

**DISEÑO DE APERTURAS DE CODIFICACIÓN PARA LA  
ADQUISICIÓN COMPRESIVA DE IMÁGENES ESPECTRALES  
DINÁMICAS**

**KARETH MARCELA LEÓN LÓPEZ**

UNIVERSIDAD INDUSTRIAL DE SANTANDER  
FACULTAD DE INGENIERÍAS FISICOMECÁNICAS  
ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA  
MAESTRÍA EN INGENIERÍA DE SISTEMAS E INFORMÁTICA  
BUCARAMANGA  
2017

**DISEÑO DE APERTURAS DE CODIFICACIÓN PARA LA  
ADQUISICIÓN COMPRESIVA DE IMÁGENES ESPECTRALES  
DINÁMICAS**

**KARETH MARCELA LEÓN LÓPEZ**

Trabajo de investigación para optar por el título de:  
Magíster en Ingeniería de Sistemas e Informática

Director:

PhD. HENRY ARGUELLO FUENTES

Codirector:

PhD(c). LAURA VIVIANA GALVIS CARREÑO

UNIVERSIDAD INDUSTRIAL DE SANTANDER  
FACULTAD DE INGENIERÍAS FISICOMECAÑICAS  
ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA  
MAESTRÍA EN INGENIERÍA DE SISTEMAS E INFORMÁTICA  
BUCARAMANGA

2017

## Agradecimientos

Agradezco a Dios por permitirme alcanzar este logro, por darme las fuerzas suficientes para continuar y alcanzar mis metas.

Agradezco a mi madre, Maria Eugenia, por ser constante fuente de inspiración y apoyo incondicional.

Agradezco especialmente al profesor Henry Arguello Fuentes por su apoyo, enseñanzas y reflexiones en este camino de la academia. Agradezco a mis compañeros y amigos del grupo HDSP por demostrar su paciencia, disciplina y dedicación, la cual inspira a muchos para seguir adelante. Agradezco a Laura Galvis por ser guía y apoyo en mi formación académica.

Agradezco al grupo HDSP por brindar su laboratorio de óptica para realizar la adquisición de datos reales con los cuales se realizaron simulaciones en este trabajo.

## Dedicatoria

*A mi madre,  
A mis hermanos,  
A Dios.*

# CONTENIDO

	pág.
<b>Introducción</b>	<b>15</b>
<b>1 Muestreo compresivo (MC) e Imágenes multidimensionales</b>	<b>18</b>
1.1 MC en Imágenes espectrales . . . . .	18
1.2 Sistema de adquisición de imágenes espectral basado en MC (CASSI) . . . . .	19
1.3 Video espectral . . . . .	19
1.3.1 Imágenes espectrales dinámicas 4D . . . . .	19
1.3.2 Correlación temporal . . . . .	21
<b>2 Muestreo Compresivo en Video espectral</b>	<b>23</b>
2.1 Modelo de adquisición . . . . .	23
2.2 Aperturas de codificación . . . . .	24
2.2.1 Aperturas de codificación basadas en patrones binarios . . . . .	24
2.2.2 Aperturas de codificación de color . . . . .	25
2.3 Representación escasa en 4D: espacial (2D), espectral y temporal . . . . .	27
2.4 Enfoque de múltiples resoluciones en la reconstrucción de video compresivo . .	29
2.4.1 Estimación de baja resolución espacial . . . . .	30
2.4.2 Estimación del movimiento: Flujo óptico . . . . .	30
2.5 Reconstrucción de video a partir de las medidas comprimidas . . . . .	31
2.5.1 Reconstrucción de video compresivo en arquitecturas de multiplexación espacial, temporal y espectral . . . . .	31
2.5.1.1 Reconstrucción de video en un sistema de multiplexación es- pacial (CS-MUVI) . . . . .	31
2.5.1.2 Reconstrucción de video en un sistema de multiplexación tem- poral (P2C2) . . . . .	32
2.5.1.3 Reconstrucción de video en un sistema de multiplexación es- pacial (CASSI) . . . . .	32
<b>3 Diseño de aperturas de codificación y enfoque de reconstrucción con   múltiples resoluciones para la adquisición compresiva de video espectral</b>	<b>33</b>
3.1 Preliminares . . . . .	33
3.2 Criterios de diseño de las aperturas de codificación . . . . .	34
3.3 Algoritmo de diseño de aperturas de codificación . . . . .	35

3.3.1	Análisis espacio-temporal del diseño . . . . .	37
3.3.2	Análisis espectro-temporal del diseño . . . . .	39
3.3.3	Coherencia mutua . . . . .	40
3.4	Reconstrucción restringida en video compresivo espectral . . . . .	40
<b>4</b>	<b>Simulaciones y Resultados</b>	<b>41</b>
4.1	Parámetros de simulación . . . . .	41
4.2	Reconstrucción usando el modelo de reconstrucción tradicional . . . . .	43
4.3	Enfoque de reconstrucción en múltiples resoluciones . . . . .	44
<b>5</b>	<b>Conclusiones, Contribuciones y Trabajo futuro</b>	<b>50</b>
5.1	Conclusiones . . . . .	50
5.2	Contribuciones . . . . .	50
5.3	Trabajo futuro . . . . .	51
	<b>Bibliografía</b>	<b>52</b>
	Referencias . . . . .	52

# LISTA DE FIGURAS

pág.

1	Ilustración del sistema de adquisición de imágenes espectrales de única captura basado en muestreo compresivo, CASSI. En el sistema, la señal incidente $f_0(x, y, \lambda)$ es codificada espacialmente por la apertura de codificación $T(x, y)$ , y luego es separada espectralmente por un elemento dispersor, generalmente un prisma. La señal codificada y espectralmente separada, $f_2(x, y, \lambda)$ , es integrada en el detector $\mathbf{Y}_{j,\ell}$ . Tomado y modificado de (Arce, Brady, Carin, Arguello, y Kittle, 2014). . . . .	20
2	Esquema ilustrativo de un video espectral. Cada imagen espectral varía con respecto a las imágenes adyacentes en cada instante de tiempo ( $t$ ). . . . .	20
3	Video espectral de vigilancia. En (a) se muestran ocho fotogramas consecutivos (mapeados a RGB) en diferentes instantes de tiempo ( $t$ ). En (b) se muestra el fotograma 5 de (a) en ocho diferentes longitudes de onda. Tomado de (Benezeth, Sidibé, y Thomas, 2014). . . . .	21
4	Ilustración de la correlación que existe entre los fotogramas de un video espectral compuesto por 16 fotogramas y 8 bandas espectrales. La correlación entre fotogramas espectrales, aumenta (es máxima) entre fotogramas cercanos, y disminuye entre fotogramas alejados en el tiempo. . . . .	22
5	Patrones binarios usualmente usados en las arquitecturas de adquisición compresiva de imágenes espectrales. El patrón aleatorio es generado siguiendo una distribución Gaussiana o de Bernoulli, a diferencia de la matriz Hadamard, que sigue una estructuración definida. El círculo muestra una versión ampliada de $3 \times 3$ píxeles de cada apertura. . . . .	25
6	Ilustración de una apertura de codificación de color, CCA. Una CCA puede verse como una matriz donde cada elemento espacial $i, j$ es un filtro óptico con una respuesta espectral ideal dada (Derecha). Para un filtro pasa-baja $\mathcal{L}$ y pasa-alta $\mathcal{H}$ , se ilustra la modulación espectral de un píxel donde el color del filtro corresponde a la frecuencia de corte del filtro (Izquierda). Tomada y adaptado de (Rueda, Arguello, y Arce, 2014). . . . .	26
7	Reconstrucción a partir de la representación <i>escasa</i> de dos videos espectrales usando las diferentes bases de representación indicadas en la Tabla 1. Para cada video espectral, se toma solo el 1% de los coeficientes más significativos en la transformación y se obtiene una aproximación con una calidad de hasta 35 dB. En la parte superior, se muestra el fotograma 1 original para cada video, sin compresión, para las bandas espectrales 1, 3 y 8; en las siguientes filas se muestran las respectivas reconstrucciones usando el 1% de los coeficientes y su respectivo PSNR con respecto a la banda original. . . . .	28
8	Coeficientes en cada base de representación indicada en la Tabla 1, para dos videos espectrales con dimensiones espaciales $M = 128, N = 128, L = 8$ bandas espectrales y $D = 8$ fotogramas. . . . .	29

9	Ilustración de la representación de flujo óptico basado en un mapa de color. El movimiento entre el fotograma 1 y 2 puede ser representado en colores partiendo del flujo óptico usando el mapa de color. . . . .	30
10	Matriz del sistema de multiplexación espectral CASSI . . . . .	34
11	Distribución espacial de filtros ópticos $\Gamma$ (algoritmo 1, línea 8). La distribución espacial se genera ubicando los filtros de manera que la correlación espacio-espectral entre ellos sea la menor. Cada color en el arreglo muestra un índice de un filtro del conjunto $\xi$ . . . . .	36
12	Conjunto de filtros ópticos seleccionados $\xi$ . (Izquierda) Se muestra la respuesta espectral de cada filtro y de forma discreta (Derecha), en elementos de paso o bloqueo con su respectivo equivalente, donde $\mathcal{L}$ denota un filtro pasa-baja, $\mathcal{H}$ un filtro pasa-alta, $\mathcal{B}$ un filtro pasa-banda y $\mathcal{S}$ un filtro elimina-banda. . . . .	38
14	Realización de una apertura T-CCA. Los elementos blancos representan el paso y los elementos negros representan el bloqueo de la información (con $8 \times 8$ píxeles espaciales). (a) Distribución espacial de los elementos de paso. (b) Característica espectro-temporal de las T-CCA: con las T-CCA se asegura un muestreo uniforme a través del tiempo haciendo que un punto espacial sea muestreado de forma complementaria en el tiempo. . . . .	38
13	Comparación de la distribución espacial de una apertura generada con el algoritmo propuesto, T-CCA, y una apertura generada con una distribución aleatoria con un 25 % de elementos de paso y 75 % de elementos de bloqueo. En la sección resaltada se muestra la característica complementaria a través del tiempo para $D = 4$ . . . . .	39
15	Representación RGB de ocho fotogramas de los videos espectrales empleados para las simulaciones. La resolución de cada video es de $128 \times 128$ píxeles espaciales, 8 bandas espectrales y 8 fotogramas. . . . .	42
16	Representación en RGB de la reconstrucción del video 1 muestreado con las diferentes aperturas: $\mathcal{LH}$ -CCA (primera fila), $\mathcal{LH}$ -randomCCA (segunda fila), T-CCA (tercera fila) y la señal original (cuarta fila). En las columnas se tiene la secuencia de fotogramas que cambian en el tiempo. . . . .	43
17	Representación en RGB de la reconstrucción del fotograma 1 del video 2. La sección ampliada muestra la calidad de reconstrucción espacial alcanzada cuando se muestrea la señal con cada apertura de codificación. . . . .	44
18	Visualización de la reconstrucción de la tercera banda espectral en 8 fotogramas del video <b>1</b> . La calidad en términos de PSNR se presenta para cada reconstrucción. . . . .	45
19	Visualización de la reconstrucción de la tercera banda espectral en 8 fotogramas del video <b>2</b> . La calidad en términos de PSNR se presenta para cada reconstrucción. . . . .	45

20	Representación RGB del fotograma 5, original y reconstruido, del video 3. La sección enfocada muestra que hay espacialmente una mejora usando la apertura T-CCA, donde se obtiene hasta 4dB de ganancia en calidad de reconstrucción con respecto a las otras aperturas. . . . .	46
21	Comparación de la respuesta espectral (pixel espectral) para los puntos P1 (fotograma 1) y P2 (fotograma 8) del video espectral 2 usando las aperturas $\mathcal{LH}$ -CCA, $\mathcal{LH}$ -random CCA y T-CCA. Note que con las T-CCA se obtiene una mejora aproximación espectral (aproximación a la curva verde) con respecto a las otras aperturas. . . . .	47
22	Comparación de la respuesta espectral (píxel espectral) para los puntos P1 (fotograma 1) y P2 (fotograma 6) del video espectral 4 usando las aperturas $\mathcal{LH}$ -CCA, $\mathcal{LH}$ -random CCA y T-CCA. . . . .	48
23	Representación RGB de la reconstrucción de baja resolución de los primeros 4 fotogramas para los 4 videos de prueba con base en las medidas tomadas desde la apertura T-CCA. La calidad promedio en términos de PSNR se presenta para cada reconstrucción. . . . .	48

## LISTA DE TABLAS

	<b>pág.</b>
1 Combinaciones de diferentes bases de representación para el producto Kronecker. W: Transformada Wavelet; D: Transformada discreta coseno. Tomado y traducido de (C. Correa, Galvis, y Arguello, 2016). . . . .	27
2 Promedio de las métricas PSNR y SSIM para cada reconstrucción con base en el problema inverso tradicional . . . . .	44
3 Promedio de las métricas PSNR y SSIM para cada reconstrucción con base en el problema inverso propuesto . . . . .	49

## Resumen

**TÍTULO:** DISEÑO DE APERTURAS DE CODIFICACIÓN PARA LA ADQUISICIÓN COMPRESIVA DE IMÁGENES ESPECTRALES DINÁMICAS <sup>1</sup>

**AUTOR:** KARETH MARCELA LEÓN LÓPEZ <sup>2</sup>

**PALABRAS CLAVE:** Video espectral, diseño de aperturas de codificación, correlación temporal, muestreo compresivo.

Los sistemas de adquisición de vídeo espectral compresivo permiten obtener información espacial, espectral y temporal a través de la codificación y dispersión de los rayos incidentes usando una apertura de codificación y un elemento dispersor, respectivamente. Las aperturas codificadas de color (CCA) se han introducido en sistemas de adquisición compresiva de imágenes espectrales implicando estrategias de codificación más significativas. A pesar de que el uso de la CCA ofrece ventajas significativas y podría aplicarse también a la adquisición de video espectral compresivo, esta última aplicación presenta aún diversos desafíos originados por la variable temporal. Durante la reconstrucción de vídeo, algunos artefactos producidos por el movimiento de la escena obtienen *alias* dañando los datos. Como resultado, se han propuesto enfoques de reconstrucción basados en múltiples resoluciones para aliviar el *aliasing* y mejorar la reconstrucción de video. Sin embargo, estos enfoques se han enfocado en arquitecturas de multiplexación espacial o temporal mientras descartando la información espectral en el video. En este trabajo, se propone un enfoque de codificación y reconstrucción para un sistema de adquisición de video espectral compresivo. Para la codificación, se propone un algoritmo que diseña las aperturas de codificación explotando la matriz de muestreo usando la propiedad de coherencia. Los patrones diseñados permiten muestrear la información espacial, espectral y temporal de una manera uniforme tal que cada fotograma espectral se muestrea espacialmente al menos una vez. Para la reconstrucción, se propone un enfoque de reconstrucción basado en multiresolución en el sistema de multiplexación espectral CASSI (extendido al video) que permite extraer estimaciones de flujo óptico para abordar una reconstrucción de mayor calidad. Los experimentos computacionales realizados sobre cuatro videos espectrales diferentes muestran una mejora de hasta 4 dB en términos de relación pico-señal a ruido (PSNR) en la calidad de reconstrucción con el enfoque multiresolución utilizando los patrones diseñados con respecto a estructuras aleatorias.

---

<sup>1</sup>Trabajo de Investigación

<sup>2</sup>Facultad de Ingenierías Fisicomecánicas. Escuela de Ingeniería de Sistemas e Informática. Director, PhD. Henry Arguello Fuentes. Codirector, PhD(c) Laura Viviana Galvis Carreño.

## Abstract

**TITLE:** CODED APERTURE DESIGN FOR COMPRESSIVE SPECTRAL VIDEO SENSING <sup>3</sup>

**AUTHOR:** KARETH MARCELA LEÓN LÓPEZ <sup>4</sup>

**KEYWORDS:** Compressive spectral video, coded aperture design, temporal correlation, compressive sensing

Compressive spectral video sensing systems allow obtaining spatial, spectral and temporal information through the encoding and dispersion of the incoming light rays using a coded aperture and a dispersive element, respectively. Colored coded apertures (CCA) have been introduced in compressive spectral imaging (CSI) systems entailing richer coding strategies. Even though the use of the CCA offers significant advantages and could also be applied to compressive spectral video sensing, this later application still exhibits diverse challenges originated by the temporal variable. During the video reconstruction, some artifacts produced by the scene motion get aliased damaging the data. As a result, multiresolution-based reconstruction approaches have been proposed in order to alleviate the aliasing and enhance the video reconstruction. However, these approaches have been focused in spatial or temporal multiplexing architectures while the spectral information in the video is discarded. In this work, it is proposed an encoding and reconstruction approach for compressive spectral video sensing system. For the encoding, it is proposed an algorithm that design the coded apertures exploiting the sensing matrix by using the coherence property. The designed patterns allow to sense the spatial, spectral and temporal information in an uniform way such that each spectral frame is spatially sensed at least once. For the reconstruction, it is proposed a multiresolution-based reconstruction approach in the spectral multiplexing system CASSI (extended to video) allowing to extract optical flow estimates to address a higher quality reconstruction. Computational experiments performed over four different spectral videos show an improvement up to 4 dB in terms of peak-signal to noise ratio (PSNR) in the reconstruction quality with the multiresolution approach using the designed patterns with respect to the optimized and random CCA patterns.

---

<sup>3</sup>Research Work

<sup>4</sup>Faculty of Physical – Mechanics Engineering. Systems engineering and informatics School. Advisor, PhD, Henry Arguello Fuentes. Co-director PhD(c). Laura Viviana Galvis Carreño.

## Introducción

Las cámaras fotográficas convencionales intentan imitar el ojo humano mediante el uso de detectores sensibles al rojo, verde y azul (RGB). De manera que, para adquirir imágenes RGB, el sensor descarta un gran número de detalles de la escena respecto al espectro restante. Asimismo, se ha encontrado que con imágenes tomadas en diferentes longitudes de onda del espectro visible y del infrarrojo es posible adquirir muchos más detalles de la escena bajo observación. Los sensores para la adquisición de imágenes espectrales (SI del inglés *spectral imaging*) combinan la adquisición de imágenes bidimensionales (2D) y espectroscopia para detectar información espacial a través de un gran número de longitudes de onda. Una imagen espectral se puede ver como un cubo de datos 3D con dos dimensiones espaciales y una dimensión espectral (Smith y cols., 2001). Tradicionalmente, este tipo de imágenes son adquiridas por medio de espectrómetros que miden la intensidad lumínica o polarización electromagnética de las ondas de la luz presentes en un rango específico de frecuencias del espectro electromagnético. Debido a esto, el uso de estas imágenes permite la identificación de características no detectables a simple vista. Sin embargo, la adquisición de imágenes espectrales de alta resolución, utilizando técnicas tradicionales tales como *whiskbroom*, de escaneo punto por punto (Green y cols., 1998), *pushbroom*, de escaneo línea por línea (Gehm, Kim, Fernandez, y Brady, 2008), o cámara de filtros, de escaneo de línea por línea espectral; implican el almacenamiento de una gran cantidad de datos. Por ende, el costo de almacenamiento, procesamiento y transmisión de estos datos es elevado. En contraste con lo anterior, con el uso de la técnica de muestreo compresivo (MC) en los sistemas de adquisición de imágenes espectrales, es posible mitigar los costos de adquisición, almacenamiento y transferencia de datos, debido a que el muestreo de la información se realiza de forma comprimida (Arce y cols., 2014).

En general, los sistemas de adquisición basados en MC combinan el muestreo y la compresión a partir la proyección de la señal en el sistema, y con su uso es posible reducir significativamente el número de medidas requeridas para obtener una señal. Entre las arquitecturas de adquisición compresiva más relevantes está el sistema de adquisición de imágenes espectrales basado en una apertura de codificación de única adquisición (CASSI). CASSI permite adquirir la información espectral 3D de la escena en una sola toma. De forma general, este sistema se basa en la codificación y dispersión de la luz incidente al sensor de la cámara. El proceso de codificación se realiza utilizando un modulador espacial de luz (SLM, *Spatial Light Modulator* en inglés) o un dispositivo digital de micro espejos (DMD, *Digital Micromirror Device* en inglés) que permite la adquisición de mediciones moduladas de forma aleatoria. El proceso de dispersión es realizado por un prisma, el cual permite la separación espectral en diferentes longitudes de onda del espectro. Matemáticamente, el proceso de adquisición compresiva de una imagen espectral puede ser modelado como la proyección vectorial de la señal espectral  $\mathbf{f}$  en la matriz del sistema  $\Phi$ , expresado como  $\mathbf{g} = \Phi\mathbf{f}$ , donde la matriz del sistema representa los efectos de codificación y dispersión; y  $\mathbf{g}$  representa las mediciones en el detector.

