

**ESTIMACIÓN DE LA MATRIZ DE CORRELACIÓN  
CRUZADA PARA DETECCIÓN DE OBJETIVOS  
USANDO UN ALGORITMO DE PROCESAMIENTO  
DE IMÁGENES ADQUIRIDAS CON UN SISTEMA  
COMPRESIVO**

KEVIN ADEMIR ARIAS ROJAS

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

2018

# **ESTIMACIÓN DE LA MATRIZ DE CORRELACIÓN CRUZADA PARA DETECCIÓN DE OBJETIVOS USANDO UN ALGORITMO DE PROCESAMIENTO DE IMÁGENES ADQUIRIDAS CON UN SISTEMA COMPRESIVO**

KEVIN ADEMIR ARIAS ROJAS

Trabajo de grado para optar por el título de  
Ingeniero de Sistemas

Director:

TATIANA CAROLINA GELVEZ BARRERA

*PhD.(c) en Ingeniería*

Codirector:

HENRY ARGUELLO FUENTES

*Ph.D en Ingeniería Eléctrica y computación*

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

2018

---

# DEDICATORIA

*Dedico este libro a Dios, a mi familia y a las personas allegadas*

*A Dios, por darme la fuerza y el coraje necesario para seguir adelante independientemente de los obstáculos en el camino. Además, por haberme hecho una persona humilde y dedicada enseñándome que las cosas especiales de la vida requieren esfuerzo.*

*A mis padres, quienes me han apoyado incondicionalmente a lo largo del trabajo realizado y representan mi mayor inspiración. La confianza que han brindado en mí, hoy se refleja en un sueño hecho realidad y los logros obtenidos reflejan su amor, compromiso y dedicación hacia mí.*

*A mis hermanos, por su fiel compañía y sus palabras de motivación. Todo el esfuerzo requerido fue inspirado en su apoyo incondicional.*

*A Dayanna, por brindarme su compañía en los momentos difíciles. Su colaboración y entrega fue de gran ayuda para culminar mis metas.*

---

## AGRADECIMIENTOS

Gracias a Dios, a mi familia y a todas las personas que contribuyeron en esta etapa de formación como estudiante.

Gracias a mi mama ELDA ROJAS JAIMES por el amor y el esfuerzo que dedico hacia mí. Por la compañía y el apoyo brindado en todo momento, sus consejos me dieron fortaleza y contribuyeron a la formación de una persona con principios y valores dispuesta a afrontar los retos de la vida con el objetivo de cumplir sus metas.

Gracias a mi papa PEDRO ALFONSO ARIAS CARVAJAL quien me ha acompañado en todos los proyectos que he emprendido, su compañía ha sido de gran importancia para culminar con éxito mis sueños.

Gracias a ustedes padres pude lograr el sueño de ser ingeniero.

Gracias a GERSSON hermano JAIR ARIAS ROJAS por cuidarme y enseñarme la importancia de la disciplina.

Gracias a mi hermano EMERSON ALFONSO ARIAS ROJAS por enseñarme que la madurez y los principios son muy importante en la búsqueda de una mejor sociedad.

Gracias a DAYANNA FUENTES MARTINEZ, por la compañía y las palabras de aliento, su actitud fue de gran ayuda en los momentos de angustia y desmayo.

Gracias a mi directora TATIANA CAROLINA GELVEZ BARRERA, por su paciencia, colaboración y profesionalidad en el desarrollo de mi proceso académico.

Gracias a mi codirector HENRY ARGUELLO FUENTES, por los consejos, la paciencia y el apoyo brindado.

Al grupo HDSP quien acompaño y respaldo en gran parte la etapa de formación en la universidad como estudiante.

A todos mis amigos, por los momentos de alegría compartidos.

# CONTENIDO

<b>INTRODUCCIÓN</b>	<b>14</b>
<b>1 MARCO TEÓRICO</b>	<b>17</b>
1.1 DEFINICIÓN Y CLASIFICACIÓN DEL RECONOCIMIENTO DE PATRONES . . . . .	17
1.2 CORRELACIÓN CRUZADA ENTRE IMÁGENES . . . . .	18
1.3 TEORÍA DE ADQUISICIÓN COMPRESIVA. . . . .	23
1.3.1 Modelo matemático de adquisición compresiva . . . . .	23
1.3.2 Escasez . . . . .	24
1.3.3 Incoherencia . . . . .	24
1.3.4 Modelo de adquisición compresiva sobre una imagen a escala de grises . . . . .	25
<b>2 METODOLOGÍA</b>	<b>27</b>
2.1 PLANTEAMIENTO MATEMÁTICO . . . . .	27
2.2 ELECCIÓN DE UN FILTRO DE CORRELACIÓN . . . . .	28
2.3 ALGORITMO PROPUESTO . . . . .	32
<b>3 MATERIALES Y MÉTODOS.</b>	<b>35</b>
3.1 MONTAJE EXPERIMENTAL PARA LA ADQUISICIÓN DE UN CONJUNTO DE DATOS REAL . . . . .	35
3.2 CONJUNTO DE DATOS SINTÉTICO . . . . .	36
3.3 EMPAREJAMIENTO DE HISTOGRAMAS . . . . .	37
3.4 MÉTRICAS DE RENDIMIENTO . . . . .	39
3.4.1 Relación pico señal a ruido . . . . .	40
3.4.2 Error de seguimiento . . . . .	40
3.4.3 Tasa de éxito . . . . .	41
<b>4 SIMULACIONES Y RESULTADOS</b>	<b>42</b>
4.1 DESCRIPCIÓN DEL CONJUNTO DE DATOS DE PRUEBA . . . . .	42
4.2 RESULTADOS VISUALES . . . . .	43
4.3 ANÁLISIS DE LAS MÉTRICAS DE CALIDAD . . . . .	48
4.3.1 Resultados de error de seguimiento . . . . .	48
4.3.2 Resultados de tasa de éxito . . . . .	49
4.3.3 Resultados de tiempo de ejecución . . . . .	50

---

4.3.4	Resultados de PSNR . . . . .	50
4.3.5	Resultados de comparación . . . . .	51
<b>5</b>	<b>CONCLUSIONES</b>	<b>52</b>
	<b>REFERENCIAS</b>	<b>53</b>
	<b>BIBLIOGRAFÍA</b>	<b>56</b>

## LISTA DE FIGURAS

Figura 1	Reconocimiento de patrones basado en la extracción de características. . . . .	17
Figura 2	Reconocimiento de patrones basado en la correspondencia de patrones. . . . .	18
Figura 3	Cálculo de la correlación cruzada entre dos imágenes. . . . .	20
Figura 4	Adquisición compresiva de una imagen . . . . .	25
Figura 5	Invarianza a la rotación incluyendo filtros armónicos circulares. . . . .	29
Figura 6	Estimación de la correlación para la detección del objetivo pez. . . . .	30
Figura 7	Estimación de la correlación para la detección del objetivo tortuga. . . . .	31
Figura 8	Estimación de la correlación para la detección del objetivo cangrejo. . . . .	31
Figura 9	Montaje experimental para obtener los videos de prueba. . . . .	35
Figura 10	Diseño de los datos sintéticos. . . . .	36
Figura 11	Sustracción de la media para la detección de objetivos. . . . .	38
Figura 12	Emparejamiento de histogramas para la detección de objetivos. . . . .	39
Figura 13	Vehículos modelo usados como patrón de referencia a lo largo de sus CHF. . . . .	42
Figura 14	Conjunto de datos de prueba video 1. . . . .	43
Figura 15	Conjunto de datos de prueba video 2. . . . .	44
Figura 16	Conjunto de datos de prueba video 3. . . . .	45
Figura 17	Conjunto de datos de prueba video 4. . . . .	46
Figura 18	Conjunto de datos de prueba video 5. . . . .	47
Figura 19	Evaluación de la métrica error de seguimiento usando el método propuesto. . . . .	48
Figura 20	Análisis de la métrica tasa de éxito usando el método propuesto. . . . .	49
Figura 21	Complejidad temporal del algoritmo de reconocimiento de patrones basado en compresión. . . . .	50
Figura 22	Análisis de la métrica PSNR sobre el método propuesto. . . . .	50

---

Figura 23 Comparación entre las métricas de calidad usando el método de correlación tradicional y propuesto. . . . .	51
--	----

## LISTA DE TABLAS

Tabla 1	Valores de discriminación obtenidos por los CHFs en la detección de los objetivos pez, tortuga y cangrejo . . . . .	32
---------	---	----

---

## RESUMEN

**TÍTULO:** " Estimación de la matriz de correlación cruzada para detección de objetivos usando un algoritmo de procesamiento de imágenes adquiridas con un sistemas compresivo".\*

**AUTOR:** Kevin Ademir Arias Rojas \*\*

**PALABRAS CLAVE:** Reconocimiento de patrones, detección de objetivos, matriz de correlación cruzada, filtros armónicos circulares, adquisición compresiva.

