y}}$ obtenida a través de la imagen contaminada observada \mathbf{Y}_{obs} . Las medidas *relación señal a ruido pico* y *raíz del error medio cuadrático* se definen en función del error medio cuadrático (MSE).

- Error medio cuadrático (del inglés, *Mean Square Error*):

$$MSE = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (Y_{i,j} - \hat{Y}_{i,j})^2 \quad (\text{A.1})$$

- Relación pico de la señal a ruido (del inglés *Peak Signal-to-Noise Ratio*):

$$PSNR = 10 \cdot \log_{10} \left(\frac{[\max(\mathbf{Y})]^2}{MSE} \right) \quad (\text{A.2})$$

- Error medio absoluto (del inglés *Mean Absolute Error*):

$$MAE = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N |Y_{i,j} - \hat{Y}_{i,j}| \quad (\text{A.3})$$

- Raíz del error medio cuadrático (del inglés *Root-Mean Square Error*):

$$RMSE = \sqrt{MSE} \quad (\text{A.4})$$

REFERENCIAS

- [1] D. L. Donoho, “Compressed Sensing,” *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [2] R. G. Baraniuk, “Compressive Sensing [lecture notes],” *IEEE Signal Processing Magazine*, vol. 24, no. 4, pp. 118–121, Jul. 2007.
- [3] J. A. Rodríguez, “Estudio de los algoritmos de reconstrucción en compressed sensing,” Trabajo Especial de Grado, Departamento de Física, Facultad de Ciencias, Universidad de Los Andes, Mérida, Venezuela, 2008.
- [4] G. Davis, S. Mallat, y M. Avellaneda, “Adaptive greedy approximations,” *Constructive Approximations*, vol. 13, no. 1, pp. 57–98, 1997.
- [5] J. A. Tropp y A. C. Gilbert, “Signal recovery from random measurements via orthogonal matching pursuit,” *IEEE Transactions on Signal Processing*, vol. 53, no. 12, pp. 4655–4666, 2007.
- [6] J. L. Paredes y G. R. Arce, “Compressive sensing signal reconstruction by weighted median regression estimates,” *IEEE Transactions on Signal Processing*, vol. 59, no. 6, pp. 2585–2601, Jun. 2011.
- [7] S.-J. Kim, K. Koh, M. Lustig, S. Boyd, y D. Gorinevsky, “An interior-point method for large-scale ℓ_1 -regularized least squares,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 606–617, 2007.
- [8] J. A. Tropp, “Topics in sparse approximation,” Tesis de Ph.D., The University of Texas at Austin, Austin, Ago. 2004.
- [9] S. Boyd y L. Vandenberghe, *Convex Optimization*. Nueva York, EE.UU.: Cambridge University Press, 2004.

- [10] S. A. Khayam, “The discrete cosine transform (DCT): Theory and application,” Michigan State University, Department of Electrical & Computer Engineering, ECE 802-602: Information Theory and Coding, Mar. 2003.
- [11] B. J. Broxon, “The Kronecker product,” Tesis de Maestría, College of Arts and Sciences, University of North Florida, 2006.
- [12] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. Haifa, Israel: Springer, 2010.
- [13] M. Aharon, M. Elad, y A. Bruckstein, “K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation,” *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.
- [14] K. Egan, S. O. Aase, y J. Hakon-Husoy, “Method of optimal directions for frame design,” en *IEEE ICASSP*, vol. 5, Phoenix, AZ, EE.UU., 1999, pp. 2443–2446.
- [15] B. A. Olshausen y D. J. Field, “Sparse coding with an overcomplete basis set: A strategy employed by V1?” *Vision Research*, vol. 37, no. 23, pp. 3311–3325, 1997.
- [16] K. Kreutz-Delgado, J. F. Murray, B. D. Rao, K. Engan, T.-W. Lee, y T. J. Sejnowski, “Dictionary learning algorithms for sparse representation,” *Neural Computation*, vol. 15, no. 2, pp. 349–396, 2003.
- [17] S. Lesage, R. Gribonval, F. Bimbot, y L. Benaroya, “Learning unions of orthonormal bases with thresholded singular value decomposition,” en *IEEE ICASSP*, vol. 5, Phoenix, AZ, EE.UU., Mar.18–23, 2005, pp. 293–296.
- [18] M. Aharon, “Overcomplete dictionaries for sparse representation of signals,” Tesis de Ph.D., The Technion - Israel Institute of Technology, Haifa, Nov. 2006.
- [19] J. S. Lim, *Two-dimensional signal and image processing*. Englewood Cliffs, NJ, EE.UU.: Prentice Hall, 1990.
- [20] R. Rubinstein, M. Zibulevsky, y M. Elad, “Efficient implementation of the K-SVD algorithm using batch orthogonal matching pursuit,” The Technion - Israel Institute of Technology, Reporte Técnico, 2008.

- [21] S. Mallat y Z. Zhang, “Matching pursuits with time-frequency dictionaries,” *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3397–3415, 1993.
- [22] A. Becerra, J. Ordenes, y J. L. Paredes, “Métodos de aceleración del algoritmo de mediana ponderada regresiva,” en *IV BIOVEN Congreso Venezolano de Bioingeniería*, San Cristóbal, Venezuela, Nov.12–15 2012.
- [23] S. Masnou, “Disocclusion: a variational approach using level lines,” *IEEE Transactions on Image Processing*, vol. 11, no. 2, pp. 68–76, 2002.
- [24] M. Bertalmio, G. Sapiro, C. Ballester, y V. Caselles, “Image Inpainting,” en *SIGGRAPH 2000*, New Orleans, EE.UU., Jul. 2000, pp. 417–424.
- [25] R. H. Chan, C. Hu, y M. Nikolova, “An iterative procedure for removing random-valued impulse noise,” *IEEE Signal Processing Letters*, vol. 11, no. 12, pp. 921–924, Dic. 2004.
- [26] M. J. Fadili, J. L. Starck, y F. Murtagh, “Inpainting and zooming using sparse representations,” *The Computer Journal*, vol. 52, no. 1, pp. 64–79, Ene. 2007.
- [27] B. Deka y P. K. Bora, “Removal of random-valued impulse noise using sparse representation,” en *National Conference on Communications (NCC)*, Bangalore, India, Ene.28–30, 2011, pp. 1–5.
- [28] R. Gonzalez, R. Woods, y S. Eddins, *Digital Image Processing Using Matlab*. Pearson Prentice Hall, 2004.
- [29] S.-J. Ko y Y. H. Lee, “Center weighted median filters and their applications to image enhancement,” *IEEE Transactions on Circuits and Systems*, vol. 38, no. 9, pp. 984–993, Sep. 1991.
- [30] G. R. Arce, J. Bacca, y J. L. Paredes, “Nonlinear filtering for image analysis and enhancement,” en *The Essential Guide to Image Processing*, 2da edición, A. C. Bovik, Editor. Academic Press, 2009, cap. 12, pp. 263–292.
- [31] T. Chen y H. R. Wu, “Adaptive impulse detection using center-weighted median filters,” *IEEE Signal Processing Letters*, vol. 8, no. 1, pp. 1–3, Ene. 2001.

- [32] J. L. Paredes y N. Pérez, “Frequency selective filtering using weighted order statistic admitting real-valued weights,” *Revista de la Facultad de Ingeniería, Universidad Central de Venezuela*, vol. 25, no. 4, pp. 95–105, Dic. 2010.