Las imágenes espectrales han sido ampliamente utilizadas en campos como medicina, teledetección (Smith y cols., 2001) y aplicaciones militares (C. Correa y cols., 2016). Por otra parte, los cambios entre períodos cortos de tiempo de la información espectral es también importante en muchas otras aplicaciones como vigilancia, estudios biológicos microscópicos, entre otros. Por lo tanto, no sólo la información espacial y espectral, sino también la temporal es hoy en día de gran interés.

A la fecha, un número limitado de trabajos se ha centrado en el desarrollo de sistemas de adquisición de video espectral basados en el muestreo compresivo. Estos enfoques son apropiados para la adquisición espectral de video debido a su alta velocidad de adquisición y, a que la captura de dicha gran cantidad de datos se realiza de forma comprimida (A. A. Wagadarikar, Pitsianis, Sun, y Brady, 2009). Entre estos trabajos se encuentran los sistemas de adquisición híbridos, que se basan en la combinación de sensores de adquisición convencional y sensores basados en muestreo compresivo como CASSI, y el uso de algoritmos robustos de reconstrucción de video desde un único conjunto de medidas codificadas. Entre las arquitecturas de este tipo está el sistema de adquisición híbrida de video espectral (HVIS del inglés *hybrid spectral video imaging system*) (Ma, Cao, Wu, y Dai, 2014; Ma, Cao, Tong, Dai, y Lin, 2014); el sistema de adquisición compresiva basada en aperturas de codificación para imágenes espectro-temporales (Tsai, Llull, Yuan, Carin, y Brady, 2015); y el sistema de adquisición de video espectral de alta velocidad (HSHS del inglés *high-speed hyperspectral video acquisition*) (L. Wang y cols., 2015).

Los sistemas de adquisición compresiva de video espectral mencionados anteriormente, usan aperturas de codificación con “independencia espectral”, esto es, que todas las componentes espectrales son codificadas de la misma manera. Dichas aperturas son conocidas como aperturas binarias, o aperturas de paso y bloqueo, las cuales pueden ser generadas con alguna distribución Gausiana o Bernoulli. Estas aperturas, como su nombre lo indica, contienen elementos de transmisión, o de paso, y elementos oclusores, o de bloqueo, distribuidos espacialmente que modulan espacialmente la luz incidente; esta propiedad de codificación implica que no existe una codificación a lo largo de la componente espectral.

Por otra parte, se han desarrollado trabajos que proponen el diseño de aperturas de codificación para la adquisición de imágenes espectrales “estáticas” para explotar la correlación que existe en la componente espectral de las imágenes espectrales. Entre dichos trabajos está el diseño para selectividad espectral (Arguello y Arce, 2013), filtrado espectral (Arguello y Arce, 2014) y el diseño basado en ruido azul (C. V. Correa, Arguello, y Arce, 2016). Sin embargo, hasta la fecha, en la literatura no se han presentado diseños específicos para la adquisición compresiva de video espectral, en donde además de tener redundancia espectral, se tiene redundancia y alta correlación en la componente temporal.

Por otro lado, en cuanto a la reconstrucción de video compresivo, existen muchos retos originados por la variable temporal, entre ellos los artefactos que aparecen en el proceso de reconstrucción. El movimiento de la escena dinámica durante la adquisición produce artefactos inherentes de movimiento, los cuales provocan un solapamiento o *aliasing* durante la reconstrucción del video, dañando la totalidad de los datos (Sankaranarayanan y cols., 2015). Por esta razón, en trabajos como (Sankaranarayanan y cols., 2015; Reddy, Veeraraghavan, y Chellappa, 2011) se propone un enfoque de múltiples resoluciones que permite reducir el solapamiento y mejorar la reconstrucción de video. La idea de interpretar los datos en múltiples resoluciones se conoce como el problema “del huevo y la gallina” (*“chicken and egg problem”*). Este problema establece que es posible obtener un video adquirido de forma compresiva, en alta calidad, agregando la correlación temporal como el flujo óptico; sin embargo, el cálculo de dicha correlación temporal requiere del video completo reconstruido. Una solución que se encontró a este problema fue realizar una reconstrucción tosca, o de baja resolución del video, para así extraer la información temporal con la que se mejora la calidad del video en alta re-

solución (Sankaranarayanan y cols., 2015). No obstante, los trabajos del estado del arte ahora que implementan esta solución, solo han sido enfocados a sistemas que realizan la compresión del video en forma espacial, como CS-MUVI (Sankaranarayanan y cols., 2015), y en forma temporal, como P2C2 (Reddy y cols., 2011; Goldstein, Xu, Kelly, y Baraniuk, 2015), y donde la componente espectral del video no se adquiere, ni se tiene en cuenta.

Con el fin de mejorar, tanto la adquisición como la reconstrucción de video espectral adquirido de forma compresiva, este trabajo propone un diseño de las aperturas de codificación en un sistema de adquisición compresiva de video espectral, que tiene en cuenta la redundancia de información espacio-espectro-temporal de la escena. Adicionalmente, plantea hacer uso del enfoque de múltiples resoluciones orientado a la reconstrucción de video espectral, donde es posible tener en cuenta la correlación temporal de la escena para mejorar la calidad de reconstrucción en términos de relación señal a pico de ruido (PSNR). Los resultados en simulaciones, con base en el diseño de aperturas de codificación propuesto, presentan una calidad de reconstrucción de hasta 33dB en términos de PSNR, y de hasta 36dB cuando se agrega la correlación temporal a la reconstrucción. De manera general, en comparación con las aperturas de codificación del estado del arte, se obtiene una ganancia en promedio de hasta 4dB en PSNR.

# 1. Muestreo compresivo (MC) e Imágenes multidimensionales

Los avances en el campo de adquisición y procesamiento de señales han producido diversos aportes a la industria tales como sensores ópticos, sensores de radio y detectores sísmicos. Estos avances se basan en el criterio de muestreo de Shannon-Nyquist, el cual establece que es posible reconstruir una señal si esta es limitada en su ancho de banda y su tasa de muestreo es superior al doble de su frecuencia máxima. No obstante, con la teoría de muestreo compresivo (MC), se ha comprobado que el límite de la tasa de muestreo impuesto por el criterio de Shannon-Nyquist puede ser altamente reducido (Qaisar, Bilal, Iqbal, Naureen, y Lee, 2013). Para esto, es necesario que la señal sea “*escasa*”, es decir, que la mayor parte de la energía de la señal esté concentrada en un conjunto pequeño de sus componentes. Si bien, la mayoría de señales naturales no son *escasas*, es posible encontrar una representación de este tipo en alguna base de representación o base de transformación dada (Candès y Wakin, 2008; Qaisar y cols., 2013; Candès, 2006).

MC ha sido ampliamente aplicado en el procesamiento de imágenes y video, específicamente, en imágenes espectrales, campos de luz, tomografía computarizada, biología computacional, astronomía y súper resolución (Thapa, Raahemifar, y Lakshminarayanan, 2015; Marwah, Wetzstein, Bando, y Raskar, 2013). El éxito de esta técnica radica en que los procesos de adquisición y compresión son realizados simultáneamente, disminuyendo considerablemente el número de muestras requeridas para reconstruir la señal deseada.

## 1.1. MC en Imágenes espectrales

Una imagen espectral se define como un conjunto de imágenes bidimensionales en  $(x, y)$ , adquiridas en diferentes longitudes de onda ( $\lambda$ ). Matemáticamente y en forma discreta, una imagen espectral dada, con  $N \times N$  píxeles espaciales y  $L$  bandas espectrales es definida como  $\mathbf{F} \in \mathbb{R}^{N \times N \times L}$ , o su representación vectorizada  $\mathbf{f} \in \mathbb{R}^n$  siendo  $n = N^2L$ ; de la imagen puede ser expresada de forma *escasa* como

$$\mathbf{f} = \mathbf{\Psi}\boldsymbol{\theta}, \quad (1.1)$$

donde  $\mathbf{\Psi} \in \mathbb{R}^{n \times n}$  es una base de transformación (Wavelet, DCT, entre otras) y  $\boldsymbol{\theta}$  el vector que representa la señal en el dominio de transformación con  $K$  coeficientes diferentes de cero ( $K \ll n$ ).

El proceso simultáneo de muestreo y compresión se puede expresar matemáticamente como la proyección de la señal  $\mathbf{f}$ , o su representación  $\mathbf{\Psi}\boldsymbol{\theta}$ , en la matrix de muestreo  $\mathbf{\Phi} \in \mathbb{R}^{m \times n}$ . Dichas proyecciones representan las medidas que se adquieren en el detector y se expresan como

$$\mathbf{g} = \mathbf{\Phi}\mathbf{f} = \mathbf{\Phi}\mathbf{\Psi}\boldsymbol{\theta} = \mathbf{A}\boldsymbol{\theta}, \quad (1.2)$$

donde  $\mathbf{g}$  representa el detector o medidas comprimidas y  $\mathbf{A} = \mathbf{\Phi}\mathbf{\Psi}$  es la matrix de muestreo. La señal comprimida puede ser recuperada resolviendo el problema inverso dado por  $\boldsymbol{\theta} = \mathbf{A}^{-1}\mathbf{g}$ , sin embargo, debido a que el número de filas de la matrix  $\mathbf{A}$  es menor que el número de columnas ( $m \ll n$ ) el sistema resultante tiene infinitas soluciones.

Con base en la teoría de MC, una aproximación de la señal se puede encontrar resolviendo un problema de minimización que busca la solución más escasa de  $\mathbf{g}$  en  $\Psi$ . El planteamiento del problema inverso para la reconstrucción de la señal  $\mathbf{f}$  puede escribirse como

$$\hat{\mathbf{f}} = \Psi(\underset{\boldsymbol{\theta}}{\operatorname{argmin}} \|\mathbf{g} - \Phi\Psi\boldsymbol{\theta}\|_2^2 + \tau\|\boldsymbol{\theta}\|_1), \quad (1.3)$$

donde  $\tau$  es un parámetro de regularización,  $\|\cdot\|_1$  es la norma  $\ell_1$  que mide la cantidad de valores diferentes de cero en el vector, y  $\|\cdot\|_2$  es la norma  $\ell_2$  usada para medir el error cuadrático medio de la estimación. La minimización de estas dos normas implica la convergencia de la función, y por lo tanto, la reconstrucción de la señal con base en las medidas.

## 1.2. Sistema de adquisición de imágenes espectral basado en MC (CASSI)

El muestreo compresivo en la adquisición de imágenes espectrales ha tenido gran acogida por la comunidad científica dado que permite disminuir costos en sensores espectrales, tiempos de adquisición y almacenamiento. Hasta la fecha, se han desarrollado diferentes arquitecturas que permiten la adquisición de imágenes espectrales en forma comprimida. Entre las arquitecturas más conocidas se encuentra el sistema de adquisición de imágenes espectrales basado en una apertura de codificación de única captura, CASSI (*Coded Aperture Snapshot Spectral Imager*) (A. Wagadarikar, John, Willett, y Brady, 2008). Este sistema permite obtener proyecciones codificadas de una imagen espectral 3D usando un detector 2D. Los componentes principales del sistema se muestran en la Figura 1.

Desde la invención del sistema CASSI, se han propuesto diferentes variaciones para mejorar la calidad de reconstrucción de las imágenes obtenidas. Entre dichas modificaciones se encuentran: la extensión de CASSI para admitir múltiples capturas con diferentes patrones de aperturas de codificación (Arguello y Arce, 2011a; Wu, Mirza, Arce, y Prather, 2011), el diseño óptimo en las aperturas de codificación para la selectividad espectral (Arguello y Arce, 2013) y el diseño de filtros ópticos para una adquisición óptima de la información espectral (Rueda, Arguello, y Arce, 2015; Arguello y Arce, 2014). Entre otros sistemas de adquisición compresiva, se encuentran SCCSI (*Snapshot Colored Compressive Spectral Imager*) (C. V. Correa, Arguello, y Arce, 2015) y SSCSI (*Spatial-spectral Encoded Compressive Hyperspectral Imaging*) (Lin, Liu, Wu, y Dai, 2014), los cuales varían con respecto a CASSI en el posicionamiento de la apertura de codificación en el camino óptico.

## 1.3. Video espectral

### 1.3.1. Imágenes espectrales dinámicas 4D

En aplicaciones como vigilancia, seguimiento de células vivas, seguridad o reconocimiento de objetivos, las imágenes bajo análisis no son completamente estáticas, dado a que muchos cambios ocurren en periodos cortos de tiempo (Leitner y cols., 2013; Arnold, De Biasio, y Leitner, 2013; Cheng, Park, y Trivedi, 2007; C. Correa y cols., 2016; Ramirez, Arguello, y Arce, 2011; Banerjee, Burlina, y Broadwater, 2009). De manera que, además de la información espacial y espectral, la información a lo largo del tiempo en estas aplicaciones toma importancia. La

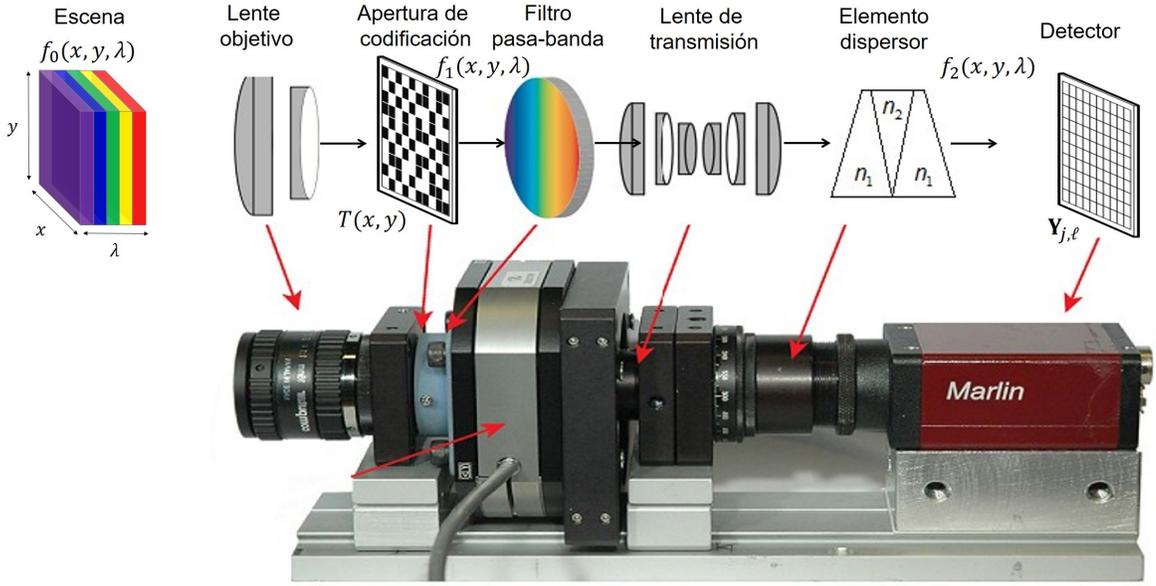


Figura 1: Ilustración del sistema de adquisición de imágenes espectrales de única captura basado en muestreo compresivo, CASSI. En el sistema, la señal incidente  $f_0(x, y, \lambda)$  es codificada espacialmente por la apertura de codificación  $T(x, y)$ , y luego es separada espectralmente por un elemento dispersor, generalmente un prisma. La señal codificada y espectralmente separada,  $f_2(x, y, \lambda)$ , es integrada en el detector  $\mathbf{Y}_{j,\ell}$ . Tomado y modificado de (Arce y cols., 2014).

adquisición de la información espacial, espectral y temporal de una escena es conocida como video espectral. Un video espectral puede ser modelado como una secuencia de imágenes espectrales, con dimensiones  $(x, y, \lambda)$ , indexadas en el tiempo ( $t$ ) como se ilustra en la Fig. 2.

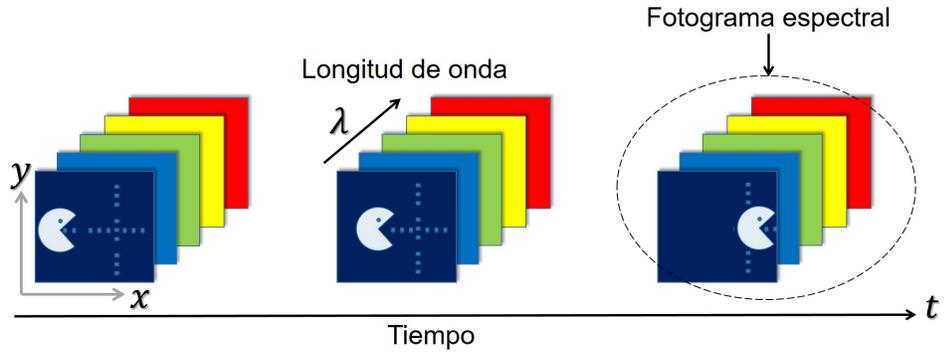


Figura 2: Esquema ilustrativo de un video espectral. Cada imagen espectral varía con respecto a las imágenes adyacentes en cada instante de tiempo ( $t$ ).

En forma discreta, un video espectral se puede expresar como el conjunto de  $D$  imágenes o fotogramas espectrales  $\mathbf{F}_k \in \mathbb{R}^{M \times N \times L}$ , con  $M \times N$  píxeles espaciales y  $L$  bandas espectrales, cada una en capturada en un tiempo  $k$  ( $k = 0, \dots, D - 1$ ). En forma compacta, un video

espectral se puede expresar como la concatenación de cada fotograma espectral en su forma vectorial,  $\mathbf{f}_t = \text{vec}(\mathbf{F}_k)$ , dado por la expresión

$$\mathbf{f}_{1:D} = [\mathbf{f}_1, \dots, \mathbf{f}_D] \quad (1.4)$$

donde  $D$  es el número de fotogramas espectrales capturados en un período (Sankaranarayanan, Turaga, Chellappa, y Baraniuk, 2013) y  $\text{vec}(\cdot)$  es un operador que vectoriza una matriz dada. La Figura 3 presenta un ejemplo de un video espectral real de vigilancia; en la parte (a) se muestra los cambios que hay de fotograma a fotograma, en una versión mapeada a RGB, y en la parte (b) se muestra el fotograma 5 en diferentes longitudes de onda (Benzeth y cols., 2014).

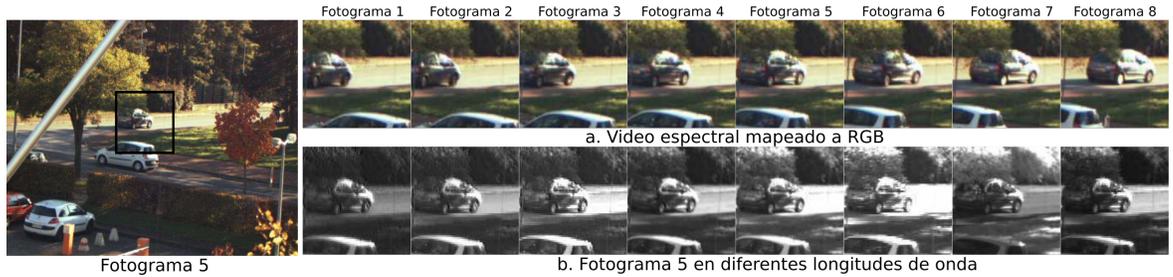


Figura 3: Video espectral de vigilancia. En (a) se muestran ocho fotogramas consecutivos (mapeados a RGB) en diferentes instantes de tiempo ( $t$ ). En (b) se muestra el fotograma 5 de (a) en ocho diferentes longitudes de onda. Tomado de (Benzeth y cols., 2014).

### 1.3.2. Correlación temporal

Entre las características más importantes del video espectral está la alta redundancia de información, ya que se adquiere información espacio-espectral de la misma escena en diferentes instantes de tiempo. Esta redundancia implica que existe una alta correlación entre la información. Para ilustrar este concepto, la Fig. 4 muestra una matriz de correlación en función de cada fotograma espectral, donde la mayor correlación equivale al valor numérico 1, para un video espectral de 16 fotogramas y 8 bandas espectrales. Cabe notar, que cuando se compara contra el mismo fotograma (diagonal de la matriz) el valor correspondiente es el máximo. Sin embargo, cuando se compara con fotogramas diferentes a este mismo, los valores son muy cercanos a 1, indicando que existe correlación (para más detalle ver (Manolakis, Lockwood, y Cooley, 2008), Ecuación 8.).