**DESCRIPCIÓN:** La correlación entre dos escenas puede ser modelada por una matriz que contiene medidas de similitud entre los pixeles espaciales de ambas imágenes. Por ende, los algoritmos de reconocimiento de patrones basados en la estimación de la matriz de correlación son útiles para aplicaciones de visión por computador como la detección de objetivos. Las limitaciones de almacenamiento y transmisión de datos cuando el tamaño de las escenas aumenta han previsto a los sistemas compresivos como el futuro de la adquisición de datos. Un sistema compresivo es aquel que permite adquirir y comprimir la escena simultáneamente. Actualmente, el cálculo de la matriz de correlación bajo la teoría de la adquisición compresiva requiere recuperar la escena a partir de las medidas comprimidas. No obstante, este procedimiento podría replantearse matemáticamente para obtener la matriz de correlación de forma directa. Con base en lo anterior, este trabajo de grado presenta un enfoque para estimar la matriz de correlación cruzada a partir de medidas comprimidas capturadas de una escena de entrada. Específicamente, se incluyó un conjunto de operadores lineales sobre el modelo matemático de adquisición para establecer adecuadamente la descomposición escasa de la escena y recuperar la matriz de correlación mediante un algoritmo de reconstrucción basado en escasez. El enfoque propuesto se implementó en un sistema de reconocimiento de patrones para el seguimiento de un objeto. Los resultados fueron comparados con respecto al cálculo de la correlación usando la imagen completa. El método propuesto logró recuperar la correlación con una calidad de hasta  $31[dB]$  en términos de PSNR y exhibió un error de seguimiento de  $6[px]$  reduciendo el porcentaje de datos adquiridos hasta en un 40 %.

---

\* Trabajo de Investigación.

\*\* Facultad de Ingenierías Fisicomecánicas. Escuela de Ingeniería de Electrónica. Director: PhD.(c) Tatiana Carolina Gelvez Barrera.

---

## ABSTRACT

**TITLE:** Cross-correlation estimation to target detection using a compressive image processing algorithm. \*

**AUTHOR:** Kevin Ademir Arias Rojas \*\* -0.3cm

**KEYWORDS:** Pattern recognition, target detection, cross-correlation matrix, circular harmonic filters, compressive sensing.

**DESCRIPTION:** Correlation between a reference pattern and an input scene can be modeled by a matrix structure whose elements contain similarity measures between spatial pixels of both images. Hence, pattern recognition algorithms based on the correlation matrix estimation are useful for some computer vision applications such as target detection. However, storage and transmission limitations of the data when the size of the scenes increases have foreseen compressive systems as the future of the data acquisition. A compressive system is one that allows you to acquire and compress the scene simultaneously through the hardware.

Currently, the correlation matrix calculation under the compressive sensing theory previously requires recovering the scene from the compressed measurements. Notwithstanding, this procedure could be mathematically reformulated to obtain the correlation in a direct way. Based on the above, this thesis work presents an approach to estimate the cross-correlation matrix from compressed measurements captured of an input scene. Specifically, a set of linear operators was included on the compressed acquisition mathematical model to properly establish the sparse decomposition of the scene and recover the correlation matrix by means of a sparse-based reconstruction algorithm. The proposed approach was implemented in a pattern recognition system for detecting and tracking an object. The results obtained were compared with respect to the correlation calculation using the complete image. The proposed method managed to recover the correlation with a quality of up to  $31[dB]$  in terms of PSNR and exhibited a comparable tracking error of  $6[px]$  reducing the percentage of data acquired by up to 40 %.

---

\* Research Work.

\*\* School of Physical-Mechanical Engineering. Department of Electronic Engineering. Advisor: Tatiana Carolina Gelvez Barrera.

# INTRODUCCIÓN

Un sistema de detección de objetivos puede ser descrito por la comunicación entre un sensor y un controlador. El sensor es el encargado de transformar las magnitudes físicas en magnitudes eléctricas para que el controlador pueda procesarlas. A su vez, el controlador incorpora un algoritmo de reconocimiento de patrones (PR, del inglés Pattern Recognition) para detectar el objetivo dentro de un campo de visión [1]. La detección del objetivo puede implementarse en el algoritmo descrito mediante la inclusión de la técnica de emparejamiento de patrones (TM, del inglés Template Matching). El TM comúnmente conocido como correlación cruzada es una técnica para determinar la ubicación de un patrón de referencia sobre una imagen de entrada utilizando un análisis pixel a pixel [2]. El proceso consiste en desplazar un patrón de referencia sobre todas las posiciones de una imagen de entrada y calcular un índice numérico del grado de similitud en cada posición [3]. Adicionalmente, la correlación cruzada permite incluir filtros para solucionar problemas de varianza debido a distorsiones geométricas y/o degradaciones de intensidad presentes en la imagen de entrada [4, 5]. En la práctica, el filtro de correlación es almacenado previamente a la adquisición de la imagen de entrada. Luego, la correlación cruzada se calcula de manera eficiente en el dominio de la frecuencia siguiendo el teorema establecido en [6]. Generalmente, las implementaciones de correlación que usan un filtro en frecuencia como en los trabajos [7, 8] han mostrado robustez en ambientes controlados o invarianza ante diferentes factores de ruido.

Por lo anterior, la correlación ha sido de utilidad en aplicaciones como detección de características faciales [9], análisis de partículas individuales [10] y tomografía de electrones [11]. A diferencia de otros métodos, la correlación presenta las siguientes ventajas: un fundamento matemático aceptable, la posibilidad de implementarse de manera óptica y su propiedad de invarianza al desplazamiento [4]. Sin embargo, las limitaciones de adquisición, almacenamiento y transmisión de una cantidad de información cada vez mayor han previsto a los sistemas compresivos como el futuro de la adquisición de datos [12]. Por lo tanto, es de gran interés incursionar en el uso de una técnica que permita obtener la correlación a partir de medidas comprimidas. Para el caso de análisis de vídeo, el seguimiento de un objetivo por correlación implica la captura de cada uno de los cuadros del video, la transferencia de los datos, y el almacenamiento del video completo para calcular la ubicación por correlación a lo largo de los cuadros. Esto demanda

---

una mayor capacidad de los recursos en comparación al seguimiento del objetivo capturando solo unas pocas mediciones de cada cuadro del video como dicta la teoría de adquisición compresiva (CS, del inglés Compressive sensing).

CS establece que una señal puede ser recuperada a partir de un conjunto de pocas observaciones si se cumple que la señal es poco densa en una base de representación apropiada y la matriz de adquisición es suficientemente incoherente [12, 13]. El modelo de adquisición compresiva conlleva a reducir la tasa de transferencia de información, el costo de almacenamiento, y posiblemente el tiempo de adquisición con la certeza de recuperar la escena original de manera exacta [12]. Tradicionalmente, la estimación de la matriz de correlación bajo la teoría de CS es llevada a cabo reconstruyendo primero la imagen original a partir de las medidas comprimidas y calculando la correlación posteriormente con un filtro armónico circular (CHF, del inglés Circular Harmonic Filter) obtenido a partir del patrón de referencia. Con el objetivo de evitar el proceso de reconstrucción de la imagen original, este trabajo de grado presenta un enfoque para estimar directamente la matriz de correlación a partir de las medidas comprimidas capturadas de una imagen de entrada, sobre la cual se desea buscar el objetivo. El problema inverso se soluciona minimizando las normas  $\ell_1$  y  $\ell_2$  con el algoritmo para reconstrucción escasa por proyección de gradiente (GPSR, del inglés Gradient Projection for Sparse Reconstruction) [14]. La exactitud del método fue evaluada sobre el seguimiento de un objetivo al calcular iterativamente su ubicación sobre los cuadros del video. Esta ubicación fue obtenida mediante la matriz de correlación recuperada, la cual incluyó un CHF derivado de la fase para proporcionar invarianza a la rotación. Finalmente, las simulaciones fueron realizadas sobre datos reales y sintéticos logrando reconstrucciones de correlación comparables respecto al enfoque tradicional con una tasa de compresión de hasta el 40%. En términos de calidad de reconstrucción y exactitud del seguimiento se obtuvo un PSNR de hasta 31[*dB*], una tasa de éxito promedio de 97% y un error de seguimiento promedio de 6[*px*].

El contenido de este trabajo se organiza como sigue: en el capítulo 1 se describen las bases teóricas incluyendo el proceso de adquisición compresiva, la estructura de un sistema de reconocimiento de patrones, y la correlación cruzada; en el capítulo 2 se establece la metodología del trabajo presentando el planteamiento matemático, el problema de minimización y el algoritmo implementado; en el capítulo 3 se enumeran los materiales y métodos utilizados como el montaje experimental, la presentación del conjunto de datos de prueba sintético, y la descripción de las métricas de calidad utilizadas; en el capítulo 4 se muestra

---

el análisis de calidad de los resultados en función de diferentes parámetros de variación. Finalmente, en el capítulo 5 se presentan las conclusiones del trabajo.

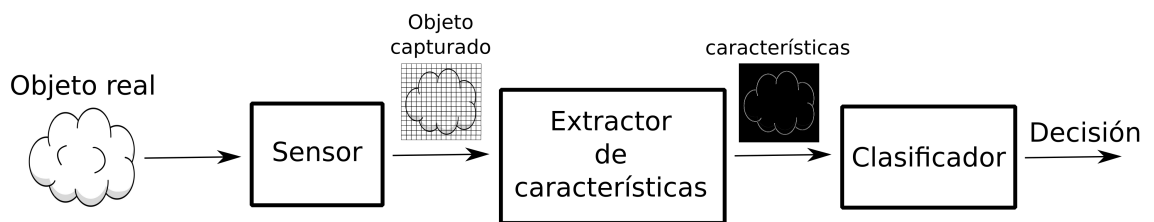
# 1. MARCO TEÓRICO

Este trabajo presenta un enfoque para obtener la matriz de correlación cruzada a partir de medidas comprimidas de un video con el fin de hacer seguimiento a un objetivo bajo el esquema de un sistema de reconocimiento de patrones. Por lo tanto, este capítulo describe los conceptos generales y teóricos de un sistema de reconocimiento de patrones, el calculo de la matriz de correlación cruzada usando filtros para reducir la varianza a la rotación, y el marco de la adquisición compresiva.

## 1.1. DEFINICIÓN Y CLASIFICACIÓN DEL RECONOCIMIENTO DE PATRONES

El reconocimiento de patrones (PR, del inglés Pattern Recognition) consiste en la identificación de propiedades específicas presentes en una señal. Este proceso puede ser llevado a cabo de dos maneras: basado en la extracción de características o por comparación de plantillas (TM, del inglés Template Matching). El PR basado en la extracción de características sigue tres procesos fundamentales: la adquisición de datos, que transforma magnitudes físicas en magnitudes eléctricas; la extracción de características de objetos conocidos para establecer propiedades de una clase de tal modo que puedan reconocerse los patrones en una señal [15, 16]; y la toma de decisiones, donde se clasifica un objeto dentro de una clase según sus características observadas [17]. Estos pasos se ilustran en la Figura 1.

**Figura 1:** Reconocimiento de patrones basado en la extracción de características.

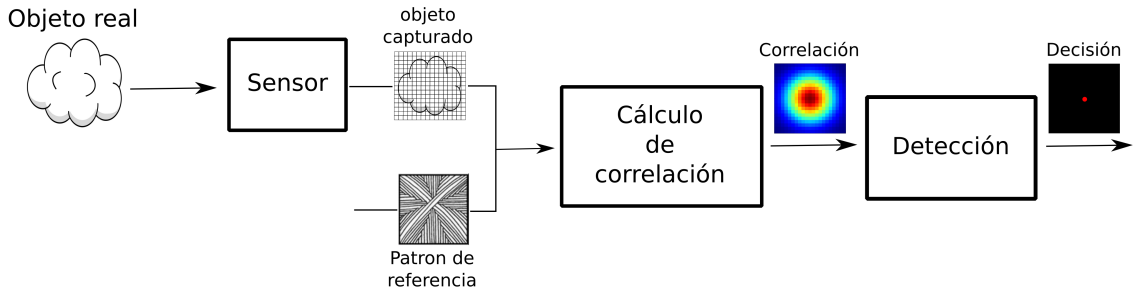


Este sistema intenta asignar a cada valor de entrada una clase predeterminada. Por lo tanto, tiene aplicaciones en los sistemas de diagnóstico asistido, el reconocimiento automático de voz y el reconocimiento de objetivos [18, 19, 20].

El PR mediante el modelo de TM considera que la gente almacena varias minicopias de patrones (plantillas) percibidas en el exterior. Así, cuando una persona percibe un patrón actual, su cerebro intenta reconocerlo mediante la búsqueda de una fuerte coincidencia entre este patrón y el conjunto de plantillas almacenadas en el cerebro. El resultado de una coincidencia corresponde generalmente al reconocimiento del patrón ya que cada una de estas plantillas fueron relacionadas a un significado específico en el momento de su almacenamiento [21]. En procesamiento de imágenes digitales, el TM define una técnica para determinar la presencia y la ubicación de un patrón de referencia o plantilla dentro de

una imagen de entrada mediante un análisis pixel a pixel [2]. El proceso consiste en desplazar el patrón de referencia sobre todas las posiciones posibles en la imagen de entrada para calcular un índice numérico que indique el grado de similitud en esa posición respecto al patrón de interés. La correlación es la métrica de similitud estándar en aplicaciones que usan TM según la literatura, ya que ha mostrado ser menos invariante respecto a otras medidas de similitud. Generalmente, la técnica de TM usando la correlación y la operación de correlación cruzada se refieren al mismo proceso de desplazamiento del patrón para obtener la matriz de coeficientes de similitud. Por lo tanto, los términos correlación y correlación cruzada serán usados indistintamente a lo largo del libro para hacer referencia a este tipo de emparejamiento. La Fig. 2 muestra los pasos generales que describen el PR basado en TM. Observe que, la decisión se toma con base en los picos de correlación calculados indicando la presencia del objetivo en esa posición.

**Figura 2:** Reconocimiento de patrones basado en la correspondencia de patrones.



Este enfoque puede usarse en el seguimiento de objetivos en videos calculando la correlación en cada cuadro del video. Algunas de sus aplicaciones más importantes se encuentran en las áreas de medicina, biología, y seguridad [22, 23, 24].