Con base en esta correlación temporal, es posible explotar la redundancia de información a través de sistemas de compresión, de tal forma que sea posible comprimir altamente la señal de video sin llegar a dañar los datos; lo que permitiría reducir costos en almacenamiento y procesamiento.

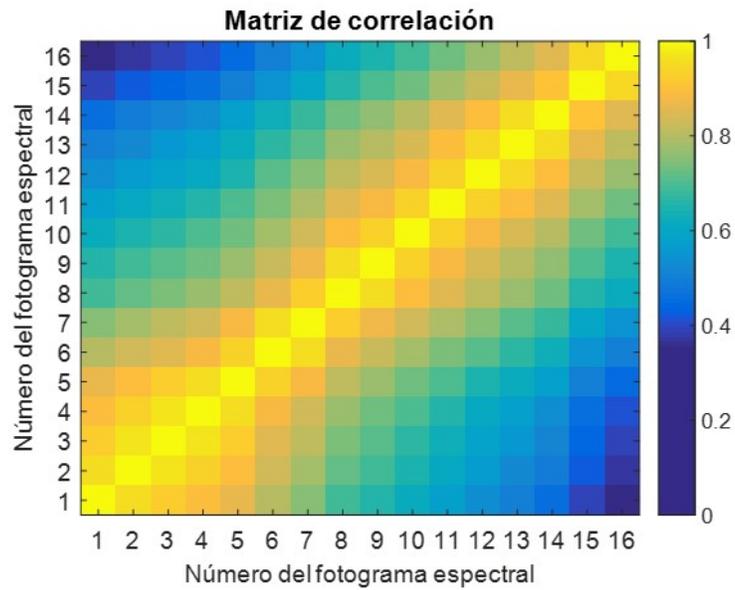


Figura 4: Ilustración de la correlación que existe entre los fotogramas de un video espectral compuesto por 16 fotogramas y 8 bandas espectrales. La correlación entre fotogramas espectrales, aumenta (es máxima) entre fotogramas cercanos, y disminuye entre fotogramas alejados en el tiempo.

## 2. Muestreo Compresivo en Video espectral

### 2.1. Modelo de adquisición

El modelo de adquisición compresiva de video espectral en la arquitectura de adquisición CASSI se asemeja al modelo de imágenes estáticas expuesto en la Ecuación 1.2, no obstante, la señal bajo observación será el video espectral expresado en la Ecuación 1.4. Por tanto, la adquisición compresiva de la escena espectral para un tiempo ( $d$ ) se expresa de la forma

$$\mathbf{g}_d = \mathbf{\Phi}_d \mathbf{f}_d, \quad (2.1)$$

donde  $\mathbf{g}_d$  representa las medidas comprimidas en el detector bidimensional,  $\mathbf{\Phi}_d$  es la matriz de muestreo y  $\mathbf{f}_d$  es el fotograma espectral, para  $d = 0, \dots, D-1$ . De manera general, para un video espectral  $\mathbf{f} \in \mathbb{R}^{\dot{n}}$ , con  $\dot{n} = NNLD$ , las medidas en el detector se pueden reescribir como

$$\underbrace{\begin{bmatrix} \mathbf{g}_0 \\ \mathbf{g}_1 \\ \vdots \\ \mathbf{g}_{D-1} \end{bmatrix}}_{\mathbf{g}} = \underbrace{\begin{bmatrix} \mathbf{\Phi}_0 & \mathbf{0}_{DN(1) \times NNLD} & \cdots & \mathbf{0}_{DN(L-1) \times NNLD} \\ \vdots & \mathbf{\Phi}_1 & \ddots & \vdots \\ \mathbf{0}_{DN(L-1) \times NNLD} & \mathbf{0}_{DN(L-2) \times NNLD} & \cdots & \mathbf{\Phi}_{D-1} \end{bmatrix}}_{\mathbf{\Phi}} \underbrace{\begin{bmatrix} \mathbf{f}_0 \\ \mathbf{f}_1 \\ \vdots \\ \mathbf{f}_{D-1} \end{bmatrix}}_{\mathbf{f}}, \quad (2.2)$$

donde  $N$  representa el número de píxeles espaciales,  $L$  las bandas espectrales y  $D$  el número de fotogramas. La matriz total de muestreo del sistema  $\mathbf{\Phi}$  tendrá dimensiones  $\dot{m} \times \dot{n}$ , con  $\dot{m} = DN(N+L-1)$ . Los elementos  $\mathbf{0}_{DN(1) \times NNLD}$  y  $\mathbf{0}_{DN(L-1) \times NNLD}$  son matrices de tamaño  $DN(1) \times NNLD$  y  $DN(L-1) \times NNLD$ , cuyos elementos son cero. La matriz  $\mathbf{\Phi}$  es la concatenación de  $D$  submatrices  $\mathbf{\Phi}_d$ , donde cada submatriz  $\mathbf{\Phi}_d$  contiene en su estructura el patrón de la apertura con el que se codifica cada fotograma del video espectral. De forma más detallada, la matriz  $\mathbf{\Phi}$  modela el efecto de dispersión ( $\mathbf{P} \in \mathbb{R}^{\dot{m} \times \dot{n}}$ ) y codificación ( $\mathbf{T} \in \mathbb{R}^{\dot{n} \times \dot{n}}$ ) en el sistema de adquisición; de manera que,  $\mathbf{\Phi}$  puede descomponerse como

$$\mathbf{\Phi} = \mathbf{P}\mathbf{T}, \quad (2.3)$$

donde  $\mathbf{P}$  es la matriz dispersión dada por  $\mathbf{P} = \text{blkdiag}\{(\tilde{\mathbf{P}}^0)^T, \dots, (\tilde{\mathbf{P}}^{D-1})^T\}^T$  donde el operador  $\text{blkdiag}\{\cdot\}$  denota una concatenación de bloques diagonales, y cada submatriz  $\tilde{\mathbf{P}}^d$ , expresada como

$$\tilde{\mathbf{P}}^d = \begin{bmatrix} \text{diag}(\mathbf{1}_{N^2}) & \mathbf{0}_{N \times N^2} & \cdots & \mathbf{0}_{N \times N^2} \\ \mathbf{0}_{N \times N^2} & \text{diag}(\mathbf{1}_{N^2}) & \cdots & \mathbf{0}_{N \times N^2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{N \times N^2} & \mathbf{0}_{N \times N^2} & \cdots & \text{diag}(\mathbf{1}_{N^2}) \end{bmatrix}_{(N+L-1)N \times N^2L}, \quad (2.4)$$

representa la dispersión para cada fotograma espectral  $d$ , donde  $\text{diag}(\mathbf{1}_{N^2})$  es una matriz diagonal de unos con dimensiones  $N^2 \times N^2$ . La matriz  $\mathbf{T}$ , es una matriz dada por  $\mathbf{T} = \text{blkdiag}\{(\tilde{\mathbf{T}}^0)^T, \dots, (\tilde{\mathbf{T}}^{D-1})^T\}^T$  donde  $\tilde{\mathbf{T}}^d$  se expresa como

$$\tilde{\mathbf{T}}^d = \begin{bmatrix} \text{diag}(\mathbf{t}_0^d) & \mathbf{0}_{N^2 \times N^2} & \cdots & \mathbf{0}_{N^2 \times N^2} \\ & \text{diag}(\mathbf{t}_1^d) & \cdots & \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{N^2 \times N^2} & \mathbf{0}_{N^2 \times N^2} & \cdots & \text{diag}(\mathbf{t}_{L-1}^d) \end{bmatrix}_{N^2L \times N^2L}, \quad (2.5)$$

siendo  $\text{diag}(\mathbf{t}_k^d)$  una matriz diagonal cuyas entradas son los valores de la apertura de codificación para  $k = 0, \dots, L - 1$ . Tradicionalmente, las entradas de la apertura de codificación son generadas siguiendo una distribución aleatoria Gaussiana o Bernoulli y satisfacen  $\mathbf{t}_0^d = \mathbf{t}_1^d = \dots = \mathbf{t}_{L-1}^d$ . Esta clase de aperturas de codificación es conocida como aperturas binarias o aperturas de bloqueo-desbloqueo. Sin embargo, recientemente y como parte de esta investigación, se propuso un trabajo, en el área de video compresivo espectral, donde dichas aperturas son cambiadas por aperturas de codificación espectro-dependientes, lo que significa que además de codificar la información espacio-temporal, se codifica la información espectral (León, Galvis, y Arguello, 2016) y satisfacen que  $\mathbf{t}_0^d \neq \mathbf{t}_1^d \neq \dots \neq \mathbf{t}_{L-1}^d$  (Arguello y Arce, 2014; Rueda y cols., 2015). Estas aperturas proveen una propiedad que es conocida como filtrado, lo que produce un procedimiento de codificación espectralmente más significativo (Rueda y cols., 2015). En la siguiente sección se ampliará la idea de este tipo de aperturas de codificación y, en el siguiente capítulo, se presentará el diseño que se propone para mejorar la calidad de reconstrucción en video espectral.

El uso de aperturas de codificación binarias para la adquisición compresiva de video espectral, y en general para la adquisición compresiva de imágenes espectrales, ha demostrado un buen desempeño en términos de calidad de reconstrucción de las imágenes. Sin embargo, en diferentes trabajos se ha presentado una mejora significativa en la calidad de reconstrucción cuando se diseña el patrón de codificación. Trabajos como (Arguello y Arce, 2014) y (Arguello y Arce, 2011a) demuestran una mejora en calidad de reconstrucción de hasta 23 decibelios en términos de relación señal a pico de ruido (PSNR) para la adquisición de imágenes espectrales estáticas. No obstante, se encuentran muy pocos trabajos directamente dedicados al diseño de las aperturas de codificación para la adquisición compresiva de video espectral, y el uso de aperturas diseñadas como las mencionadas anteriormente para imágenes estáticas en la adquisición de video espectral, implica un error significativo en el proceso de reconstrucción dado que la imagen cambia en el tiempo. Por otro lado y con respecto a la naturaleza de movimiento de las imágenes dinámicas o video, en un trabajo reciente se demostró que teniendo en cuenta la correlación temporal entre los fotogramas de un video se reducen errores y se eliminan artefactos en la reconstrucción, con lo cual se aumenta la calidad de reconstrucción del video (Sankaranarayanan y cols., 2015). No obstante, el concepto de correlación temporal aún no ha sido previamente utilizado, ni aprovechado en la adquisición de video compresivo espectral.

## 2.2. Aperturas de codificación

Las aperturas de codificación juegan un papel fundamental en la adquisición de datos basado en muestreo compresivo. Este elemento óptico, ilustrado en la Fig. 1 del capítulo 1, contiene una distribución de elementos que codifican la escena bajo observación.

### 2.2.1. Aperturas de codificación basadas en patrones binarios

Tradicionalmente, en la adquisición compresiva de video espectral, se emplean aperturas de codificación binarias. Estas aperturas permiten o bloquean el paso de luz que atraviesa el camino óptico hacia el detector. Matemáticamente, una apertura de codificación binaria puede

representarse como un arreglo matricial  $T_{i,j} \in \{0, 1\}$ , donde los elementos con valor igual a 1 funcionan como elementos transmisivos, y los elementos con valor 0 como ocluidores. Entre las distintas estructuras de las aperturas de codificación binarias están los patrones aleatorios y las estructuras Hadamard.

Las aperturas de codificación aleatorias son producidas con alguna distribución Gaussiana o de Bernoulli. Por otro lado, los patrones estructurados como Hadamard son usados por su eficiencia para reducir la varianza del ruido presente en las medidas (Willett, Gehm, y Brady, 2007; Sloane y Harwit, 1976). En la Fig. 5 se muestra un ejemplo de las aperturas anteriormente descritas, con su respectiva versión ampliada de  $3 \times 3$  píxeles.

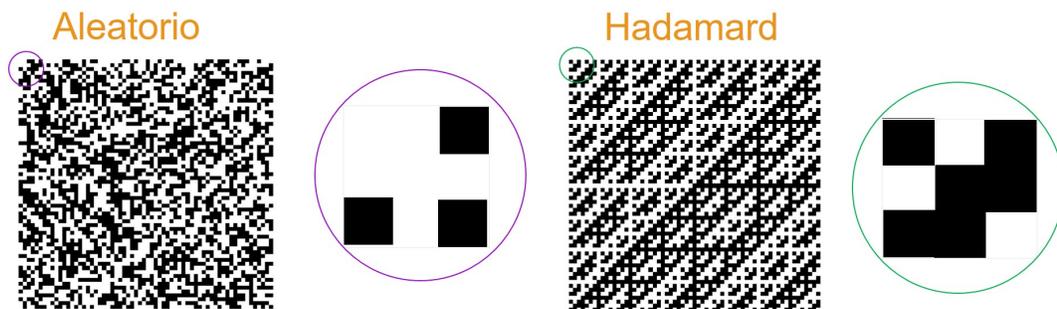


Figura 5: Patrones binarios usualmente usados en las arquitecturas de adquisición compresiva de imágenes espectrales. El patrón aleatorio es generado siguiendo una distribución Gaussiana o de Bernoulli, a diferencia de la matriz Hadamard, que sigue una estructuración definida. El círculo muestra una versión ampliada de  $3 \times 3$  píxeles de cada apertura.

El uso de patrones binarios distribuidos aleatoriamente para la adquisición compresiva de video espectral, y en general para la adquisición compresiva de imágenes espectrales, ha demostrado un buen desempeño con respecto a la calidad de reconstrucción de las imágenes. Sin embargo, en diferentes trabajos se ha presentado una mejora significativa en la calidad de reconstrucción cuando se diseña el patrón de codificación (Arguello y Arce, 2014, 2011a).

### 2.2.2. Aperturas de codificación de color

Las aperturas de codificación de color (CCA del inglés *Colored coded apertures*) se desarrollaron recientemente para la adquisición y compresión de imágenes “estáticas” espectrales usando una o múltiples capturas. Estas aperturas han demostrado una notable mejora en la calidad de reconstrucción con respecto a las aperturas de codificación binarias, en escenarios tanto simulados como reales (Arguello y Arce, 2014; Rueda y cols., 2015). La principal característica de las CCA está en que no solo se realiza un bloqueo o paso de la información espacialmente, sino que también se da este bloqueo o paso en distintas regiones del espectro electromagnético. Esta característica se interpreta como la modulación de la luz a través de un arreglo de filtros ópticos donde cada filtro modula la intensidad de luz incidente de acuerdo a su frecuencia de corte, es decir, cuales longitudes de onda se dejan pasar. Para ilustrar esta característica, la Fig. 6 muestra un ejemplo de una apertura de codificación de color con las correspondientes iniciales del filtro óptico al que corresponde cada píxel de la apertura:

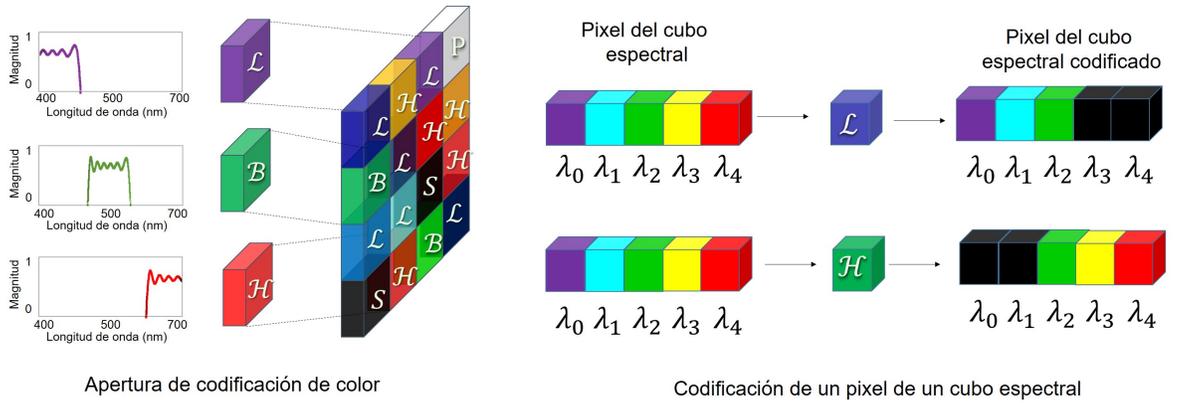


Figura 6: Ilustración de una apertura de codificación de color, CCA. Una CCA puede verse como una matriz donde cada elemento espacial  $i, j$  es un filtro óptico con una respuesta espectral ideal dada (Derecha). Para un filtro pasa-baja  $\mathcal{L}$  y pasa-alta  $\mathcal{H}$ , se ilustra la modulación espectral de un píxel donde el color del filtro corresponde a la frecuencia de corte del filtro (Izquierda). Tomada y adaptado de (Rueda y cols., 2014).

$\mathcal{L}$  filtro pasa-baja,  $\mathcal{H}$  filtro pasa-alta,  $\mathcal{B}$  filtro pasa-banda,  $\mathcal{S}$  pasa-nada y  $\mathcal{P}$  pasa-todo. En la parte derecha de la figura, se ejemplifica la modulación para dos píxeles, usando un filtro pasa-bajas y uno pasa-altas. Cuando el píxel es codificado, cada punto donde el filtro no deja pasar se torna en color negro, lo que equivale a que se omita esa parte de la información como se muestra en la parte izquierda de la figura. Además, note que el color del filtro representa la frecuencia de corte del filtro.

Las CCA son diseñadas a partir de las frecuencias de corte de los filtros ópticos. Este diseño se basa en la minimización de la correlación entre las respuestas espectrales de un ensamble de filtros ópticos. Asimismo, cuando hay múltiples capturas, los píxeles de las CCA se hacen complementarios entre sí implicando que algunos píxeles sean bloqueados completamente. Con esto, el número de muestras tomadas es minimizado mientras la calidad de reconstrucción es maximizada (Arguello y Arce, 2014; Rueda y cols., 2015).

### 2.3. Representación escasa en 4D: espacial (2D), espectral y temporal

La representación en otro dominio de la señal submuestreada es un punto clave en la correcta recuperación de la señal. En video espectral, es importante aprovechar la correlación existente entre las cuatro dimensiones: espacial (2D), espectral y temporal, por lo que, se es necesario una base de representación que aproveche la redundancia de información. Recientemente, en (C. Correa y cols., 2016) se propuso una base de representación apropiada para la representación de video espectral. Allí, se propone hacer el producto Kronecker entre diferentes transformadas discretas para formar así una base 4D. Matemáticamente, la representación escasa de  $\mathbf{f}$  se escribe como

$$\mathbf{f} = \{\Psi_1 \otimes \Psi_2 \otimes \Psi_3 \otimes \Psi_4\}\boldsymbol{\theta} = \Psi_{4D}\boldsymbol{\theta}, \quad (2.6)$$

donde  $\{\Psi_i\}_{i=1}^4$  es un conjunto de diferentes transformadas unidimensionales (C. Correa y cols., 2016). Usualmente, para imágenes espectrales, se emplea una transformación Wavelet 2D para las dos dimensiones espaciales y una transformación discreta del Coseno 1D para la dimensión espectral; por lo que para la cuarta dimensión, la dimensión temporal, se realizó un análisis de las diferentes posibilidades de transformación (C. Correa y cols., 2016).

Con el objetivo de emplear dicha transformación 4D en la representación de video espectral para este trabajo, se replicaron las pruebas hechas en (C. Correa y cols., 2016), para así, analizar y escoger la mejor representación escasa. Para esto se toman diferentes combinaciones de transformadas en  $\Psi_{4D}$  y se mide la compresión de la señal en esa base. Específicamente, se toman en cuenta las combinaciones de transformación para cada dimensión según la Tabla 1.

Tabla 1: Combinaciones de diferentes bases de representación para el producto Kronecker. W: Transformada Wavelet; D: Transformada discreta coseno. Tomado y traducido de (C. Correa y cols., 2016).

Nombre de la base	Dimensión espacial $\Psi_1 \otimes \Psi_2$	Dimensión espectral $\Psi_3$	Dimensión temporal $\Psi_4$
WWDD	Wavelet 2D	DCT 1D	DCT 1D
WWWW	Wavelet 2D	Wavelet 1D	Wavelet 1D
WWWD	Wavelet 2D	Wavelet 1D	DCT 1D
WWDW	Wavelet 2D	DCT 1D	Wavelet 1D

Con el objetivo de mostrar la *escasez* de un video espectral en las bases de representación indicadas, la Fig.8 muestra los primeros 100 coeficientes ordenados de mayor a menor, para cada base y dos videos espectrales seleccionados. Tomando solo el 1 % del total de coeficientes de  $\boldsymbol{\theta}$  más significativos de la representación se puede recuperar una aproximación de la señal, esto es que, con tan solo el 1 % de los coeficientes más grandes es posible obtener una buena aproximación de la señal en términos de PSNR (un mayor PSNR indica una mejor calidad de reconstrucción). La Figura 7 presenta la reconstrucción de los videos espectrales usados como ejemplo, a partir del 1 % de coeficientes en las diferentes bases de representación.

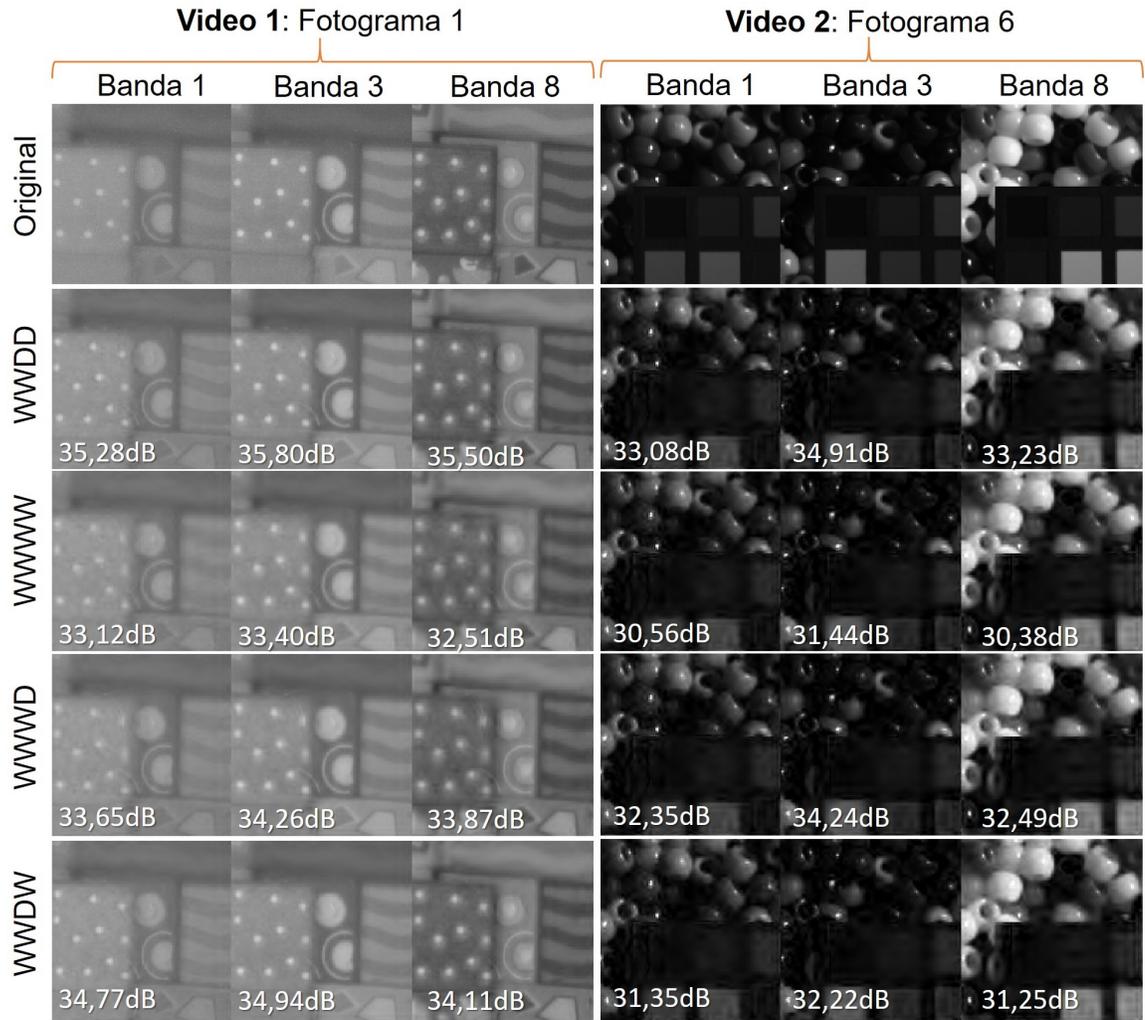


Figura 7: Reconstrucción a partir de la representación *escasa* de dos videos espectrales usando las diferentes bases de representación indicadas en la Tabla 1. Para cada video espectral, se toma solo el 1% de los coeficientes más significativos en la transformación y se obtiene una aproximación con una calidad de hasta 35 dB. En la parte superior, se muestra el fotograma 1 original para cada video, sin compresión, para las bandas espectrales 1, 3 y 8; en las siguientes filas se muestran las respectivas reconstrucciones usando el 1% de los coeficientes y su respectivo PSNR con respecto a la banda original.

Finalmente, este análisis deja a la vista las propiedades de compresión que tienen dichas transformadas presentadas; aunque todas las transformaciones proveen resultados similares, las transformaciones WWDD y WWWD presentan un decaimiento más pronunciado en la Fig. 8 y una mayor calidad de reconstrucción en la Fig. 7, por lo que se puede concluir que dichas transformaciones proveen una mejor representación *escasa* de la señal.

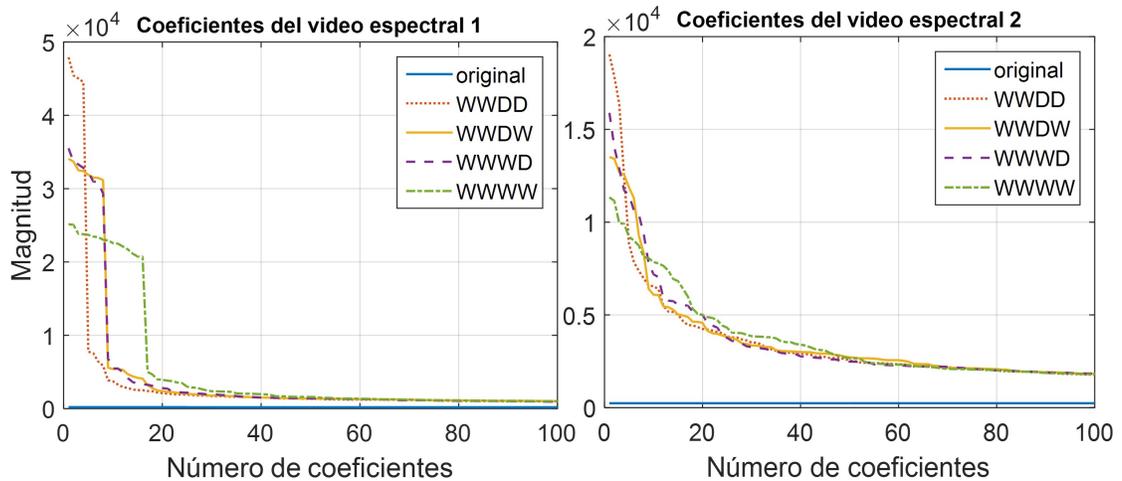


Figura 8: Coeficientes en cada base de representación indicada en la Tabla 1, para dos videos espectrales con dimensiones espaciales  $M = 128, N = 128, L = 8$  bandas espectrales y  $D = 8$  fotogramas.

#### 2.4. Enfoque de múltiples resoluciones en la reconstrucción de video compresivo

En la última década, diferentes trabajos relacionados con la adquisición de video compresivo han enfocado sus esfuerzos en mejorar la calidad de reconstrucción ya que este tipo de datos exhibe una gran cantidad de retos originados por la variable temporal. El problema principal está en la naturaleza de movimiento que presenta un video. El movimiento durante la adquisición produce artefactos, y estos artefactos, durante el proceso de reconstrucción, producen un fenómeno llamado *aliasing* que daña la calidad del video reconstruido (Sankaranarayanan y cols., 2015). Una de las propuestas que ha mejorado considerablemente este problema es hacer uso de la correlación temporal que existe entre cada fotograma del video y estimar el movimiento para corregir dichos artefactos en la reconstrucción. Sin embargo, dado que este enfoque requiere el video reconstruido, este trabajo propone realizar una reconstrucción de baja resolución con el fin de mitigar los gastos de cómputo en la reconstrucción inicial del video. En otras palabras, dado que se requiere mejorar la calidad de reconstrucción de un video adquirido en forma compresiva, se propuso estimar una aproximación de baja resolución del video con la que se pueda calcular el movimiento para luego usar esta información en la reconstrucción de alta resolución, específicamente en la corrección de artefactos inducidos por la naturaleza de movimiento de la escena. A este enfoque se le llamó **enfoque de múltiples resoluciones**, ya que se hace uso de una estimación tosca o de baja resolución de la señal en movimiento para así mejorar la señal en la reconstrucción de alta resolución (Reddy y cols., 2011; Sankaranarayanan y cols., 2015).

### 2.4.1. Estimación de baja resolución espacial

Dadas las medidas adquiridas en el detector  $\mathbf{g} = \Phi\mathbf{f}$ , como en la Ecuación 2.2, se agrega un operador decimador  $\mathbf{Q}$  para la reducción de dimensión espacial en la señal  $\mathbf{f}$ , esto es

$$\mathbf{g} = \Phi\mathbf{Q}^T\mathbf{Q}\mathbf{f}, \quad (2.7)$$

siendo  $\mathbf{Q}^T$  el operador  $\mathbf{Q}$  transpuesto. Dado el sistema en la Ecuación 2.7, la estimación de baja resolución se encuentra haciendo la operación pseudoinversa expresada como

$$\tilde{\mathbf{f}} = (\Phi\mathbf{Q}^T)^\dagger\mathbf{g}, \quad (2.8)$$

donde  $\tilde{\mathbf{f}}$  es la estimación de baja resolución y el superíndice  $(\cdot)^\dagger$  indica la pseudoinversa de la matriz.

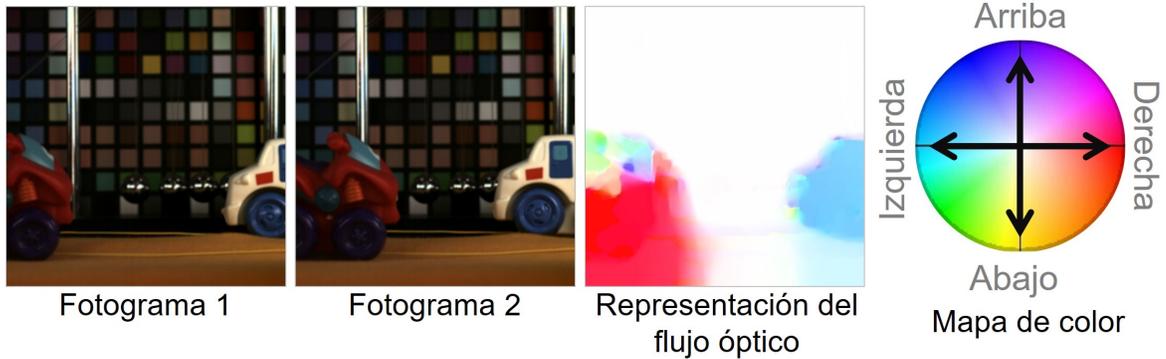


Figura 9: Ilustración de la representación de flujo óptico basado en un mapa de color. El movimiento entre el fotograma 1 y 2 puede ser representado en colores partiendo del flujo óptico usando el mapa de color.

### 2.4.2. Estimación del movimiento: Flujo óptico

Gracias al enfoque de múltiples resoluciones es posible obtener características del video comprimido con un bajo costo computacional, ya que no es necesario recuperar la señal en alta resolución. Entre las características que se pueden estimar en baja resolución está el movimiento innato de los objetos dentro de la escena, y una forma de estimar este movimiento es a través del flujo óptico. El flujo óptico permite determinar el movimiento horizontal y vertical en una secuencia de imágenes dinámicas con base en los cambios de luminosidad de la escena (Horn y Schunck, 1981). Los dos tipos de movimiento, horizontal y vertical, se representan como el cambio en el tiempo de las posiciones espaciales  $x, y$  y se pueden expresar como

$$u = \frac{dx}{dt}, v = \frac{dy}{dt}. \quad (2.9)$$

Los mapas de color son una forma alterna de representar el movimiento en una escena, y se basan en las componentes  $u$  y  $v$  calculadas a través del flujo óptico (Baker y cols., 2011; Liu,

2009). Para ilustrar el concepto de mapa de color, en la Fig. 9 se muestran dos fotogramas en tiempos diferentes, donde el fotograma 1 muestra dos carros en una posición 1 y el fotograma 2, muestra los carros en una posición diferente a la inicial. En el fotograma 2, el carro rojo se ha movido a la izquierda y el carro blanco se ha movido a la derecha. El tercer cuadro de la Fig. 9 presenta la matriz de color del movimiento de los carros con base en el mapa de color de la izquierda, donde un movimiento a la derecha se torna de color rojo, un movimiento a la izquierda se torna color azul y un movimiento arriba o abajo se torna morado o amarillo, respectivamente (Sankaranarayanan y cols., 2015; Liu, 2009).

## 2.5. Reconstrucción de video a partir de las medidas comprimidas

El problema inverso de reconstrucción consiste en recuperar el video espectral a partir de los datos adquiridos de forma compresiva. En esta sección se presenta el planteamiento del problema de reconstrucción para diferentes sistemas de multiplexación, o compresión, de datos usando el enfoque de múltiples resoluciones, con el fin de mitigar los artefactos en la reconstrucción de video compresivo.

### 2.5.1. Reconstrucción de video compresivo en arquitecturas de multiplexación espacial, temporal y espectral

Las arquitecturas de multiplexación de datos están basadas en el muestreo compresivo. Existen diferentes formas de realizar multiplexación de datos en sistemas de muestreo compresivo. Para el caso de video, la multiplexación puede ser de forma espacial o temporal, donde la multiplexación espacial corresponde a la disminución de medidas en el sensor y la temporal, corresponde a la disminución de los fotogramas por segundo. Arquitecturas de adquisición compresiva de video como CS-MUVI (*Compressive Sensing Multi-scale Video*) y P2C2 (*Programmable Pixel Compressive Camera*) son ejemplos de sistemas de multiplexación espacial y temporal, respectivamente (Sankaranarayanan y cols., 2015; Reddy y cols., 2011), donde no se agrega la componente espectral al video. Debido al problema de aliasing mencionado en la sección anterior, estos sistemas ya se han implementado haciendo uso del enfoque de múltiples resoluciones para mejorar la calidad de reconstrucción (Sankaranarayanan y cols., 2015; Reddy y cols., 2011; Goldstein y cols., 2015). Sin embargo, en sistemas de multiplexación espectral, para la adquisición de video espectral, como lo es el ya mencionado sistema CASSI (*Coded Aperture Snapshot Spectral Imager*), este enfoque de múltiples resoluciones no ha sido explorado.

#### 2.5.1.1. Reconstrucción de video en un sistema de multiplexación espacial (CS-MUVI)

El problema de reconstruir un video adquirido de forma compresiva por un sistema de multiplexación espacial es expresado como

$$\hat{\mathbf{f}} = \begin{cases} \text{mín} & \sum \|\Psi^T \hat{\mathbf{f}}\|_1 \\ \text{sujeto a} & \|\langle \phi_t, \mathbf{f}_t \rangle - g_t\|_2 \leq \epsilon_1, \\ & \|\hat{\mathbf{f}}_i(x, y) - \hat{\mathbf{f}}_j(x + u, y + v)\|_2 \leq \epsilon_2, \forall i, j, \end{cases} \quad (2.10)$$

donde se minimiza la representación escasa de  $\hat{\mathbf{f}}$  con respecto a las medidas adquiridas y a la relación de cambio dada por  $\hat{\mathbf{f}}_i(x, y) - \hat{\mathbf{f}}_j(x + u, y + v)$ , siendo  $u$  y  $v$  la información de cambio horizontal y vertical, respectivamente; y donde los subíndices  $i, j$  representan dos fotogramas tomados en dos tiempos distintos, con  $i > j$ . Los parámetros  $\epsilon_1$  y  $\epsilon_2$  corresponden a indicadores de nivel de ruido y brillo (Sankaranarayanan y cols., 2015). Cabe resaltar, que la información  $u$  y  $v$  es estimada calculando el flujo óptico del video en baja resolución espacial, por lo que estos parámetros son información obtenida previamente al proceso de reconstrucción. Este tipo de problemas son llamado problemas inversos con restricción, ya que están restringidos o sujetos a una condición determinada.

### 2.5.1.2. Reconstrucción de video en un sistema de multiplexación temporal (P2C2)

El problema inverso para la adquisición con el sistema de multiplexación temporal se expresa como

$$\hat{\mathbf{f}}^k = \arg \min \sum_{k=1}^D \beta \|\Psi^{-1} \mathbf{f}_t\|_1 + \|\mathbf{g} - \Phi \mathbf{f}\|_2^2 + \lambda \|\Omega^{k-1} \mathbf{f}\|_2^2 \quad (2.11)$$

donde  $\beta$  y  $\lambda$  son regularizadores y se busca minimizar la energía de la función usando la matriz  $\Omega^{k-1}$  obtenida en una iteración inicial  $\hat{\mathbf{f}}^0$ . Es decir, La matriz  $\Omega$  contiene la información del movimiento del video estimado con base en la reconstrucción de la iteración inicial  $\hat{\mathbf{f}}^0$  (Reddy y cols., 2011).

### 2.5.1.3. Reconstrucción de video en un sistema de multiplexación espectral (CASSI)

Ahora bien, cuando se agrega la componente espectral a la adquisición de video, sistemas como CASSI, realizan una multiplexación espectral para adquirir la información de forma comprimida. El problema de reconstrucción tradicional se expresa como

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \|\mathbf{g} - \Phi \Psi \boldsymbol{\theta}\|_2^2 + \tau \|\boldsymbol{\theta}\|_1 \quad (2.12)$$

donde  $\mathbf{g}$ ,  $\Phi$ ,  $\Psi$  y  $\boldsymbol{\theta}$  son las medidas, la matriz de muestreo, la base de representación y la representación *escasa*, en orden respectivo;  $\tau$  es un parámetro de regularización,  $\|\cdot\|_1$  es la norma  $\ell_1$  que mide la cantidad de valores diferentes de cero en el vector y  $\|\cdot\|_2$  es la norma  $\ell_2$  usada para medir el error cuadrático medio de la estimación. La reconstrucción del video esta dada por  $\hat{\mathbf{f}} = \Psi^T \hat{\boldsymbol{\theta}}$ . Sin embargo, esta solución presenta problemas ya que agrega artefactos por la variable temporal. En el capítulo siguiente, se presentará el problema inverso planteado en la Ecuación 2.12 modificado con el fin de agregar información de la variable temporal, como el flujo óptico, en la reconstrucción del video espectral.

### 3. Diseño de aperturas de codificación y enfoque de reconstrucción con múltiples resoluciones para la adquisición compresiva de video espectral

En este capítulo se presenta el desarrollo del algoritmo de diseño de aperturas de codificación para el sistema de multiplexación espectral CASSI. Estas aperturas diseñadas contienen una distribución uniforme de los elementos de codificación haciendo que mejore el muestreo y por lo tanto, la calidad de reconstrucción en términos de PSNR, para la adquisición compresiva y reconstrucción de video espectral. Por otro lado, el diseño permite aplicar el enfoque de reconstrucción con múltiples resoluciones ya que es posible obtener una reconstrucción de baja resolución de buena calidad. Por lo que, adicionalmente, se presenta la modificación al problema tradicional de minimización para que sea posible agregar la correlación temporal estimada de la reconstrucción de baja resolución.

#### 3.1. Preliminares

La selección de la base de representación es fundamental en el proceso de reconstrucción de la señal sub-muestreada. Sin embargo, adicional a esta selección, la matriz del sistema de multiplexación también influye en la calidad de las reconstrucciones. La matriz del sistema de multiplexación se define como

$$\mathbf{A} = \mathbf{\Phi}\mathbf{\Psi} = \mathbf{P}\mathbf{T}\mathbf{\Psi}, \quad (3.1)$$

con  $\mathbf{A} \in \mathbb{R}^{m \times n}$ , donde  $\mathbf{\Psi}$  es la base de representación (Ec. 2.6),  $\mathbf{P}$  la matriz de dispersión y  $\mathbf{T}$  la matriz de codificación (Ec. 2.3). Observe que las matrices  $\mathbf{\Psi}$  y  $\mathbf{P}$  son matrices inalterables y la matriz  $\mathbf{T}$  depende del patrón de la apertura de codificación.