## 1.2. CORRELACIÓN CRUZADA ENTRE IMÁGENES

La correlación es una operación matemática que calcula un índice de similitud entre un patrón y un conjunto de regiones de pixeles que pertenecen a una imagen. Su estimación mediante el desplazamiento del patrón sobre la imagen puede modelarse mediante la operación de convolución cuando una de las señales de entrada es reflejada. Debido a esta semejanza de métricas se opta por definir independientemente cada operación, ya que sus resultados tienen significados muy diferentes. Matemáticamente los coeficientes de la correlación discreta  $2D$  entre dos imágenes  $\mathbf{G} \in \mathbb{R}^{m \times n}$  y  $\mathbf{F} \in \mathbb{R}^{M \times N}$  son calculados como,

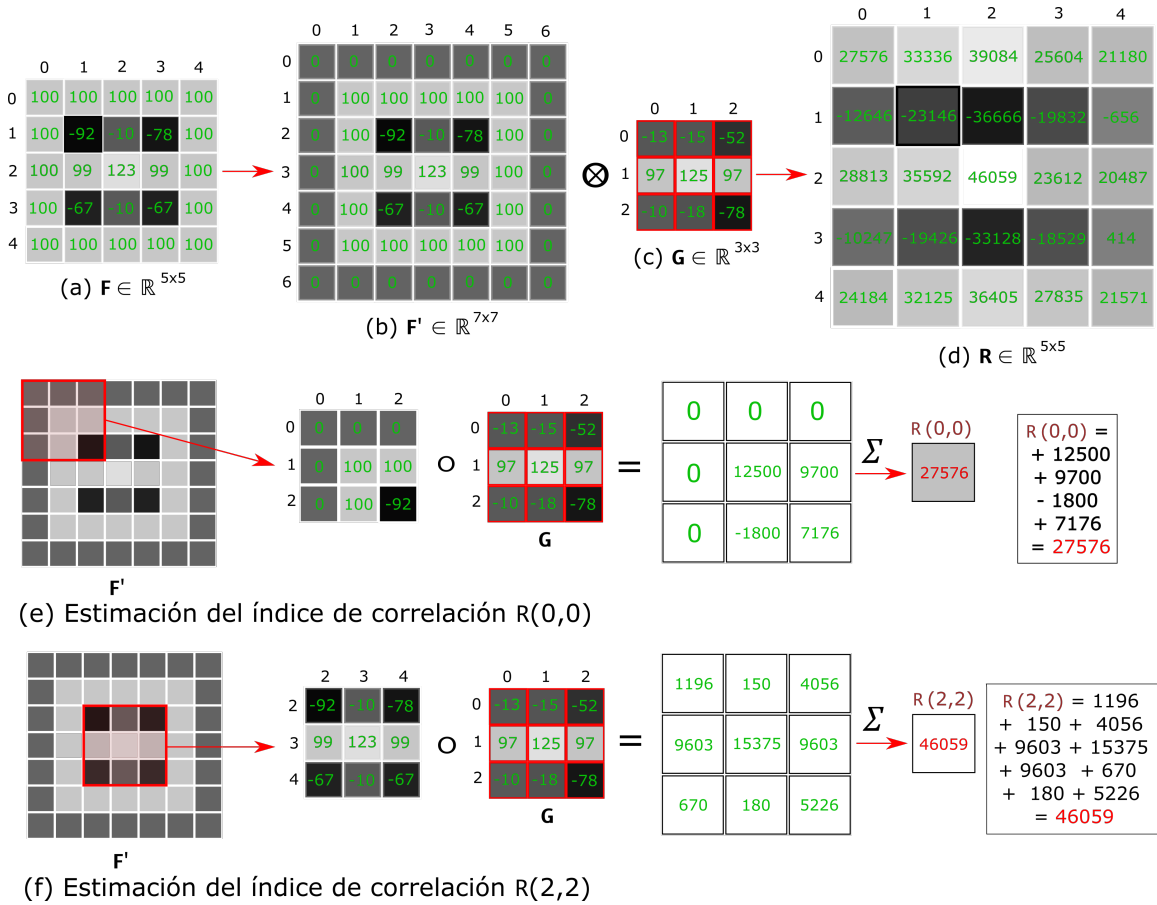
$$R_{GF}(\tau, \beta) = \mathbf{G} \otimes \mathbf{F} = \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} G(x, y)^* \circ F'(x + \tau, y + \beta), \quad (1.1)$$

$$\tau = 0, 1, 2, \dots, M - 1, \quad \beta = 0, 1, 2, \dots, N - 1.$$

donde  $R_{GF}(\tau, \beta)$  representa el coeficiente de correlación entre las imágenes cuando la imagen  $G$  es desplazada  $\tau$  unidades en  $x$  y  $\beta$  unidades en  $y$  siendo  $(x, y)$  las coordenadas espaciales,  $\otimes$  es el símbolo que denota la operación de correlación,  $G(x, y)^*$  denota la conjugada compleja de  $G(x, y)$ ,  $\circ$  representa la multiplicación Hadamard y  $F' \in \mathbb{R}^{M+m-1 \times N+m-1}$  es la imagen  $F$  redimensionada por algún método de relleno (del inglés padding). Observe que cada uno de los coeficientes de correspondencia almacenados en la correlación  $R_{GF}$  se calcula mediante la suma de multiplicaciones entre pares de coeficientes de las dos imágenes en función de los desplazamientos  $\tau$  y  $\beta$ .

En la literatura, términos como autocorrelación y correlación cruzada son utilizados para establecer una diferencia en Eq.(1.1) entre la correlación de una imagen con si misma  $G = F$  y la correlación de dos señales diferentes  $G \neq F$  respectivamente [25]. Estos términos son generalizados a cualquier dimensión para especificar la igualdad o diferencia de las señales de entrada. Específicamente, el desarrollo del libro estará enfocado en la correlación cruzada 2D, ya que los algoritmos de reconocimiento de patrones para la detección de objetivos involucran operaciones sobre imágenes diferentes. Visualmente, la correlación cruzada entre imágenes puede entenderse como se ilustra en la Figura 3,

**Figura 3:** Cálculo de la correlación cruzada entre dos imágenes.



Observe en Fig. 3(b) que la operación de relleno conocida como *zero padding*

es aplicada sobre la imagen  $\mathbf{F} \in \mathbb{R}^{5 \times 5}$  en la Fig. 3(a) para solucionar problemas tradicionales de bordes, obteniendo así la nueva imagen  $\mathbf{F}' \in \mathbb{R}^{7 \times 7}$  lista para correlacionar con el patrón de referencia  $\mathbf{G} \in \mathbb{R}^{3 \times 3}$  en Fig. 3(c). El resultado se observa en la Fig. 3(d). Note que, el cálculo de los coeficiente de correlación  $R(0,0)$  y  $R(2,2)$  se detalla en las Figs. 3(e) y 3(f) respectivamente, en donde  $\mathbf{G}$  es multiplicado punto a punto por una región de  $\mathbf{F}'$  delimitada en cada momento por una ventana que se mueve de izquierda-derecha y de arriba-abajo. Luego, el coeficiente de correlación almacenado en la matriz  $\mathbf{R} \in \mathbb{R}^{5 \times 5}$  en Fig. 3(d) es estimado como una suma de las multiplicaciones punto a punto obtenidas para una posición específica de la ventana. Cada índice de correlación en  $R(x,y)$  es calculado como una correspondencia entre la imagen  $\mathbf{G}$  y la región de  $\mathbf{F}'$  delimitada por la ventana en la posición  $(x,y)$ . De acuerdo a lo anterior, se comprueba la ubicación del pico de correlación en la posición  $(2,2)$  de la matriz  $\mathbf{R}$ , en donde los pixeles involucrados tienen la mayor correspondencia.