Por lo tanto, el sistema de la Ecuación 2.2 se puede reescribir como

$$\mathbf{g} = \mathbf{A}\boldsymbol{\theta}. \quad (3.2)$$

Una posible solución al sistema 3.2 puede ser calculada resolviendo la pseudoinversa de la matriz  $\mathbf{A}$  como

$$\begin{aligned} \mathbf{A}\boldsymbol{\theta} &= \mathbf{g} \\ \mathbf{A}^T\mathbf{A}\boldsymbol{\theta} &= \mathbf{A}^T\mathbf{g} \\ \boldsymbol{\theta} &= (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{g}, \end{aligned}$$

donde la pseudoinversa de  $\mathbf{A}$  es  $\mathbf{A}^\dagger = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$ . Sin embargo, esta aproximación de inversión depende directamente de las entradas de la matriz  $\mathbf{T}$ , ya que las matrices  $\mathbf{\Psi}$  y  $\mathbf{P}$  son matrices fijas en  $\mathbf{A}$  (Fig. 10). Cabe recalcar que la matriz  $\mathbf{T}$  conserva en su diagonal los elementos de paso y bloqueo de la apertura de codificación. Esto es, si se tiene una apertura de codificación  $\mathbf{C} \in \{0, 1\}$  de tamaño  $M \times N$ , con un número de elementos de paso en un 50% del total de elementos, o una transmitancia de 50%, los elementos  $\mathbf{t}_k^d$  que corresponden a la diagonal de  $\mathbf{T}$ , se toman de la vectorización de  $\mathbf{C}$  como

$$\mathbf{t}_k^d = [\mathbf{c}_0, \dots, \mathbf{c}_{N-1}], \quad (3.3)$$

donde  $\mathbf{c}_n \in \mathbb{R}^M$  corresponde a cada columna de  $\mathbf{C}$  para  $n = 0, \dots, N - 1$ ,  $k = 0, \dots, L - 1$  y  $d = 0, \dots, D - 1$ . Por lo tanto, para video espectral,  $\mathbf{T} = \text{blkdiag}\{(\tilde{\mathbf{T}}^0)^T, \dots, (\tilde{\mathbf{T}}^{D-1})^T\}^T$  se compone de  $D - 1$  matrices estructuradas de la siguiente manera

$$\tilde{\mathbf{T}}^d = \begin{bmatrix} \text{diag}(\mathbf{t}_0^d) & \mathbf{0}_{N^2 \times N^2} & \cdots & \mathbf{0}_{N^2 \times N^2} \\ & \text{diag}(\mathbf{t}_1^d) & \cdots & \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{N^2 \times N^2} & \mathbf{0}_{N^2 \times N^2} & \cdots & \text{diag}(\mathbf{t}_{L-1}^d) \end{bmatrix}_{N^2 L \times N^2 L}$$

Con el fin de mejorar la calidad de las reconstrucciones, se realiza una mejora en la condicionalidad de la matriz del sistema de multiplexación para la adquisición de video espectral a través del diseño de la matriz  $\mathbf{C}$  correspondiente a las diagonales en  $\mathbf{T}$ . Entre los trabajos que proponen el diseño específico de la matriz  $\mathbf{T}$  para mejorar la inversión de la matriz  $\mathbf{A}^T \mathbf{A}$ , está el desarrollado en (Arguello y Arce, 2014), donde se establece que minimizando la correlación de los elementos en la matriz  $\mathbf{C}$ , o la diagonal en  $\mathbf{T}$ , mejora la condicionalidad de  $\mathbf{A}$  y por lo tanto, la inversión del sistema. Otro trabajo plantea la adquisición de múltiples capturas (*multishot*) de la señal para mejorar la condicionalidad del sistema (Arguello y Arce, 2011b; C. V. Correa y cols., 2016). Sin embargo, ninguno de estos trabajos ha sido enfocado al diseño de la apertura de codificación en la adquisición compresiva de video espectral o escenas espectrales dinámicas, donde existe una variable temporal que no puede ser ignorada en el proceso de adquisición y donde no es posible realizar múltiples capturas debido a esa misma variable.

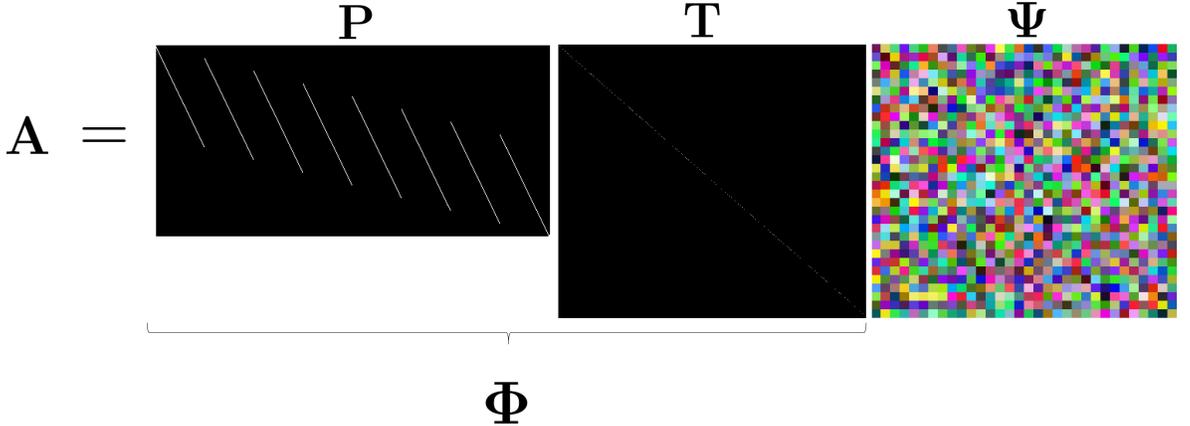


Figura 10: Matriz del sistema de multiplexación espectral CASSI

### 3.2. Criterios de diseño de las aperturas de codificación

La distribución espacial, espectral y temporal de los elementos de paso (1) y bloqueo (0) en la apertura de codificación  $\mathbf{C} \in \{0, 1\}$ , establecen el muestreo de las características de la escena bajo observación, donde los elementos de paso son los que determinan la información de la escena que se adquiere en el detector. Por consiguiente, para video espectral compresivo, además de realizar el muestreo y compresión al mismo tiempo, se busca adquirir la información

más relevante de la escena. Es por esto que se realiza un diseño de apertura de codificación que siga los siguientes criterios:

1. *Criterio de muestreo espacial y espectral:* El muestreo espacial y espectral de cada fotograma se realizará a través de la combinación de un conjunto de filtros ópticos. De modo que, además de tener un muestreo compresivo espacial se agrega la modulación espectral que brinda al sistema un muestreo más representativo a través de la componente espectral, en comparación con las aperturas binarias.
2. *Criterio de muestreo temporal:* El muestreo temporal se realizará de forma complementaria entre fotogramas y, dado a la naturaleza efímera de una escena dinámica, cada banda espectral de cada fotograma será muestreada espacialmente al menos una vez. Esto es, para cada fotograma, existirá información espacial en alguna de las bandas espectrales para asegurar la adquisición de objetos que entran, se desplazan o desaparecen en la escena.

### 3.3. Algoritmo de diseño de aperturas de codificación

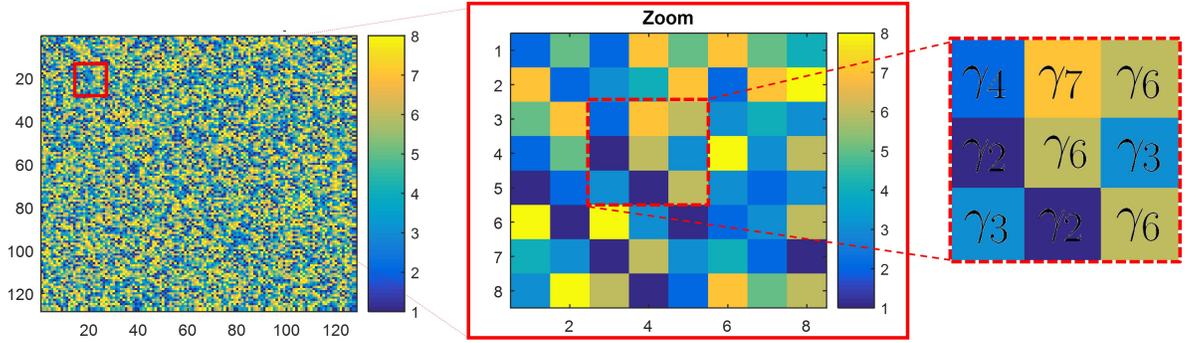
En el algoritmo 1, se presenta de manera formal el desarrollo de la creación de una apertura de codificación que sigue los criterios mencionados generando la estructura que considera el aspecto espacial, espectral y temporal de la escena. Las entradas del algoritmo son: un conjunto finito de filtros ópticos  $\xi$ , el número total de filtros  $F_N$  y la resolución espacial, espectral y temporal del video espectral como  $N, L, D$ . De forma general, el algoritmo tiene como objetivo generar una apertura de codificación ubicando filtros ópticos espacialmente. En detalle, el algoritmo inicia generando un vector  $\gamma = [\gamma_{\Omega_1}, \gamma_{\Omega_2}, \dots, \gamma_{\Omega_w}, \dots, \gamma_{\Omega_z}, \dots, \gamma_{\Omega_{F_N}}]$  de valores aleatorios entre 1 y  $F_N$ , donde los índices/subíndices  $\Omega_w, \Omega_z$  indican la aleatoriedad del vector para diferentes posiciones  $w, z$  de este. A su vez, el vector  $\gamma$  representa el índice de cada filtro óptico. En general, el proceso de diseño lo conforman tres bucles iterativos, dos de ellos anidados. En los bucles anidados, las variables  $i, j$  recorren espacialmente la matriz  $\mathbf{\Gamma} \in \mathbb{R}^{N \times N}$  posicionando uno por uno los filtros ópticos a través de su respectivo subíndice  $\gamma_{\Omega_w}$ . La asignación de filtros para la primera columna de  $\mathbf{\Gamma}$  se hace de forma aleatoria, esto es  $\mathbf{\Gamma}_{i,0} = \gamma_{\Omega_w}$ . Por otro lado, la asignación para la primera fila de  $\mathbf{\Gamma}$ , es decir  $\mathbf{\Gamma}_{0,j}$  ( $j = 0, \dots, N - 1$ ), contempla la minimización de la correlación entre la selección de los filtros. La mínima correlación obedece al mínimo valor de la norma  $\ell_2$  entre dos filtros adyacentes. Esto es, si para la posición  $(0, j - 1)$  se ha escogido un filtro óptico  $\gamma_{\Omega_w}$  como  $\xi_{\gamma_{\Omega_w}}$  desde  $\xi$ , la selección de un siguiente filtro óptico  $\gamma_{\Omega_z}$  como  $\xi_{\gamma_{\Omega_z}}$  para el píxel inmediatamente siguiente  $(0, j)$ , debe satisfacer que

$$\mathbf{\Gamma}_{i,j} = \underset{r}{\operatorname{argmin}} \|\xi_{\mathbf{\Gamma}_{i,j-1}} \xi_r^T\|_2, \quad (3.4)$$

para  $r = 1, \dots, F_N$ . Cuando se tiene que  $i > 0$  y  $j > 0$ , la restricción de la norma  $\ell_2$  en la selección de filtros se extiende a

$$\mathbf{\Gamma}_{i,j} = \underset{r}{\operatorname{argmin}} (\|\xi_{\mathbf{\Gamma}_{i,j-1}} \xi_r^T\|_2 + \delta \cdot \|\xi_{\mathbf{\Gamma}_{i-1,j}} \xi_r^T\|_2), \quad (3.5)$$

donde  $\delta$  es una constante auxiliar con valor diferente de 0 solo cuando se cumple que  $i$  es mayor a 0 ( $i > 0$ ), y donde se agrega el término  $\|\xi_{\mathbf{\Gamma}_{i-1,j}} \xi_r^T\|_2$ , que tiene en cuenta la selección



Distribución Espacial de filtros ópticos

Figura 11: Distribución espacial de filtros ópticos  $\Gamma$  (algoritmo 1, línea 8). La distribución espacial se genera ubicando los filtros de manera que la correlación espacio-espectral entre ellos sea la menor. Cada color en el arreglo muestra un índice de un filtro del conjunto  $\xi$ .

del filtro en la posición  $(i-1, j)$ . En otras palabras, la extensión de 3.4 a 3.5 se realiza con el fin de seleccionar y ubicar los filtros espacialmente con la mínima correlación tanto vertical como horizontal. De esta manera, la restricción en la Ecuación 3.5 presenta, no solo la minimización con respecto al filtro ubicado en la posición anterior  $(i, j-1)$ , sino también la minimización con el filtro ubicado inmediatamente en la parte superior de la posición  $(i, j)$ , es decir, el filtro ubicado en  $(i-1, j)$ . La Figura 11 presenta un ejemplo de la distribución espacial de filtros  $\Gamma$  obtenida al finalizar las  $N \times N$  iteraciones en los bucles anidados. Observe que aunque se minimiza la correlación de la selección y ubicación espacio-espectral de los filtros, la matriz  $\Gamma$  almacena solo los índices  $\gamma = [\gamma_{\Omega_1}, \gamma_{\Omega_2}, \dots, \gamma_{\Omega_w}, \dots, \gamma_{\Omega_z}, \dots, \gamma_{\Omega_{F_N}}]$  que corresponden al filtro seleccionado desde  $\xi$ . Hay que mencionar, además, que la componente aleatoria de la primera columna induce un valor aleatorio en el diseño de la apertura, lo que es favorable para el diseño, ya que un patrón con una estructura definida, o con escasa aleatoriedad, implica un aumento en la correlación de los elementos (para más detalles ver (Rauhut, 2010)). El tercer bucle corresponde a la componente temporal, donde se realiza una repetición de  $\Gamma$  con base en el modulo para  $d = 0, \dots, D-1$  fotogramas espectrales para asegurar una máxima separación de los elementos en la componente temporal. Dicho de otra forma, la distribución espacial realizada es aprovechada a lo largo de dicha componente temporal haciendo una repetición de esta. Por último, con la operación  $\mathbf{C} \leftarrow \mathbf{A}(\xi_r)$  se lleva el diseño desde la distribución espacial de índices  $\Gamma$  a la apertura de codificación final  $\mathbf{C}$ , con  $r = 1, \dots, F_N$ . La apertura de codificación de salida tiene  $N \times N$  píxeles espaciales,  $L$  matrices para la codificación espectral, y  $D$  matrices espacio-espectrales, para la codificación temporal ( $\mathbf{C} \in \mathbb{R}^{N \times N \times L \times D}$ ).

Las aperturas de codificación diseñadas a partir del Algoritmo 1 son denominadas aperturas de codificación de color temporal (T-CCA del inglés *Temporal Colored Coded Apertures*), y reciben este nombre debido a que son aperturas de codificación diseñadas utilizando filtros ópticos que varían en el tiempo. Es importante resaltar que las T-CCA permiten el muestreo espacial al menos una vez, lo que significa que los cambios en cada fotograma son siempre muestreados. El conjunto de filtros ópticos óptimo para el Algoritmo 1 es presentado en la Fig. 12. Note que el conjunto de filtros óptimos tiene un número finito de filtros, lo cual permite que sea posible realizar una implementación real de la apertura de codificación.

---

**Algoritmo 1** Diseño de una apertura de codificación espacio-espectro-temporal

---

**Entrada:**  $\xi$  (conjunto finito de filtros ópticos),  $F_N$  (número de filtros) y  $N, L, D$  (resolución espacio-espectro-temporal).

**Salida:**  $\mathbf{C} \in \mathbb{R}^{N \times N \times L \times D}$ : Apertura de codificación temporal de color (T-CCA)

- 1:  $\gamma = [\gamma_{\Omega_1}, \gamma_{\Omega_2}, \dots, \gamma_{\Omega_w}, \dots, \gamma_{\Omega_z}, \dots, \gamma_{\Omega_{F_N}}]$ ,  $\Omega_w \neq \Omega_z, w \neq z$
  - 2:  $\delta = 0$
  - 3: **para**  $i \leftarrow 0$  hasta  $N - 1$  **hacer**
  - 4:    $\Gamma_{i,0} = \gamma_{\Omega_w}$
  - 5:   **para**  $j \leftarrow 1$  hasta  $N - 1$  **hacer**
  - 6:      $\hat{r} = \underset{r}{\operatorname{argmin}}(\|\xi_{\Gamma_{i,j-1}} \cdot \xi_r^T\|_2 + \delta \cdot \|\xi_{\Gamma_{i-1,j}} \cdot \xi_r^T\|_2)$
  - 7:     con  $r = 1, \dots, F_N$
  - 8:      $\Gamma_{i,j} \leftarrow \hat{r}$
  - 9:   **fin para**
  - 10:    $\delta = 1$
  - 11: **fin para**
  - 12: **para**  $d \leftarrow 0$  hasta  $D - 1$  **hacer**
  - 13:    $\mathbf{A}_d \leftarrow \Gamma$
  - 14:    $\Gamma \leftarrow (\Gamma \bmod F_N) + 1$
  - 15: **fin para**
  - 16:  $\mathbf{C} \leftarrow \mathbf{A}(\xi_r)$  con  $r = 1, \dots, F_N$
- 

Además, observe en la Fig. 12, que la distancia entre los elementos de paso entre los filtros ópticos es máxima lo que provee una mínima correlación espectral (C. V. Correa y cols., 2016).

### 3.3.1. Análisis espacio-temporal del diseño

La distribución de los elementos de paso en una apertura de codificación  $\mathbf{C}$  diseñada para video espectral se puede analizar en el espacio, el espectro y el tiempo. Como se ha mencionado antes, la máxima separación de los elementos de paso en  $\mathbf{C}$  implica una correlación baja de los mismos, lo que conlleva una mejor condicionalidad de la matriz  $\mathbf{A}$  como se demuestra en (Arguello y Arce, 2014; C. V. Correa y cols., 2016). Por lo que para demostrar el funcionamiento del algoritmo de diseño con respecto a la mejora de la condicionalidad de  $\mathbf{A}$  desde  $\mathbf{T}$ , se presenta una realización de una apertura T-CCA y una realización de una apertura generada con una distribución aleatoria con un 25% de elementos de paso del 100% de elementos que conforma la distribución, con  $N = 8, L = 8$  y  $D = 8$  para ambas realizaciones.

La Figura 13 presenta la distribución espacial de las realizaciones mencionadas, donde se puede apreciar la distribución uniforme que alcanza las T-CCA en comparación con la distribución aleatoria. Note que las T-CCA no presentan agrupamientos de elementos de paso, los cuales están representados por los cuadros blancos, lo cual contribuye a la característica de máxima separación entre elementos de paso. Además, note en el recuadro ampliado en la parte inferior, la característica de complementariedad a través del tiempo  $d = 1, \dots, 4$  con la que se diseñan las T-CCA, lo que significa que cada punto espacial es muestreado al menos una vez para cada fotograma.

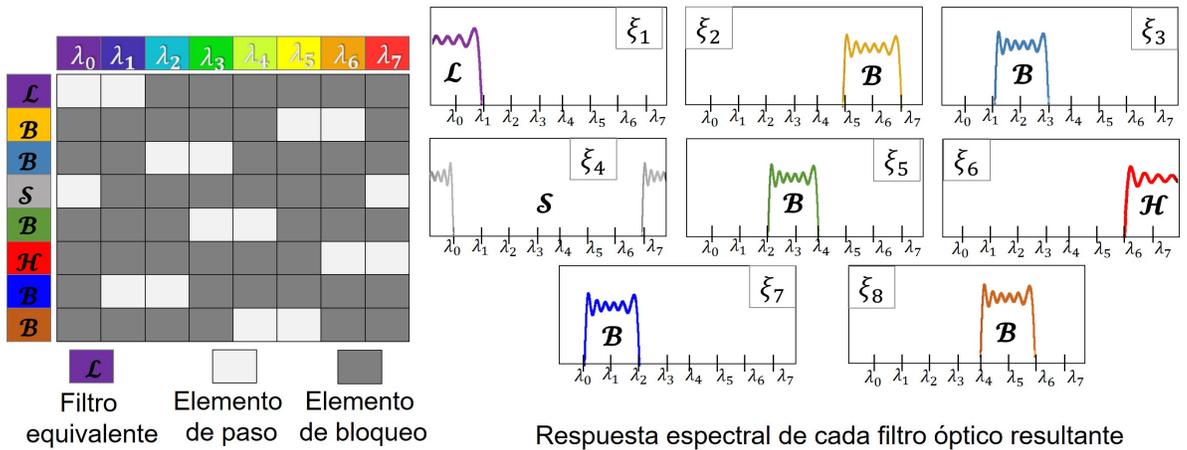


Figura 12: Conjunto de filtros ópticos seleccionados  $\xi$ . (Izquierda) Se muestra la respuesta espectral de cada filtro y de forma discreta (Derecha), en elementos de paso o bloqueo con su respectivo equivalente, donde  $\mathcal{L}$  denota un filtro pasa-baja,  $\mathcal{H}$  un filtro pasa-alta,  $\mathcal{B}$  un filtro pasa-banda y  $\mathcal{S}$  un filtro elimina-banda.

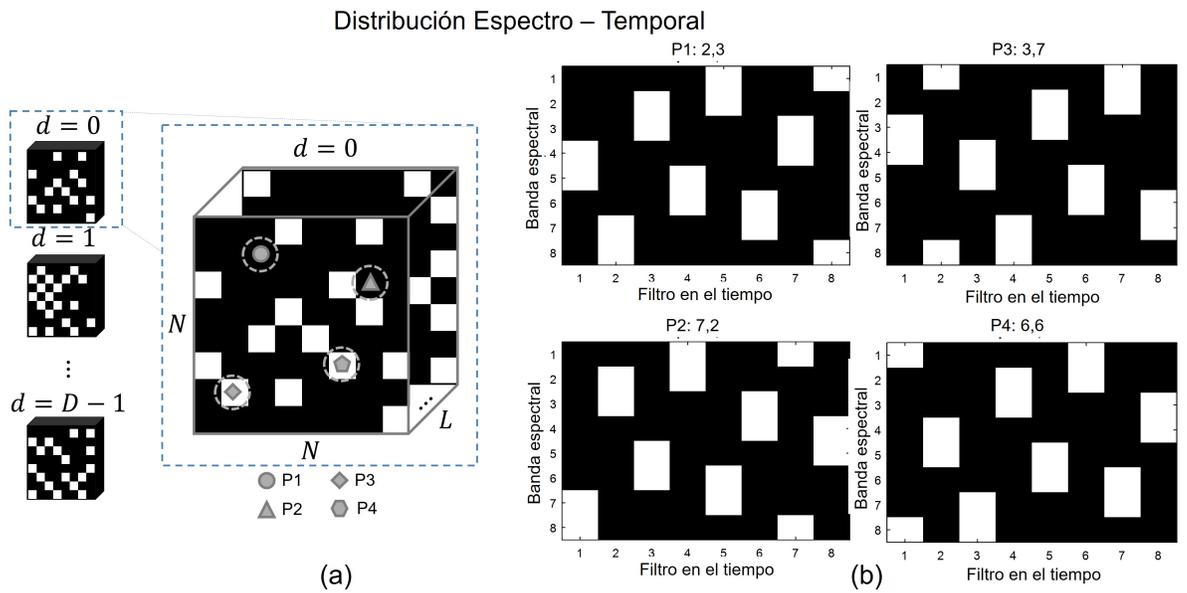


Figura 14: Realización de una apertura T-CCA. Los elementos blancos representan el paso y los elementos negros representan el bloqueo de la información (con  $8 \times 8$  píxeles espaciales). (a) Distribución espacial de los elementos de paso. (b) Característica espectro-temporal de las T-CCA: con las T-CCA se asegura un muestreo uniforme a través del tiempo haciendo que un punto espacial sea muestreado de forma complementaria en el tiempo.

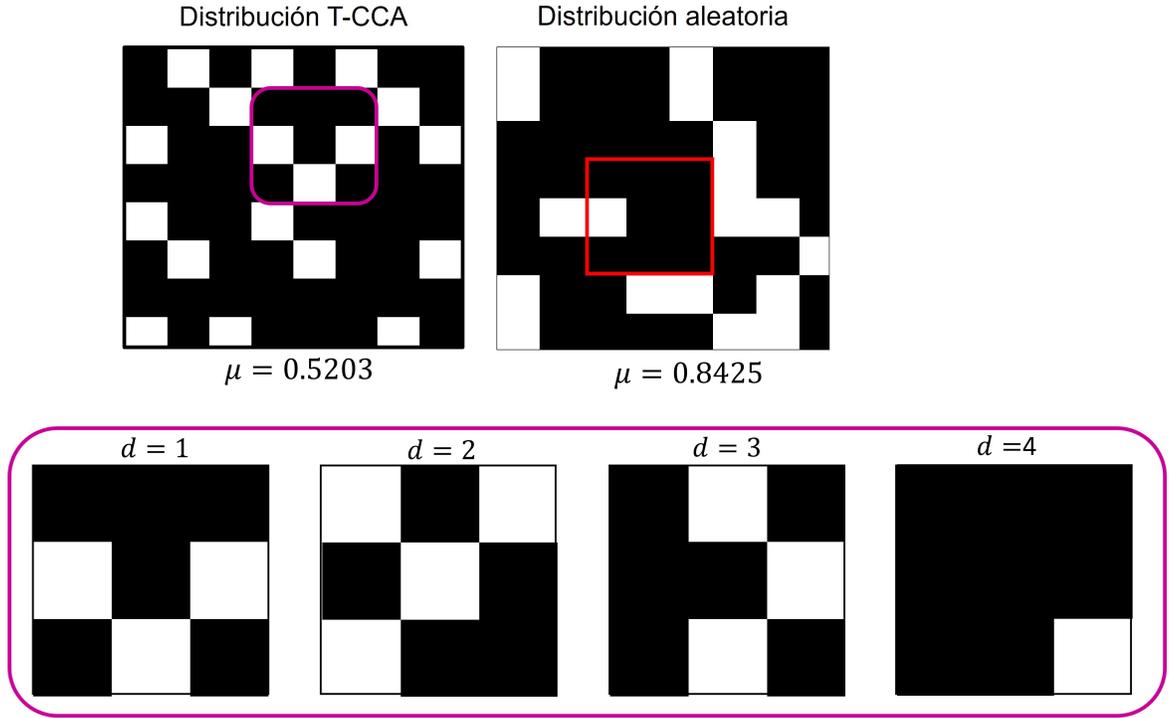


Figura 13: Comparación de la distribución espacial de una apertura generada con el algoritmo propuesto, T-CCA, y una apertura generada con una distribución aleatoria con un 25 % de elementos de paso y 75 % de elementos de bloqueo. En la sección resaltada se muestra la característica complementaria a través del tiempo para  $D = 4$ .

### 3.3.2. Análisis espectro-temporal del diseño

La característica temporal se ilustra de forma más amplia en la Figura 14. Dada la realización de la apertura T-CCA mostrada en la parte (a) de la Fig. 14, se toman 4 puntos espaciales P1, P2, P3 y P4, a través del espectro en la apertura para  $d = 0$ , esto es, tomar el filtro óptico que se ubicó en esos cuatro puntos seleccionados del “fotograma” 0. Específicamente, los puntos espaciales que se seleccionan de la apertura en  $d = 0$  son:  $i = 2, j = 3$  (P1),  $i = 7, j = 2$  (P2),  $i = 3, j = 7$  (P3) y  $i = 6, j = 6$  (P4).

Luego, a esa selección espacial, se le hace seguimiento a través del tiempo, lo que resulta en el seguimiento de la selección de filtros en cada apertura espacio-espectral con  $d = 0, \dots, D - 1$ , en los puntos específicos mencionados. En la parte (b) de la figura se ilustra dicho seguimiento, donde cada gráfico representa cada punto seleccionado; los ejes de cada gráfico están constituidos como sigue: el eje horizontal representa el punto en diferentes instante de tiempo  $d = 1, \dots, 8$ , y el eje vertical representa la respuesta espectral del filtro en ese punto. De la misma forma que en el análisis anterior, se destaca la distribución uniforme de las aperturas T-CCA en el tiempo, ya que la distribución temporal mantiene una máxima separación entre los elementos de paso (elementos blancos) a lo largo del tiempo, como se muestra en la Fig. 14(b).

### 3.3.3. Coherencia mutua

La coherencia mutua es un parámetro que permite la caracterización del grado de similitud o correlación que existe entre las columnas de  $\mathbf{A}$ . Una forma de cuantificar la calidad del diseño propuesto es calculando la coherencia mutua de  $\mathbf{A}$ , siendo deseable un valor bajo de coherencia que indique mejor calidad de  $\mathbf{A}$ . La coherencia mutua de una matriz  $\mathbf{A}$ ,  $\mu(\mathbf{A})$ , es el valor más grande resultante del producto interno entre dos columnas cualquiera  $\mathbf{a}_i, \mathbf{a}_j$  de  $\mathbf{A}$ :

$$\mu(\mathbf{A}) = \max_{1 \leq i, j \leq N} \frac{|\langle \mathbf{a}_i, \mathbf{a}_j \rangle|}{\|\mathbf{a}_i\|_2 \|\mathbf{a}_j\|_2}. \quad (3.6)$$

Esta definición se puede interpretar como el máximo elemento fuera de la diagonal de la matriz  $\mathbf{A}^T \mathbf{A}$  con sus columnas normalizadas (Gribonval y Nielsen, 2003; Eldar y Kutyniok, 2012).

La Figura 13 muestra la coherencia  $\mu$  para cada realización, donde una coherencia baja representa una menor correlación entre los elementos de la apertura de codificación. Observe que la coherencia en la distribución T-CCA presenta un valor menor con respecto a la distribución aleatoria.

## 3.4. Reconstrucción restringida en video compresivo espectral

Esta investigación, además de proponer un algoritmo para el diseño del elemento de codificación, plantea un problema inverso restringido basado en múltiples resoluciones y flujo óptico. Este problema, se plantea agregando una restricción al problema inverso típico en el estado del arte para reconstruir señales espectrales adquiridas con un sistema de multiplexación espectral. Con el objetivo de aprovechar la correlación temporal, mencionada anteriormente, se calcula el movimiento entre los fotogramas con base en la estimación de baja resolución. El cálculo se realiza usando el flujo óptico, con el cual se obtienen las matrices  $u$  y  $v$  que representan el movimiento horizontal y vertical de los objetos dentro la escena, respectivamente. Esta información temporal es agregada al problema en la Ecuación 2.12 como la minimización de la diferencia entre dos fotogramas en tiempos  $i$  y  $j$  ( $i > j$ ) con los desplazamientos  $u$  y  $v$ . El problema de minimización agregando dicha información estimada *a priori* se escribe como

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} (\|\mathbf{g} - \Phi \Psi \boldsymbol{\theta}\|_2^2 + \tau \|\boldsymbol{\theta}\|_1 + \beta \|\Lambda(\Psi \boldsymbol{\theta})_i(x, y) - \Lambda(\Psi \boldsymbol{\theta})_j(x + u, y + v)\|_2^2), \quad (3.7)$$

donde además de minimizar la diferencia entre la estimación y las medidas y la *escasez* de la señal, se minimiza el error causado por la variable temporal. El parámetro  $\beta$  es un regularizador de la restricción temporal y  $\Psi \boldsymbol{\theta}$  se define como

$$\Psi \boldsymbol{\theta} = [(\Psi \boldsymbol{\theta})_1^T, \dots, (\Psi \boldsymbol{\theta})_i^T, \dots, (\Psi \boldsymbol{\theta})_j^T, \dots, (\Psi \boldsymbol{\theta})_D^T], \quad (3.8)$$

con  $i > j$ . El operador  $\Lambda(\cdot)$  permite reorganizar un vector en forma matricial. Básicamente, la modificación planteada en la Ecuación 3.7 corrige la estimación de la señal, iterativamente, con base en la información de movimiento de los píxeles dada por  $u$  y  $v$ . La reconstrucción del video espectral se expresa como  $\hat{\mathbf{f}} = \Psi \hat{\boldsymbol{\theta}}$ .

## 4. Simulaciones y Resultados

La evaluación del desempeño de las aperturas de codificación diseñadas se realizó mediante simulaciones siguiendo el modelo de muestreo planteado en la Ecuación 2.2. Para comprobar la ganancia y mejora que se alcanza con los diseños realizados, se utilizaron diversas configuraciones y diferentes videos espectrales tanto sintéticos como reales. Las configuraciones abarcan la reconstrucción con el problema de reconstrucción tradicional de la Ecuación 2.12 y el propuesto en este trabajo, Ecuación 3.7.

### 4.1. Parámetros de simulación

Para la ejecución de las simulaciones se usaron cuatro videos espectrales  $\mathbf{F}$  con resolución espacial  $N = 128, N = 128$ . Los videos 1 y 2 corresponden a 2 secciones diferentes de un video espectral del estado del arte que contiene objetos en movimiento (Mian y Hartley, 2012). El video 3, es un video sintético compuesto por dos imágenes espectrales estáticas donde una de ellas realiza un desplazamiento sobre la otra (Yasuma, Mitsunaga, Iso, y Nayar, 2008). Por último, el video 4 es una secuencia de imágenes espectrales adquirida en el laboratorio de óptica del grupo de investigación en Diseño de Algoritmos y Procesamiento de Datos Multidimensionales (HDSP) de la Universidad Industrial de Santander. Todos los conjuntos de datos fueron adquiridos con una cámara CCD y un filtro sintonizable en longitudes de onda desde 400nm hasta 700nm en pasos de 10nm. La Figura 15 muestra la representación RGB de ocho fotogramas de cada video espectral. Para las simulaciones, se usaron 8 fotogramas espectrales y 8 bandas de las 31 bandas de la dimensión espectral. Las bandas seleccionadas corresponden a las longitudes de onda  $\lambda_0 = 430nm$ ,  $\lambda_1 = 460nm$ ,  $\lambda_2 = 490nm$ ,  $\lambda_3 = 520nm$ ,  $\lambda_4 = 550nm$ ,  $\lambda_5 = 580nm$ ,  $\lambda_6 = 610nm$  y  $\lambda_7 = 640nm$ .

Con el objetivo de comparar los diseños obtenidos con el algoritmo de propuesto, se generaron dos ensamblajes adicionales al utilizado para probar los T-CCA: el primero, corresponde a un código de apertura de color óptimo con filtros pasa baja y pasa alta ( $\mathcal{LH}$ -CCA del inglés *Low-High pass Colored Coded Apertures*). El segundo, corresponde a una realización con filtros pasa baja y pasa alta ordenados aleatoriamente ( $\mathcal{LH}$ -random CCA del inglés *Low-High pass filters random Colored Coded Apertures*), ambas aperturas propuestas en (Arguello y Arce, 2013).

Para cada apertura de codificación:  $\mathcal{LH}$ -CCA,  $\mathcal{LH}$ -random CCA y, los propuestos en este trabajo, T-CCA; para cada video espectral de la Fig.15, se simuló el proceso de adquisición y reconstrucción. Las reconstrucciones fueron realizadas usando el problema inverso tradicional presentado en la Ecuación 2.12 y el problema inverso propuesto, que tiene en cuenta la correlación temporal para la reconstrucción, presentado en la Ecuación 3.7. El algoritmo de reconstrucción que se utilizó y modificó fue el algoritmo proyección de gradiente para la reconstrucción escasa (GPSR del inglés *Gradient Projection for Sparse Reconstruction*) con 500 iteraciones (Figueiredo, Nowak, y Wright, 2007) para la reconstrucción del video en alta resolución; para la reconstrucción del video en baja resolución se usaron 30 iteraciones del algoritmo Lagrangiano Aumentado de Separación (C-SALSA del inglés *Constrained Split Augmented Lagrangian Shrinkage Algorithm*) (Afonso, Bioucas-Dias, y Figueiredo, 2011). La base de representación seleccionada para la representación *escasa* de los videos espectrales

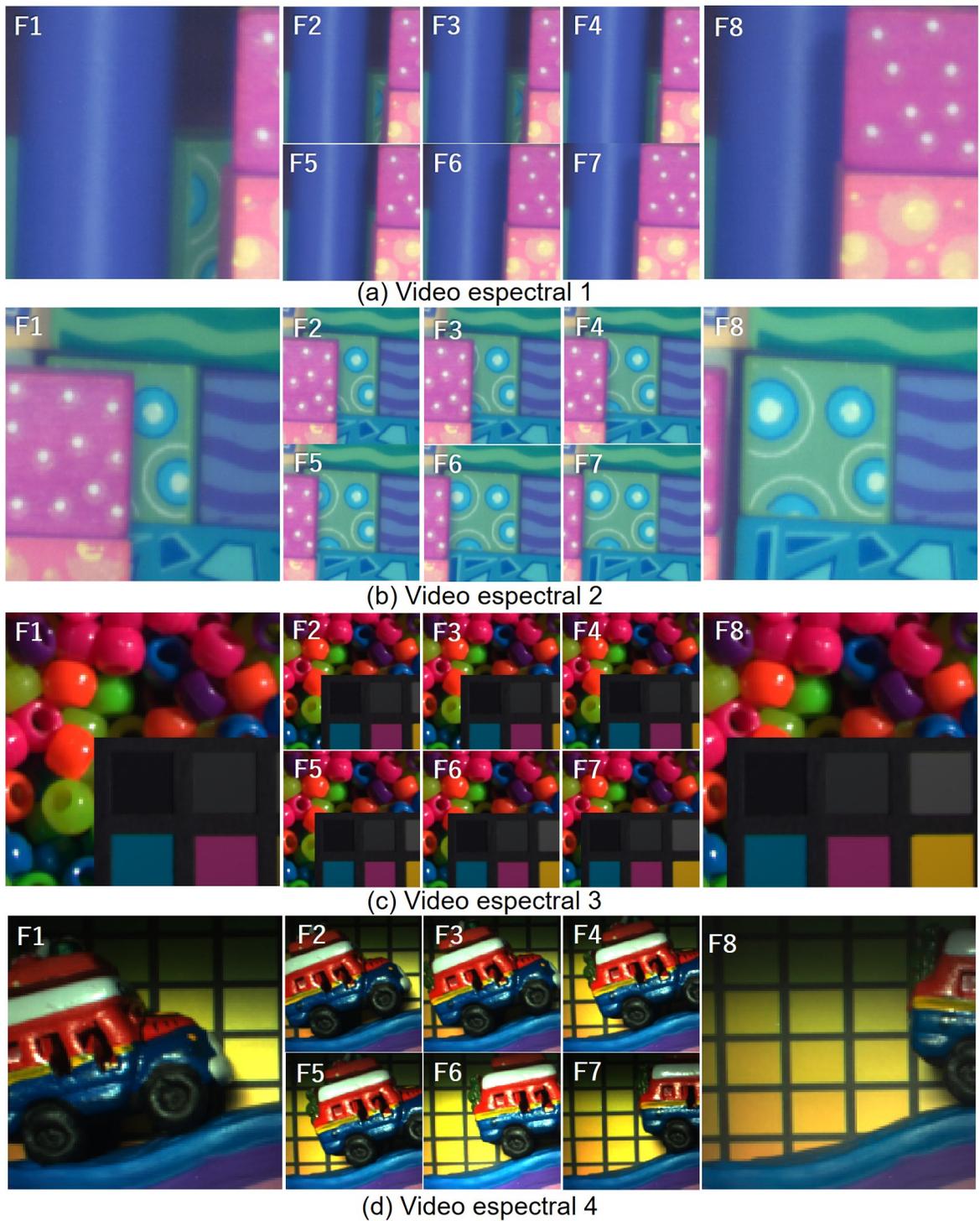


Figura 15: Representación RGB de ocho fotogramas de los videos espectrales empleados para las simulaciones. La resolución de cada video es de  $128 \times 128$  píxeles espaciales, 8 bandas espectrales y 8 fotogramas.

fue la WWDD de la Tabla 1. Las métricas para cuantificar la calidad de reconstrucción son el PSNR y la similitud estructural (SSIM), donde el PSNR presenta un valor alto (mayor a 30dB) para imágenes con gran similitud a la original y el SSIM presenta un valor máximo de 1 al ser comparado contra la misma imagen (Z. Wang, Bovik, Sheikh, y Simoncelli, 2004).