Observe que el cálculo de la correlación en el dominio espacial requiere una gran cantidad de cálculos. Para una imagen rellena  $\mathbf{F}' \in \mathbb{R}^{N \times N}$  y un patrón de referencia  $\mathbf{G} \in \mathbb{R}^{M \times M}$  se requieren cerca de  $M^2(N - M + 1)^2$  productos y  $M^2(N - M + 1)^2$  adiciones. Por lo tanto, en la literatura se propone calcular la correlación en el dominio de la frecuencia para reducir la cantidad de operaciones realizadas [26]. Esta alternativa fue derivada del teorema de convolución, el cual establece que la transformada de Fourier de una convolución entre dos señales es el producto Hadamard entre las transformadas de Fourier de las señales [6]. Así, el teorema de convolución es expresado matricialmente como,

$$\mathcal{F}_2(\mathbf{G}' * \mathbf{F}') = \mathcal{F}_2(\mathbf{G}') \circ \mathcal{F}_2(\mathbf{F}') \quad (1.2)$$

donde  $\mathcal{F}_2(\cdot) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  es un operador que calcula la transformada de Fourier 2D de sus argumento,  $\mathbf{G}' \in \mathbb{R}^{N \times N}$  es el patrón de referencia luego de un relleno de ceros suponiendo que  $N > M$ , y  $(*)$  denota el operador convolución. Realizando la transformada inversa de Forier 2D  $\mathcal{F}_2^{-1}(\cdot)$  en ambos lados de la igualdad en la Eq. (1.2) se obtiene una expresión para calcular la convolución entre dos imágenes en el dominio de la frecuencia,

$$\mathbf{G}' * \mathbf{F}' = \mathcal{F}_2^{-1}(\mathcal{F}_2(\mathbf{G}') \circ \mathcal{F}_2(\mathbf{F}')) \quad (1.3)$$

Luego, la correlación cruzada es obtenida realizando una inversión en el tiempo sobre una de las imágenes de entrada,

$$-\mathbf{G}' * \mathbf{F}' = \mathbf{G}' \otimes \mathbf{F}' = \mathcal{F}_2^{-1}(\mathcal{F}_2^*(\mathbf{G}') \circ \mathcal{F}_2(\mathbf{F}')) \quad (1.4)$$

donde  $\mathcal{F}_2(-\mathbf{G}') = \mathcal{F}_2^*(\mathbf{G}')$  siempre y cuando  $\mathbf{G}'$  sea una imagen con valores reales, tal que  $\mathbf{G}'^* = \mathbf{G}'$ . La correlación cruzada entre las imágenes  $\mathbf{G}'$  y  $\mathbf{F}'$  en el dominio de Fourier puede ser obtenida por  $12 \cdot N^2 \cdot \log_2(N)$  productos y  $18 \cdot N^2 \cdot \log_2(N)$  adiciones / sustracciones. Observe que la correlación en el dominio espacial puede llegar a ser más eficiente cuando la dimensión  $N$  es comparativamente mayor a la dimension  $M$ . Sin embargo, en casos más comunes en donde las dimensiones se aproximan, la correlación en el dominio de la frecuen-

cia es la opción más eficiente [26]. Por consiguiente, la elección de uno de los métodos depende de las dimensiones de la imagen de entrada.

### Inclusión de filtros en el cálculo de la correlación

La correlación fue inicialmente usada sobre un conjunto limitado de aplicaciones ya que los métodos basados en correlación son sensibles a degradaciones que modifican niveles de intensidad de la imagen y a distorsiones geométricas sobre los objetos de interés [4]. Estas limitaciones fueron abordadas mediante el diseño de filtros de correlación, acción promovida con la aparición del filtro espacial complejo u holográfico en 1965, también conocido como filtro acoplado [27]. El interés en desarrollar filtros cada vez más eficientes produjeron avances en el campo del reconocimiento de patrones ópticos para detectar un objeto ante distorsiones de escalamiento, rotación y otro tipo de ruido. En aplicaciones donde el escalamiento de objetos no es evidente, un filtro invariante a la rotación podría ser suficiente para detectar con una aproximación aceptable objetos sobre un ambiente controlado [28].

Específicamente, el problema de respuesta variante ante distorsiones de rotación puede ser mitigado usando un filtro armónico circular (CHF, del inglés Circular Harmonic Filter). Este filtro es incluido en el cálculo de la correlación en (1.4) sustituyendo al patrón de referencia  $\mathbf{G}'$  en el dominio de la frecuencia. Los CHFs están basados en la descomposición de un patrón de referencia 2D en coordenadas polares  $\tilde{\mathbf{G}}(r, \vartheta)$  en sus componentes armónicas circulares (CHC, del inglés Circular Harmonic Components) formuladas como,

$$\tilde{\mathbf{G}}(r, \vartheta) = \sum_{s=-\infty}^{\infty} \tilde{\mathbf{G}}_s(r) e^{is\vartheta}, \quad (1.5)$$

donde  $(r, \vartheta)$  denotan las coordenadas radial y angular, y  $\tilde{\mathbf{G}}_s(r)$  es la componente radial de orden  $s$  calculada como,

$$\tilde{\mathbf{G}}_s(r) = \frac{1}{2\pi} \int_0^{2\pi} \tilde{\mathbf{G}}(r, \vartheta) e^{-is\vartheta} d\vartheta. \quad (1.6)$$

De acuerdo a (1.6) las componentes radiales son solo afectadas cuando el patrón está escalado [28]. Por lo tanto, objetos con distorsiones de rotación son tenidos en cuenta por el CHF. Basado en la descomposición CHC en (1.5), la respuesta al impulso  $\mathbf{h}_s(r, \vartheta)$  y la respuesta en frecuencia  $\mathbf{H}_s(\rho, \phi)$  para un CHF clásico de orden  $s$  están dados por,

$$\mathbf{h}_s(r, \vartheta) = \tilde{\mathbf{G}}_s(r) e^{is\vartheta}, \quad (1.7)$$

$$\mathbf{H}_s(\rho, \phi) = \mathcal{F}_2 \left( \tilde{\mathbf{G}}_s(r) e^{is\vartheta} \right) = \tilde{\mathcal{G}}_s(\rho) e^{is\phi}, \quad (1.8)$$

donde  $(\rho, \phi)$  denotan coordenadas en frecuencia [29]. En las Eqs. (1.9), (1.10) y

(1.11) se muestran algunos diseños de filtros de correlación basados en el CHF clásico presentados en [29, 30, 31] respectivamente. Cada filtro promueve un comportamiento particular para fortalecer los resultados respecto a una métrica específica.

- CHF solo fase:

$$\mathbf{H}_s(\rho, \phi) = \frac{\tilde{\mathcal{G}}_s(\rho)e^{is\phi}}{|\tilde{\mathcal{G}}_s(\rho)|}. \quad (1.9)$$

- CHF de transferencia óptima:

$$\mathbf{H}_s(\rho, \phi) = \frac{\tilde{\mathcal{G}}_s(\rho)e^{is\phi}}{\int_0^{2\pi} |\tilde{\mathcal{G}}_s(\rho, \phi)|^2 d\phi + K}, \quad (1.10)$$

donde  $K$  es una constante sintonizable.

- CHF derivado de la fase:

$$\tilde{\mathbf{H}}(\rho, \phi) = \frac{\tilde{\mathcal{G}}_s(\rho, \phi)}{|\tilde{\mathcal{G}}_s(\rho, \phi)|},$$

$$\mathbf{H}_s(\rho, \phi) = \tilde{\mathbf{H}}_s(\rho)e^{is\phi}. \quad (1.11)$$

La correlación cruzada en (1.4) es reformulada incluyendo un CHF  $\mathbf{H}_s(\rho, \phi)$  de la siguiente manera,

$$\mathbf{R} = sh \left( \mathcal{F}_2^{-1} \left( \mathbf{H}_s^*(u, v) \circ \left\{ \mathcal{F}_2 \left( \frac{sh(\mathbf{F})}{max(\mathbf{F})} \right) \right\} \right) \right), \quad (1.12)$$

donde  $sh(\cdot)$  es un operador que realiza un corrimiento diagonal de los cuadrantes de su argumento,  $\mathbf{H}_s(u, v) \in \mathbb{R}^{N \times N}$  denota el CHF  $H_s(\rho, \phi)$  en coordenadas rectangulares y  $max(\cdot)$  retorna el elemento de mayor magnitud en su argumento.

### 1.3. TEORÍA DE ADQUISICIÓN COMPRESIVA.

La teoría de adquisición compresiva (CS, del inglés compressive sensing) establece que es posible recuperar señales a partir de una menor cantidad de información en comparación a la que utilizan los métodos tradicionales generalmente guiados por el teorema de Nyquist-Shannon. Para hacer esto posible, dos principios fundamentales: escasez e incoherencia, son establecidos sobre el protocolo de adquisición de datos. El principio de escasez establece que la señal de interés debe tener una representación poco densa cuando es expresada en una base de representación apropiada. El principio de incoherencia establece que las formas de onda de muestreo deben tener una representación extremadamente densa en la base de representación establecida. La observación crucial sobre estos principios es la posibilidad de diseñar eficientemente protocolos de muestreo que

capturen la información útil contenida en la señal escasa, y luego la condensan en un conjunto de datos de poca dimensión. En resumen, el CS es un protocolo concreto de muestreo para adquirir y comprimir los datos simultáneamente soportado en el uso de optimización numérica para recuperar la señal completa.

**1.3.1 Modelo matemático de adquisición compresiva** El mecanismo de muestreo modela la adquisición de la señal de interés, correlacionándola con un conjunto de formas de onda contenidas en la base de muestreo. Matemáticamente, sea  $\mathbf{f} = \mathcal{J}(\mathbf{F}) \in \mathbb{R}^{n=N^2}$  la representación vectorial de una imagen de entrada  $\mathbf{F} \in \mathbb{R}^{N \times N}$ , donde  $\mathcal{J}(\cdot) : \mathbb{R}^2 \rightarrow \mathbb{R}$  es un operador para reorganizar los coeficientes de una estructura 2D adentro de una estructura vectorial, entonces la configuración estándar para adquirir la información acerca de la imagen está dada de la siguiente manera

$$\mathbf{y} = \Phi \mathbf{f} + \mathbf{w} \quad (1.13)$$

donde  $\mathbf{y} \in \mathbb{R}^m$  son las mediciones adquiridas,  $\Phi \in \mathbb{R}^{m \times n}$  es la matriz de muestreo que modela el sistema óptico y  $\mathbf{w} \in \mathbb{R}^m$  es un término de error desconocido. Si las formas de onda, vistas como las filas de la base  $\Phi$  son funciones delta de Dirac,  $\mathbf{y}$  será un vector de valores muestreados de  $\mathbf{f}$  en el dominio del tiempo. Si las formas de onda son funciones de indicadores de pixeles,  $\mathbf{y}$  corresponde a la datos típicos coleccionados por sensores en una cámara digital. Si las formas de onda son sinusoides,  $\mathbf{y}$  es un vector de coeficientes de Fourier.

Debido a problemas comunes tales como: una cantidad limitada de sensores, un proceso de muestreo lento, o incluso una reducción en requerimientos de almacenamiento, existe un gran interés en situaciones de submuestreo en donde la cantidad de mediciones adquiridas parece ser insuficiente  $m \ll n$ . Estas circunstancias sugieren preguntas importantes. ¿es posible diseñar una base de muestreo  $\Phi$  que capture toda la información de  $\mathbf{f}$ ? ¿ $\mathbf{f}$  puede ser estimada a partir de un número  $m$  de muestras comparablemente menor a la dimensión  $n$  de la señal?, las cuales son abordadas por los principios de escasez e incoherencia.

**1.3.2 Escasez** La escasez afirma que la señal de interés  $\mathbf{f}$  debe ser poco densa cuando es expresada en una base apropiada  $\Psi \in \mathbb{R}^{n \times n}$ , también conocida como la base de representación. Por lo tanto, la señal  $\mathbf{f}$  es expandida sobre la base  $\Psi$  de la siguiente manera,

$$\mathbf{f} = \Psi \boldsymbol{\theta} \quad (1.14)$$

donde  $\boldsymbol{\theta} \in \mathbb{R}^n$  denota la representación escasa de  $\mathbf{f}$  en la base de representación  $\Psi$ . Los conceptos de representación poco densa, representación concisa o representación escasa se refieren a la característica de la señal  $\boldsymbol{\theta}$  en la cual la

---

mayoría de sus coeficientes son despreciables y solo unos pocos coeficientes relativamente grandes capturan la mayoría de la información. Por lo tanto, cuando una señal es escasa se pueden descartar sus coeficientes despreciables sin mucha pérdida perceptual. El principio de escasez actúa como una medida para determinar que tan eficiente se pueden adquirir señales no adaptativas, ya que las señales con alta escasez pueden ser adquiridas por un número menor de coeficientes a los que requieren las señales poco escasas.

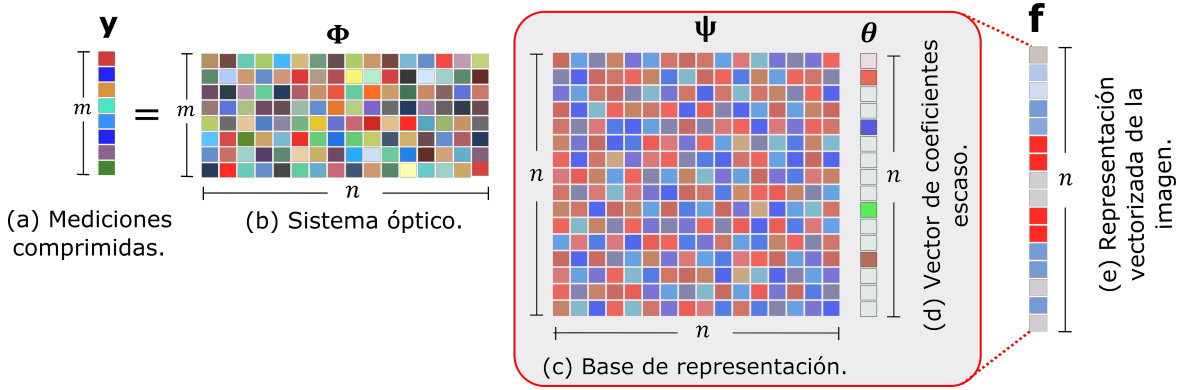
**1.3.3 Incoherencia** El principio de incoherencia establece una restricción sobre el par de bases  $(\Phi, \Psi)$  usadas para adquirir/muestrear en (1.13) y representar en (1.14) el objeto  $f$  respectivamente. Esta restricción logra simplificar el tratamiento requerido. La restricción consiste en obtener un valor de coherencia  $\mu(\Phi, \Psi)$  mínimo entre este par de bases el cual es calculado como

$$\mu(\Phi, \Psi) = \sqrt{n} \cdot \max(|\Phi\Psi|), \quad (1.15)$$

donde  $\max(\cdot)$  retorna el valor máximo en su argumento y  $\mu(\Phi, \Psi) \in [1, \sqrt{n}]$ . La coherencia mide la correlación mas grande entre cualquier par de elementos de  $\Phi$  y  $\Psi$ . Si este par de bases contiene elementos correlacionados su valor de coherencia será grande, de otro modo su valor será pequeño satisfaciendo las restricciones de CS ya que entre mayor sea el valor de incoherencia menor serán las medidas necesarias para capturar la señal  $f$  [32]. Algunos ejemplos tradicionales de pares con baja coherencia son presentados a continuación. El primer ejemplo consiste de  $\Phi$  como la base canónica y  $\Psi$  como la base de Fourier. Debido a que  $\Phi$  es la matriz de muestreo/adquisición éste corresponde al esquema clásico de muestreo en tiempo. Este par tiempo-frecuencia obtiene una máxima incoherencia  $\mu(\Phi, \Psi) = 1$ . El segundo ejemplo establece a  $\Phi$  como bases noiselets y a  $\Phi$  como bases wavelets. La coherencia entre esta combinación de bases puede obtener hasta valores de  $\sqrt{2}$ . Un tercer ejemplo son las matrices aleatorias las cuales poseen la característica importante de tener alta incoherencia con cualquier base fija  $\Psi$ . Seleccionando una base ortogonal  $\Phi$  uniformemente aleatoria se puede llegar a obtener con una alta probabilidad un valor de coherencia aproximado de  $\sqrt{2 \cdot \log n}$ .