## 4.2. Reconstrucción usando el modelo de reconstrucción tradicional

La Figura 16 presenta la reconstrucción en RGB usando el problema inverso tradicional con los tres códigos de apertura para el primer fotograma espectral del video 1 con su respectiva evaluación de calidad de reconstrucción. El PSNR que se presenta equivale al valor promedio de PSNR evaluado para cada banda espectral del fotograma. Como se muestra en la figura, la reconstrucción de una señal muestrada con las aperturas de codificación propuesta presenta una mejor reconstrucción con respecto a las otras aperturas. La Figura 17 presenta la representación en falso color de la reconstrucción del fotograma 1 del video 2. En la sección ampliada, es posible notar la calidad de reconstrucción espacial. En la parte inferior se muestra el PSNR promedio, donde se alcanza una mayor PSNR realizando el muestreo con una apertura de codificación T-CCA.

Para resumir y destacar la mejora en reconstrucción realizando el muestreo con las aperturas propuestas, en la Tabla 2 se presenta la calidad de reconstrucción promedio en PSNR y SSIM para cada video reconstruido.

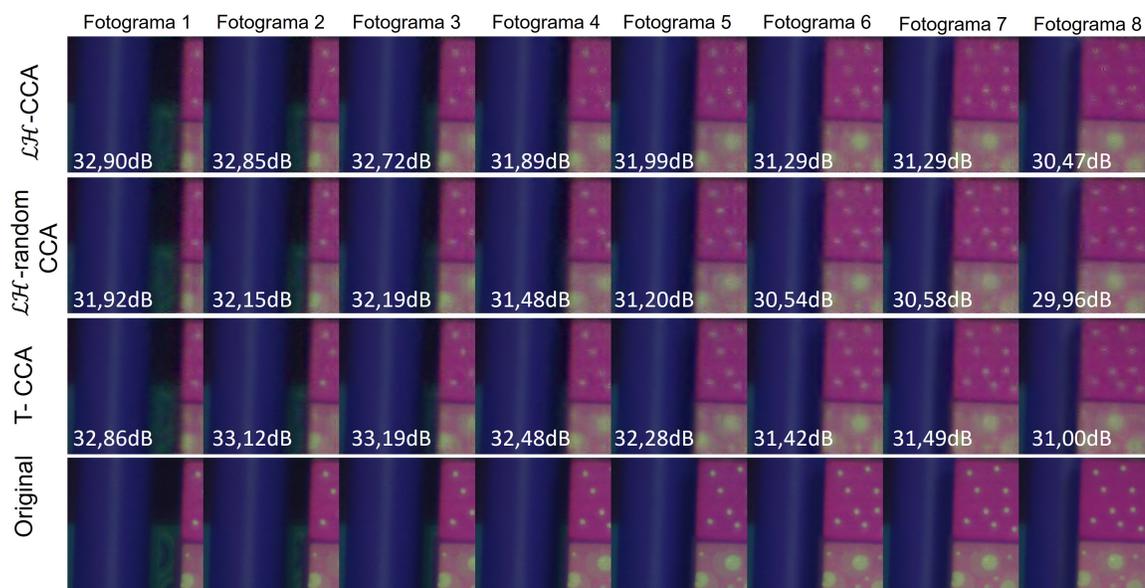


Figura 16: Representación en RGB de la reconstrucción del video 1 muestreado con las diferentes aperturas:  $\mathcal{LH}$ -CCA (primera fila),  $\mathcal{LH}$ -randomCCA (segunda fila), T-CCA (tercera fila) y la señal original (cuarta fila). En las columnas se tiene la secuencia de fotogramas que cambian en el tiempo.

Tabla 2: Promedio de las métricas PSNR y SSIM para cada reconstrucción con base en el problema inverso tradicional

<i>Apertura</i>	PSNR (dB)			SSIM		
	LH-CCA	LH-Random CCA	T-CCA	LH-CCA	LH-Random CCA	T-CCA
<b>Video 1</b>	32,17	31,46	<b>32,50</b>	0,96	0,96	0,97
<b>Video 2</b>	31,60	31,60	<b>32,72</b>	0,94	0,94	0,95
<b>Video 3</b>	26,68	26,41	<b>27,98</b>	0,89	0,87	0,89
<b>Video 4</b>	30,07	30,20	<b>30,58</b>	0,81	0,81	0,83

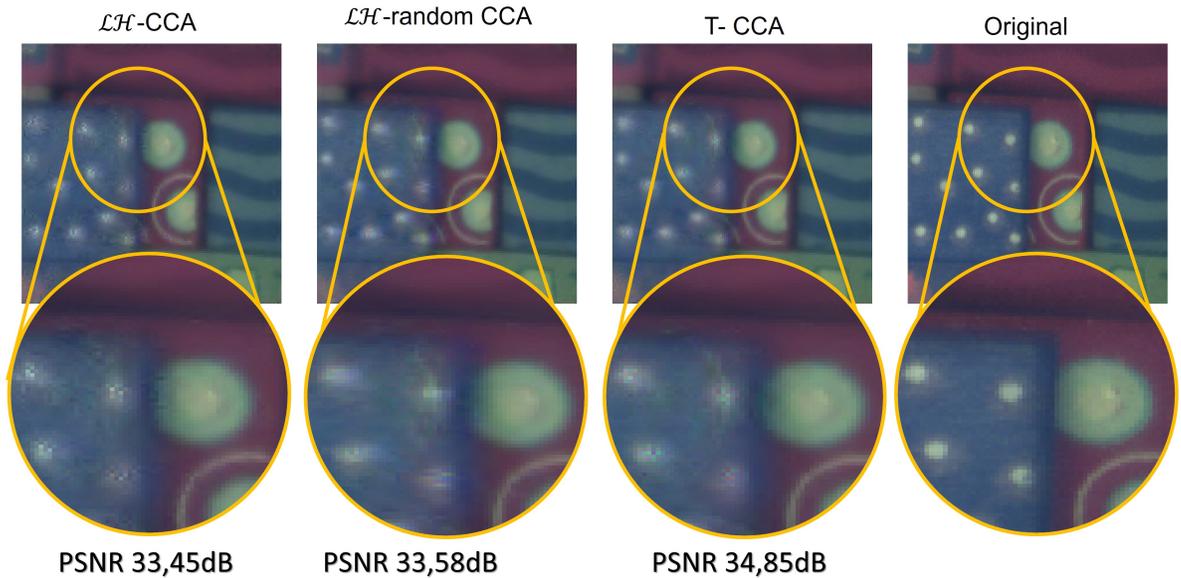


Figura 17: Representación en RGB de la reconstrucción del fotograma 1 del video 2. La sección ampliada muestra la calidad de reconstrucción espacial alcanzada cuando se muestrea la señal con cada apertura de codificación.

### 4.3. Enfoque de reconstrucción en múltiples resoluciones

Usando el problema inverso propuesto (Ec. 3.7), donde se agrega la correlación temporal del video al problema de reconstrucción, se realiza la reconstrucción de los videos de prueba obteniendo las medidas con las diferentes aperturas de codificación ya descritas. La Figura 18 presenta la reconstrucción de la banda espectral 3 para 8 fotogramas consecutivos del video 1 con su respectiva calidad de reconstrucción. Asimismo, la Figura 19 presenta la reconstrucción de la banda espectral 3 de 8 fotogramas consecutivos del video 2. Se puede observar que las reconstrucciones de las medidas tomadas con las aperturas T-CCA presentan una mejora con respecto a las otras aperturas. La Figura 20 ilustra la representación RGB del fotograma 5 para el video 3 y el promedio de la calidad de PSNR a lo largo de la dimensión espectral;

donde la sección ampliada y el PSNR muestra la calidad espacio-espectral que se alcanza con las T-CCA con una ganancia de hasta 4dB.

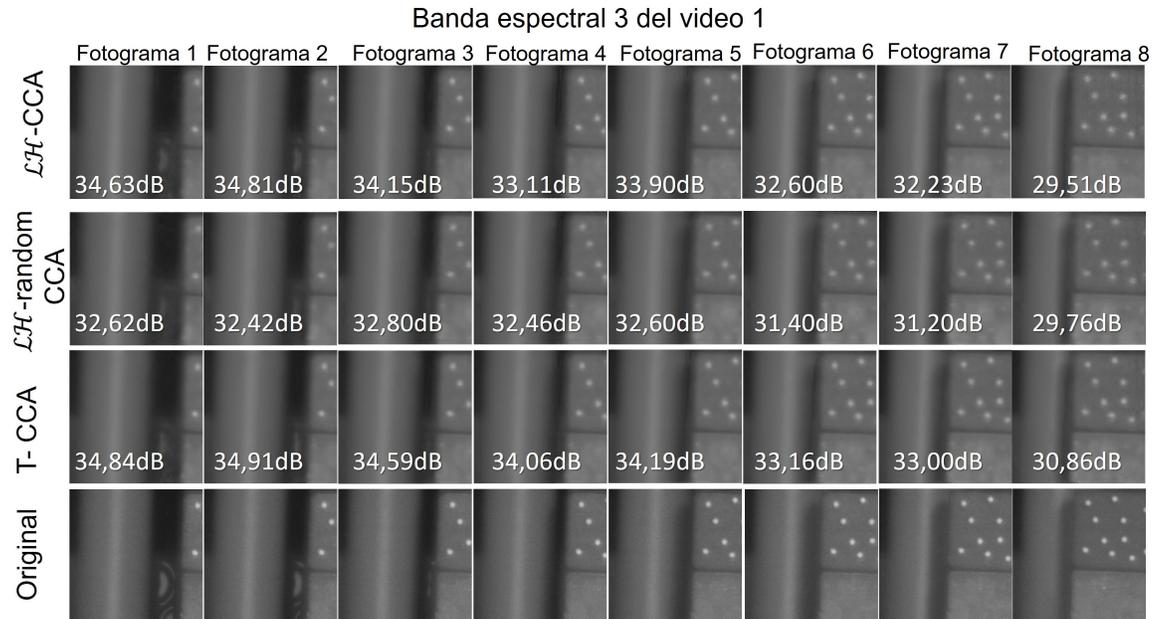


Figura 18: Visualización de la reconstrucción de la tercera banda espectral en 8 fotogramas del video 1. La calidad en términos de PSNR se presenta para cada reconstrucción.

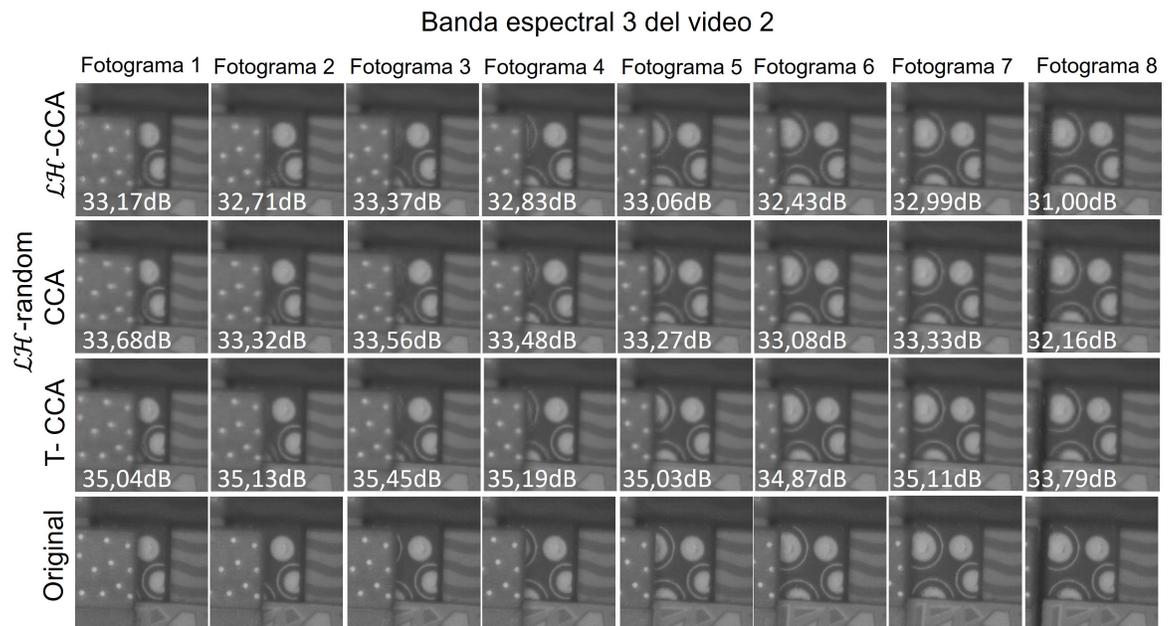


Figura 19: Visualización de la reconstrucción de la tercera banda espectral en 8 fotogramas del video 2. La calidad en términos de PSNR se presenta para cada reconstrucción.

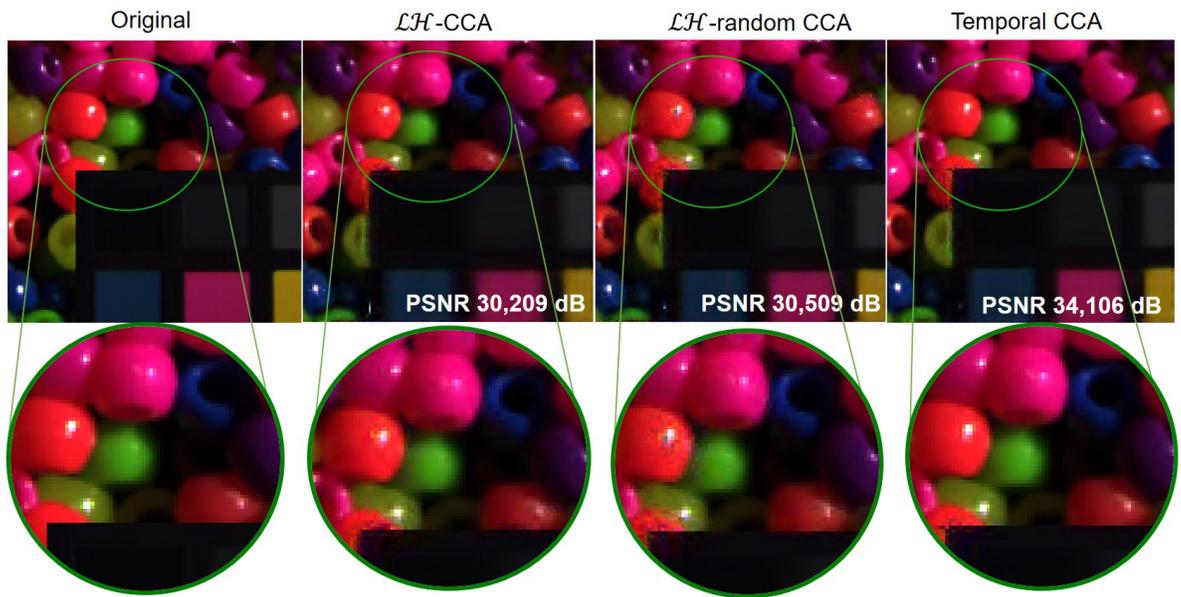


Figura 20: Representación RGB del fotograma 5, original y reconstruido, del video 3. La sección enfocada muestra que hay espacialmente una mejora usando la apertura T-CCA, donde se obtiene hasta 4dB de ganancia en calidad de reconstrucción con respecto a las otras aperturas.

Para evaluar la calidad espectral de las reconstrucciones se toman dos fotogramas del video 2 en el cual se escogen dos puntos espaciales: P1 del fotograma 1 y P2 del fotograma 8, como se ilustra en la parte superior de la Fig. 21. En estos dos puntos, se toma el píxel espectralmente, estos es, la respuesta espectral que tiene cada banda en dicho punto, y se grafica formando las curvas que se muestra en la parte inferior de la Figura 21. Observe que la curva verde corresponde a la respuesta espectral del fotograma original, y las otras curvas representan las reconstrucciones de la señal muestreada con las diferentes aperturas de codificación. Note que la curva más cercana a la original es la que corresponde a T-CCA, para los dos puntos seleccionados. Dado que las T-CCA tienen una modulación espectral más significativa, la reconstrucción de la señal es más precisa con respecto a las otras modulaciones. De igual forma, en la Fig.22 se presenta el mismo análisis espectral para la reconstrucción del video 4 en los puntos: P1 del fotograma 1 y P2 del fotograma 6. A modo de ilustración, se muestra la reconstrucción de baja de las medidas tomadas con las aperturas T-CCA en la Fig. 23.

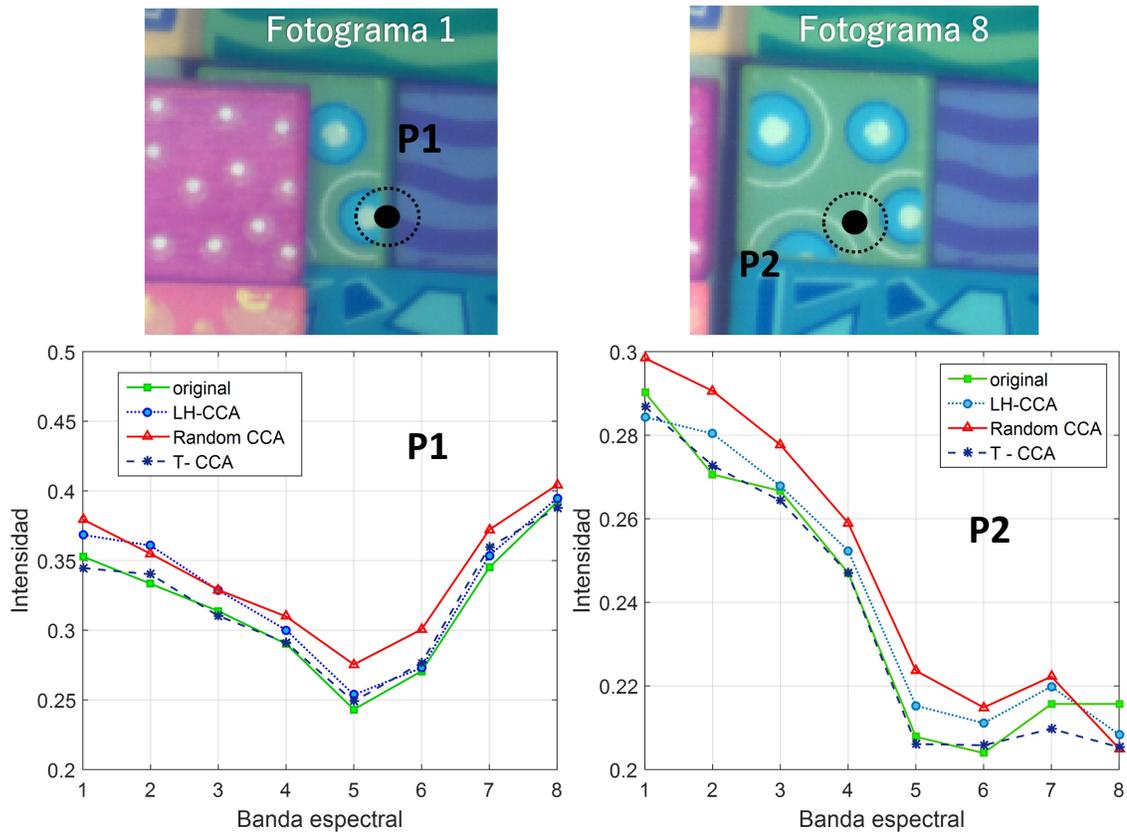


Figura 21: Comparación de la respuesta espectral (pixel espectral) para los puntos P1 (fotograma 1) y P2 (fotograma 8) del video espectral 2 usando las aperturas  $\mathcal{LH}$ -CCA,  $\mathcal{LH}$ -random CCA y T-CCA. Note que con las T-CCA se obtiene una mejora aproximación espectral (aproximación a la curva verde) con respecto a las otras aperturas.

El resumen de los resultados de las reconstrucciones para los 4 videos espectrales son presentados en la Tabla 3. La parte superior de la tabla muestra la calidad de reconstrucción del video en alta resolución y la parte inferior, muestra la calidad de reconstrucción de la versión de baja resolución para cada video. Note que las medidas tomadas con las aperturas T-CCA presentan una mejor calidad de reconstrucción tanto en alta como en baja resolución, donde se obtiene un PSNR alto y un SSIM muy cercano a 1.