**1.3.4 Modelo de adquisición compresiva sobre una imagen a escala de grises** La Fig. 4 ilustra el proceso de adquisición de una imagen a escala de grises bajo el modelo de CS.

**Figura 4:** Adquisición compresiva de una imagen



La imagen de entrada  $\mathbf{F} \in \mathbb{R}^{N \times N}$  es reorganizada en el vector  $\mathbf{f} \in \mathbb{R}^n$  mostrado en la Fig. 4(e). A su vez, la señal  $\mathbf{f}$  es representada de forma concisa por la representación escasa  $\theta \in \mathbb{R}^n$  en la en la Fig. 4(d) usando la base de representación  $\Psi \in \mathbb{R}^{n \times n}$  en la Fig. 4(c). Este par corresponde a la descomposición escasa de  $\mathbf{f}$ , la cual es correlacionada con funcionales lineales contenidos en el sistema óptico  $\Phi \in \mathbb{R}^{m \times n}$  en la Fig. 4(b) para obtener las medidas comprimidas  $\mathbf{y} \in \mathbb{R}^m$  en la Fig. 4(a). Una vez se ha adquirido el subconjunto de medidas  $\mathbf{y}$  a partir de los  $n$  coeficientes de  $\mathbf{f}$ , la recuperación de la señal es llevada a cabo por la minimización de la norma  $\ell_1$ . Esta operación consiste en la búsqueda de una solución  $\hat{\theta}$  al problema de optimización convexo formulado como,

$$\hat{\theta} = \underset{\theta \in \mathbb{R}^n}{\operatorname{argmin}} \|\theta\|_1 \quad \text{sujeto a} \quad \mathbf{y} = \Phi \Psi \theta \quad (1.16)$$

La solución corresponde a la secuencia de coeficientes que producen un valor mínimo de norma  $\ell_1$  entre el conjunto de soluciones candidatas que cumplen la restricción  $\mathbf{y} = \Phi \Psi \tilde{\theta}$ , siendo  $\tilde{\theta}$  una solución candidata. Finalmente, la propuesta de reconstrucción esta dada por  $\hat{\mathbf{f}} = \Psi \hat{\theta}$ . Si  $\mathbf{f}$  es suficientemente escasa, su recuperación usando la minimización- $\ell_1$  es con alta probabilidad exacta.

## 2. METODOLOGÍA

### 2.1. PLANTEAMIENTO MATEMÁTICO

El seguimiento de un objetivo a lo largo de un video puede realizarse calculando la correlación entre cada cuadro del video y el patrón de referencia para estimar la secuencia de posiciones del objetivo en el tiempo. Suponga un video  $\mathfrak{F} \in \mathbb{R}^{N \times N \times L}$  compuesto de  $L$  cuadros a escala de grises  $\mathbf{F}_k$ ,  $k = 1, \dots, L$  que denotan las imágenes de entrada sobre las cuales se encuentra el objetivo. La detección del objetivo en el cuadro  $k$  requiere la adquisición de la imagen  $\mathbf{F}_k$  modelada por,

$$\mathbf{y}_k = \Phi \mathbf{f}_k, \quad \mathbf{f}_k = \mathcal{J}(\mathbf{F}_k) \quad (2.1)$$

Teóricamente, el modelo matemático de adquisición en (2.1) captura la información importante de la representación escasa del cuadro de video  $\mathbf{F}_k$  y la condensa en el conjunto de mediciones  $\mathbf{y}_k$ . Con base en lo anterior, se propone incluir un conjunto de operadores lineales sobre la imagen vectorizada  $\mathbf{F}_k$  con el objetivo de establecer la correlación cruzada como su representación escasa en algún dominio particular. La inclusión de estos operadores al modelo de adquisición compresiva es descrita en tres pasos. Primero, se introduce el operador de transformación Fourier sobre el cuadro vectorizado  $\mathbf{f}_k$  con su operador inverso para mantener la consistencia del modelo tal que

$$\mathbf{y}_k = \Phi \mathcal{F}^{-1} \mathcal{F} \mathbf{f}_k \quad (2.2)$$

donde  $\mathcal{F} \in \mathbb{R}^{n \times n}$  y  $\mathcal{F}^{-1} \in \mathbb{R}^{n \times n}$  corresponden a las representaciones matriciales de los operadores  $\mathcal{F}_2(\cdot)$  y  $\mathcal{F}_2^{-1}(\cdot)$ , respectivamente. Segundo, un CHF se incluye entre el par de operadores introducidos en el paso anterior con el fin de obtener invarianza a la rotación. En consecuencia, la ecuación (2.2) es reformulada como,

$$\mathbf{y}_k = \Phi \mathcal{F}^{-1} \mathbf{H}^{-1} \mathbf{H} \mathcal{F} \mathbf{f}_k \quad (2.3)$$

donde  $\mathbf{H} \in \mathbb{R}^{n \times n}$  es una matriz diagonal que contiene los coeficientes del CHF  $\mathbf{H}_s$ . Finalmente, el par de transformaciones opuestas de Fourier son nuevamente incluidas sobre el modelo (2.3) para establecer la representación escasa deseada obteniendo

$$\mathbf{y}_k = \Phi \mathcal{F}^{-1} \mathbf{H}^{-1} \mathcal{F} \mathcal{F}^{-1} \mathbf{H} \mathcal{F} \mathbf{f}_k \quad (2.4)$$

La inclusión del conjunto de operadores lineales descrito implica definir la descomposición escasa  $(\Psi, \theta_k)$  de la siguiente manera,  $\Psi = \mathcal{F}^{-1} \mathbf{H}^{-1} \mathcal{F} \in \mathbb{R}^{n \times n}$  y  $\theta = \mathcal{F}^{-1} \mathbf{H} \mathcal{F} \mathbf{f}_k \in \mathbb{R}^n$ . Observe que, la inclusión de estos operadores lineales permite representar a  $\mathbf{f}_k$  de manera escasa mediante la correlación cruzada  $\theta = \mathbf{R}$ , siendo  $\mathbf{R}_k \in \mathbb{R}^{N \times N}$  la correlación cruzada estimada en el cuadro  $k$ . Por lo tanto, la salida de un algoritmo de reconstrucción basado en escasez será la correlación cruzada siempre y cuando se use la descomposición escasa propuesta para establecer a  $\mathbf{y}_k$  como el conjunto de información condensada más importante de

la correlación cruzada  $\mathbf{R}_k$ .

$$\mathbf{y}_k = \Phi \mathcal{T}(\Psi \mathbf{R}_k) \quad (2.5)$$

Otras descomposiciones  $(\Psi, \theta)$  solo usan la propiedad de escasez para encontrar una solución al problema inverso y requieren la proyección de la representación escasa estimada sobre la base de representación  $\hat{\mathbf{f}} = \Psi \hat{\theta}$ .

El problema inverso para estimar la correlación cruzada a partir de las medidas comprimidas es formulado mediante la minimización con restricción de igualdad lineal en (1.16), el cual es relajado a un problema de optimización sin restricciones que combina la norma  $\ell_1$  para inducir escasez, con la norma  $\ell_2$  para mantener la fidelidad de los datos [14]. En consecuencia, el problema de optimización convexo para obtener la matriz de correlación correspondiente al cuadro  $\mathbf{F}_k$  es reescrito como,

$$\hat{\theta}_k = \underset{\theta_k \in \mathbb{R}^n}{\operatorname{argmin}} \quad \|\mathbf{y}_k - \Phi \Psi \theta_k\|_2^2 + \tau_k \|\theta_k\|_1 \quad (2.6)$$