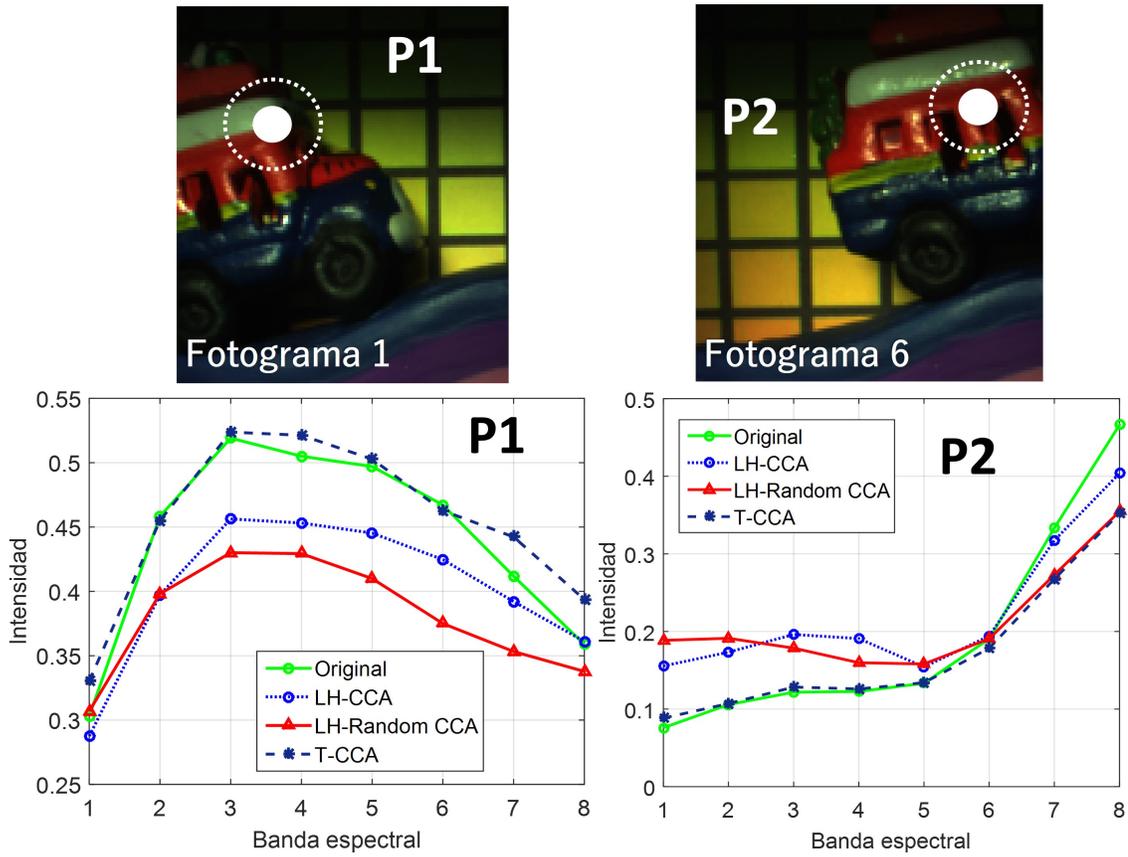


Figura 22: Comparación de la respuesta espectral (píxel espectral) para los puntos P1 (fotograma 1) y P2 (fotograma 6) del video espectral 4 usando las aperturas  $\mathcal{LH}$ -CCA,  $\mathcal{LH}$ -random CCA y T-CCA.

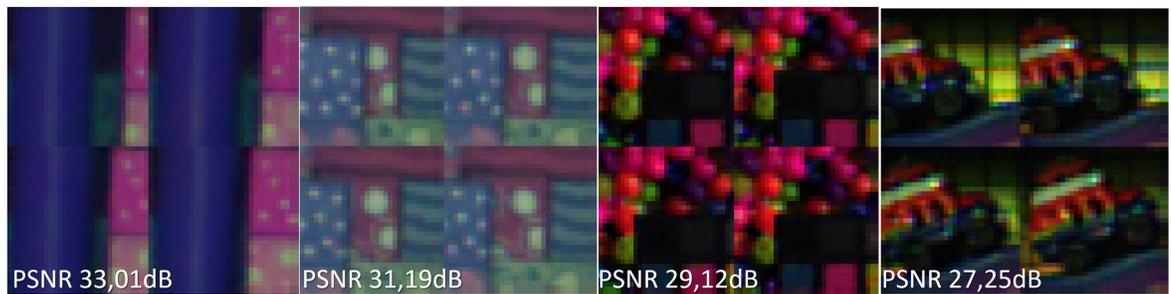


Figura 23: Representación RGB de la reconstrucción de baja resolución de los primeros 4 fotogramas para los 4 videos de prueba con base en las medidas tomadas desde la apertura T-CCA. La calidad promedio en términos de PSNR se presenta para cada reconstrucción.

Tabla 3: Promedio de las métricas PSNR y SSIM para cada reconstrucción con base en el problema inverso propuesto

Reconstrucción de alta resolución						
	PSNR (dB)			SSIM		
<i>Apertura</i>	$\mathcal{LH}$ -CCA	$\mathcal{LH}$ -random CCA	T-CCA	$\mathcal{LH}$ -CCA	$\mathcal{LH}$ -random CCA	T-CCA
<b>Video 1</b>	34,51	34,70	<b>35,62</b>	0,980	0,962	<b>0,982</b>
<b>Video 2</b>	33,45	33,58	<b>34,86</b>	0,962	0,959	<b>0,970</b>
<b>Video 3</b>	30,52	30,09	<b>33,24</b>	0,938	0,944	<b>0,966</b>
<b>Video 4</b>	31,17	31,06	<b>32,48</b>	0,875	0,858	<b>0,894</b>
Reconstrucción de baja resolución						
	PSNR (dB)			SSIM		
<i>Apertura</i>	$\mathcal{LH}$ -CCA	$\mathcal{LH}$ -random CCA	T-CCA	$\mathcal{LH}$ -CCA	$\mathcal{LH}$ -random CCA	T-CCA
<b>Video 1</b>	30,89	31,42	<b>33,01</b>	0,959	0,965	<b>0,974</b>
<b>Video 2</b>	29,78	27,47	<b>31,19</b>	0,949	0,920	<b>0,961</b>
<b>Video 3</b>	28,87	27,84	<b>29,12</b>	0,939	0,928	<b>0,942</b>
<b>Video 4</b>	27,835	25,46	27,251	0,932	0,897	0,921

## 5. Conclusiones, Contribuciones y Trabajo futuro

### 5.1. Conclusiones

Se presentó un algoritmo que permite realizar el diseño de las aperturas de codificación para un sistema de adquisición compresiva de video espectral. Para el diseño de las aperturas de codificación, la simulación del sistema de adquisición compresivo de video espectral permitió realizar una caracterización de criterios y parámetros, con lo cual se estableció el algoritmo de diseño que considera la componente espacial, espectral y temporal para generar una apertura de codificación espacio-espectro-temporal. Adicionalmente, se presentó la modificación al problema de reconstrucción de video espectral tradicional donde se agrega la correlación temporal. Dicho problema de reconstrucción propuesto está basado en el enfoque de múltiples resoluciones, que permite agregar la correlación temporal del video con base en una versión de baja resolución para mejorar la calidad del video en alta resolución. El diseño de las aperturas propuesto permitió realizar una buena estimación de baja resolución, lo que implicó mejorar la calidad de reconstrucción del video en alta resolución con respecto a aperturas de codificación del estado del arte. Con base en las aperturas diseñadas, los resultados en simulaciones, para 4 videos espectrales con diferentes características, mostraron una calidad de reconstrucción de hasta 33dB con el problema de reconstrucción tradicional, y hasta 36dB con la reconstrucción de múltiples resoluciones. A su vez, se presentó una mejora en promedio de hasta 4dB en comparación con las aperturas de color optimizadas y aleatorias, realizadas con filtros ópticos pasa baja y pasa alta.

Adicional al trabajo realizado, se presentó un análisis de las posibles representaciones escasas para el video espectral con base en un trabajo previo del estado del arte. Esto, para establecer la mejor representación escasa de la señal 4D.

### 5.2. Contribuciones

Las contribuciones de este trabajo de investigación han sido publicadas en los siguientes artículos:

- Kareth M León, Laura Galvis, and Henry Arguello. Spectral dynamic scenes reconstruction based in compressive sensing using optical color filters. In *SPIE Commercial+ Scientific Sensing and Imaging*, pages 98600D–98600D. International Society for Optics and Photonics, 2016
- Kareth M León, Laura Galvis, and Henry Arguello. Spatio-spectro-temporal coded aperture design for multiresolution compressive spectral video sensing. *Submitted to EUSIPCO*, 2017.
- Kareth M León, Laura Galvis, and Henry Arguello. Coded aperture design and optical flow constrained reconstruction in compressive spectral video sensing. In *preparation to be submitted to IEEE Transactions on Geoscience and Remote Sensing*, 2017.

### **5.3. Trabajo futuro**

El trabajo desarrollado puede ser ampliado con un análisis profundo de la estimación del movimiento basado en otras técnicas. Un trabajo futuro, buscaría tener en cuenta otras formas de estimar el movimiento debido a que este parámetro podría permitir un calculo más preciso a partir otros métodos.

## Referencias

- Afonso, M. V., Bioucas-Dias, J. M., y Figueiredo, M. A. (2011). An augmented lagrangian approach to the constrained optimization formulation of imaging inverse problems. *IEEE Transactions on Image Processing*, 20(3), 681–695.
- Arce, G. R., Brady, D. J., Carin, L., Arguello, H., y Kittle, D. S. (2014). Compressive coded aperture spectral imaging: An introduction. *IEEE Signal Processing Magazine*, 31(1), 105–115.
- Arguello, H., y Arce, G. (2014). Colored coded aperture design by concentration of measure in compressive spectral imaging. *IEEE Transactions on Image Processing*, 23(4), 1896–1908. doi: 10.1109/TIP.2014.2310125
- Arguello, H., y Arce, G. R. (2011a). Code aperture optimization for spectrally agile compressive imaging. *Journal of the Optical Society of America A*, 28(11), 2400. doi: 10.1364/JOSAA.28.002400
- Arguello, H., y Arce, G. R. (2011b). Code aperture optimization for spectrally agile compressive imaging. *JOSA A*, 28(11), 2400–2413.
- Arguello, H., y Arce, G. R. (2013). Rank minimization code aperture design for spectrally selective compressive imaging. *IEEE Transactions on Image Processing*, 22(3), 941–954. doi: 10.1109/TIP.2012.2222899
- Arnold, T., De Biasio, M., y Leitner, R. (2013). Hyper-spectral video endoscopy system for intra-surgery tissue classification. En *Proceedings of the international conference on sensing technology, icst* (pp. 145–150). doi: 10.1109/ICSensT.2013.6727632
- Baker, S., Scharstein, D., Lewis, J., Roth, S., Black, M. J., y Szeliski, R. (2011). A database and evaluation methodology for optical flow. *International Journal of Computer Vision*, 92(1), 1–31.
- Banerjee, A., Burlina, P., y Broadwater, J. (2009). Hyperspectral video for illumination-invariant tracking. En *Whispers '09 - 1st workshop on hyperspectral image and signal processing: Evolution in remote sensing*. doi: 10.1109/WHISPERS.2009.5289103
- Benezeth, Y., Sidibé, D., y Thomas, J.-B. (2014). Background subtraction with multispectral video sequences. En *Ieee international conference on robotics and automation workshop on non-classical cameras, camera networks and omnidirectional vision (omnivis)* (pp. 6–p).
- Candès, E. J. (2006). Compressive sampling. *Proceedings of the International Congress of Mathematicians*, 1433–1452. doi: 10.4171/022-3/69
- Candès, E. J., y Wakin, M. (2008). An Introduction To Compressive Sampling. *IEEE Signal Processing Magazine*, 25(2), 21–30. doi: 10.1109/MSP.2007.914731
- Cheng, S. Y., Park, S., y Trivedi, M. M. (2007). Multi-spectral and multi-perspective video arrays for driver body tracking and activity analysis. *Computer Vision and Image Understanding*, 106(2-3), 245–257. doi: 10.1016/j.cviu.2006.08.010
- Correa, C., Galvis, D., y Arguello, H. (2016). Sparse representations of dynamic scenes for compressive spectral video sensing. *Dyna*, 83(195), 42.
- Correa, C. V., Arguello, H., y Arce, G. R. (2015). Snapshot colored compressive spectral imager. *Journal of the Optical Society of America A*, 32(10), 1754–1763. doi: 1084-7529/15/101754
- Correa, C. V., Arguello, H., y Arce, G. R. (2016). Spatiotemporal blue noise coded aperture

- design for multi-shot compressive spectral imaging. *JOSA A*, 33(12), 2312–2322.
- Eldar, Y. C., y Kutyniok, G. (2012). *Compressed sensing: theory and applications*. Cambridge University Press.
- Figueiredo, M. A. T., Nowak, R. D., y Wright, S. J. (2007). Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems. *IEEE Journal on Selected Topics in Signal Processing*, 1(4), 586–597. doi: 10.1109/JS-TSP.2007.910281
- Gehm, M., Kim, M., Fernandez, C., y Brady, D. (2008). High-throughput, multiplexed pushbroom hyperspectral microscopy. *Optics express*, 16(15), 11032–11043.
- Goldstein, T., Xu, L., Kelly, K. F., y Baraniuk, R. (2015). The stone transform: Multi-resolution image enhancement and compressive video. *IEEE Transactions on Image Processing*, 24(12), 5581–5593.
- Green, R. O., Eastwood, M. L., Sarture, C. M., Chrien, T. G., Aronsson, M., Chippendale, B. J., . . . others (1998). Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (aviris). *Remote Sensing of Environment*, 65(3), 227–248.
- Gribonval, R., y Nielsen, M. (2003). Sparse representations in unions of bases. *IEEE transactions on Information theory*, 49(12), 3320–3325.
- Horn, B. K., y Schunck, B. G. (1981). Determining optical flow. *Artificial intelligence*, 17(1-3), 185–203.
- Leitner, R., Biasio, M. D., Arnold, T., Dinh, C. V., Loog, M., y Duin, R. P. W. (2013). Multi-spectral video endoscopy system for the detection of cancerous tissue. *Pattern Recognition Letters*, 34(1), 85–93. doi: 10.1016/j.patrec.2012.07.020
- León, K. M., Galvis, L., y Arguello, H. (2016). Spectral dynamic scenes reconstruction based in compressive sensing using optical color filters. En *Spie commercial+ scientific sensing and imaging* (pp. 98600D–98600D).
- Lin, X., Liu, Y., Wu, J., y Dai, Q. (2014). Spatial-spectral encoded compressive hyperspectral imaging. *ACM Trans. Graph.*, 33(6), 233:244. Descargado de <http://doi.acm.org/10.1145/2661229.2661262> doi: 10.1145/2661229.2661262
- Liu, C. (2009). *Beyond pixels: exploring new representations and applications for motion analysis* (Tesis Doctoral no publicada). Citeseer.
- Ma, C., Cao, X., Tong, X., Dai, Q., y Lin, S. (2014). Acquisition of high spatial and spectral resolution video with a hybrid camera system. *International Journal of Computer Vision*, 110(2), 141–155.
- Ma, C., Cao, X., Wu, R., y Dai, Q. (2014). Content-adaptive high-resolution hyperspectral video acquisition with a hybrid camera system. *Optics letters*, 39(4), 937–40. doi: 10.1364/OL.39.000937
- Manolakis, D., Lockwood, R., y Cooley, T. (2008). On the spectral correlation structure of hyperspectral imaging data. En *Geoscience and remote sensing symposium, 2008. igarss 2008. ieee international* (Vol. 2, pp. II–581).
- Marwah, K., Wetzstein, G., Bando, Y., y Raskar, R. (2013). Compressive light field photography using overcomplete dictionaries and optimized projections. *ACM Transactions on Graphics (TOG)*, 32(4), 46.
- Mian, A., y Hartley, R. (2012). Hyperspectral video restoration using optical flow and sparse coding. *Optics express*, 20(10), 10658–10673.
- Qaisar, S., Bilal, R. M., Iqbal, W., Naureen, M., y Lee, S. (2013). *Compressive sensing: From theory to applications, a survey* (Vol. 15) (n.º 5). doi: 10.1109/JCN.2013.000083

- Ramirez, A. B., Arguello, H., y Arce, G. (2011). Video anomaly recovery from compressed spectral imaging. En *Icassp, ieee international conference on acoustics, speech and signal processing - proceedings* (pp. 1321–1324). doi: 10.1109/ICASSP.2011.5946655
- Rauhut, H. (2010). Compressive sensing and structured random matrices. *Theoretical foundations and numerical methods for sparse recovery*, 9, 1–92.
- Reddy, D., Veeraraghavan, A., y Chellappa, R. (2011). P2c2: Programmable pixel compressive camera for high speed imaging. En *Computer vision and pattern recognition (cvpr), 2011 ieee conference on* (pp. 329–336).
- Rueda, H., Arguello, H., y Arce, G. R. (2014). Compressive spectral imaging based on colored coded apertures. En *Acoustics, speech and signal processing (icassp), 2014 ieee international conference on* (pp. 7799–7803).
- Rueda, H., Arguello, H., y Arce, G. R. (2015). Dmd-based implementation of patterned optical filter arrays for compressive spectral imaging. *JOSA A*, 32(1), 80–89.
- Sankaranarayanan, A. C., Turaga, P. K., Chellappa, R., y Baraniuk, R. G. (2013). Compressive Acquisition of Linear Dynamical Systems. *Siam Journal on Imaging Sciences*, 6(4), 2109–2133. Descargado de <Go to ISI>://WOS:000328890700010\http://epubs.siam.org/doi/pdf/10.1137/120863307 doi: Doi 10.1137/120863307
- Sankaranarayanan, A. C., Xu, L., Studer, C., Li, Y., Kelly, K. F., y Baraniuk, R. G. (2015). Video compressive sensing for spatial multiplexing cameras using motion-flow models. *SIAM Journal on Imaging Sciences*, 8(3), 1489–1518.
- Sloane, N. J., y Harwit, M. (1976). Masks for Hadamard transform optics, and weighing designs. *Applied optics*, 15(1), 107–114. doi: 10.1364/AO.15.000107
- Smith, W. L., Zhou, D. K., Harrison, F. W., Revercomb, H. E., Larar, A. M., Huang, H.-L., y Huang, B. (2001). Hyperspectral remote sensing of atmospheric profiles from satellites and aircraft. En *Second international asia-pacific symposium on remote sensing of the atmosphere, environment, and space* (pp. 94–102).
- Thapa, D., Raahemifar, K., y Lakshminarayanan, V. (2015). Less is more: compressive sensing in optics and image science. *Journal of Modern Optics*, 62(6), 415–429.
- Tsai, T.-H., Llull, P., Yuan, X., Carin, L., y Brady, D. J. (2015). Coded Aperture Compressive Spectral-Temporal Imaging. *Imaging and Applied Optics 2015*(2), CTh2E.5. Descargado de <http://www.osapublishing.org/abstract.cfm?uri=COSI-2015-CTh2E.5> doi: 10.1364/COSI.2015.CTh2E.5
- Wagadarikar, A., John, R., Willett, R., y Brady, D. (2008). Single disperser design for coded aperture snapshot spectral imaging. *Applied optics*, 47, B44–B51. doi: 10.1364/AO.47.000B44
- Wagadarikar, A. A., Pitsianis, N. P., Sun, X., y Brady, D. J. (2009). Video rate spectral imaging using a coded aperture snapshot spectral imager. *Optics Express*, 17(8), 6368–6388.
- Wang, L., Xiong, Z., Gao, D., Shi, G., Zeng, W., y Wu, F. (2015). High-speed Hyperspectral Video Acquisition with a Dual-camera Architecture. *Cvpr*. doi: 10.1109/CVPR.2015.7299128
- Wang, Z., Bovik, A. C., Sheikh, H. R., y Simoncelli, E. P. (2004). Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4), 600–612. doi: 10.1109/TIP.2003.819861
- Willett, R. M., Gehm, M. E., y Brady, D. J. (2007). Multiscale recons-

- truction for computational spectral imaging. En *5th, computational imaging conference spie* (Vol. 6498, pp. 64980L–64980L–15). Descargado de <http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=1336211>  
doi: 10.1117/12.715711
- Wu, Y., Mirza, I. O., Arce, G. R., y Prather, D. W. (2011). Development of a digital-micromirror-device-based multishot snapshot spectral imaging system. *Optics letters*, *36*(14), 2692–2694. doi: 10.1364/OL.36.002692
- Yasuma, F., Mitsunaga, T., Iso, D., y Nayar, S. K. (2008). *CAVE Projects: Multispectral Image Database*. Descargado de <http://www.cs.columbia.edu/CAVE/databases/multispectral/>