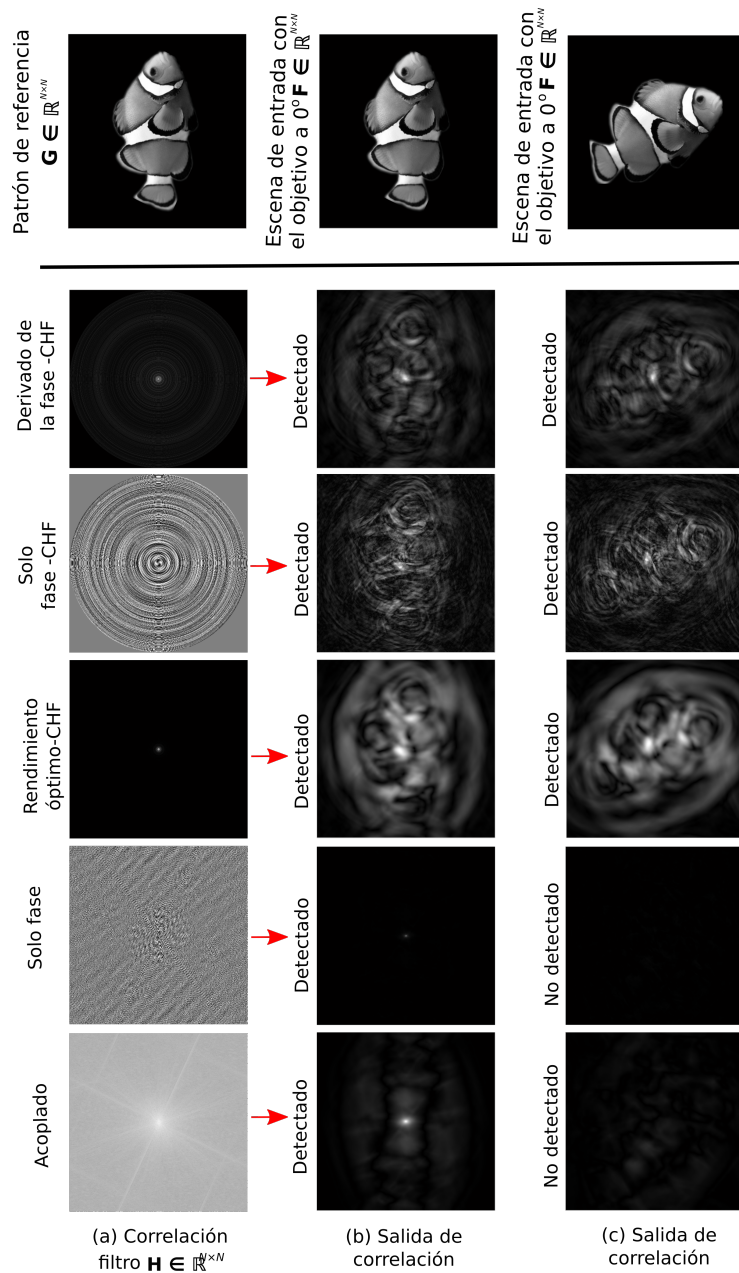
donde  $\tau_k > 0$  es un parámetro de regularización para balancear los objetivos de minimización y  $\theta = \mathcal{T}(\mathbf{R}_k)$  es la correlación cruzada estimada. La propuesta de reconstrucción es obtenida redimensionando la solución estimada  $\hat{\theta}$  a una estructura matricial  $\hat{\mathbf{R}}_k = \mathcal{T}^{-1}(\hat{\theta})$ . Observe que la salida del problema de optimización convexa es directamente la correlación cruzada y no requiere su proyección sobre la base de representación como sucede en la reconstrucción de la imagen original. Además, evita el paso de estimación de la matriz de correlación luego de obtener la solución mediante un algoritmo de reconstrucción. El problema de minimización en (2.6) fue solucionado iterativamente en cada cuadro del video por el algoritmo GPSR el cual se basa en restricciones implícitas que guían un proceso iterativo en la búsqueda de una solución aproximada. Específicamente, el GPSR usa un algoritmo de proyección del gradiente para la formulación de la programación cuadrática limitada y restringida (BCQP, del inglés Bound-Constrained Quadratic Programming) en la cual la búsqueda de un camino en cada iteración es obtenida proyectando la dirección del gradiente negativo sobre el conjunto factible [14]. Otros solucionadores potenciales en la literatura para resolver el problema en (2.6) son presentados en [33, 34, 35].

## 2.2. ELECCIÓN DE UN FILTRO DE CORRELACIÓN

El diseño de filtros de correlación es guiado hacia la optimización de las métricas de exactitud presentadas en [36] que evalúan parámetros como la cantidad de energía transferida, la concentración del pico de correlación, entre otros. La elección del filtro  $\mathbf{H}_s$  en (1.11) en este trabajo obedece a dos parámetros críticos en la detección de objetivos: La invarianza a la rotación y la habilidad para discriminar el ruido, donde el ruido hace referencia a cualquier objeto diferente al patrón de referencia. Para la elección del filtro de correlación óptimo fueron llevadas a cabo simulaciones con tres objetos diferentes: el objeto pez, el objeto tortuga y el objeto cangrejo. Una primera fase de simulaciones consistió en evaluar la invarianza a la rotación de cinco diseños diferentes de filtros de correlación: el filtro

acoplado, el filtro solo fase, el CHF de transferencia óptima, el CHF solo fase y el CHF derivado de la fase. Para ello se estimó la correlación entre el objeto pez y la escena de entrada luego de aplicar distorsiones de rotación sobre el objetivo. Una segunda fase de simulaciones fue realizada para observar la discriminación al ruido que exhibe cada uno de los filtros mencionados. Por lo tanto, se realizó la detección del objetivo por correlación en presencia de varios objetos sobre la escena de entrada. La Fig. 5 ilustra una comparación de rendimiento entre filtros convencionales (solo fase y acoplado) variantes a distorsiones de rotación y CHFs (derivado de la fase, solo fase y transferencia óptima).

**Figura 5:** Invarianza a la rotación incluyendo filtros armónicos circulares.

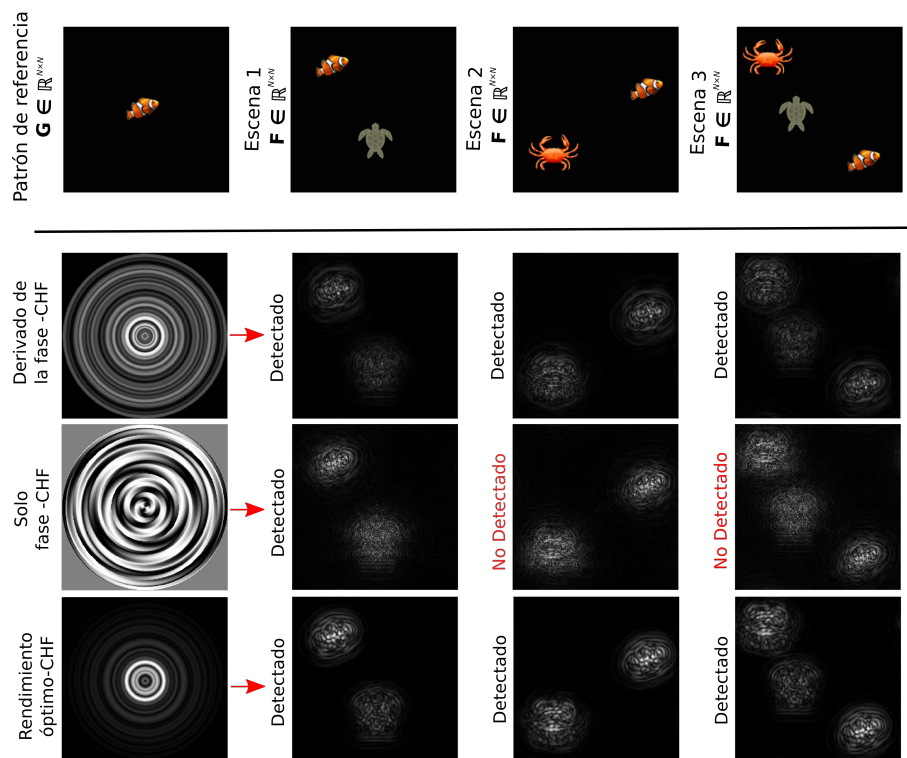


Observe en la Fig. 5(a) el patrón de referencia de prueba y la apariencia visual

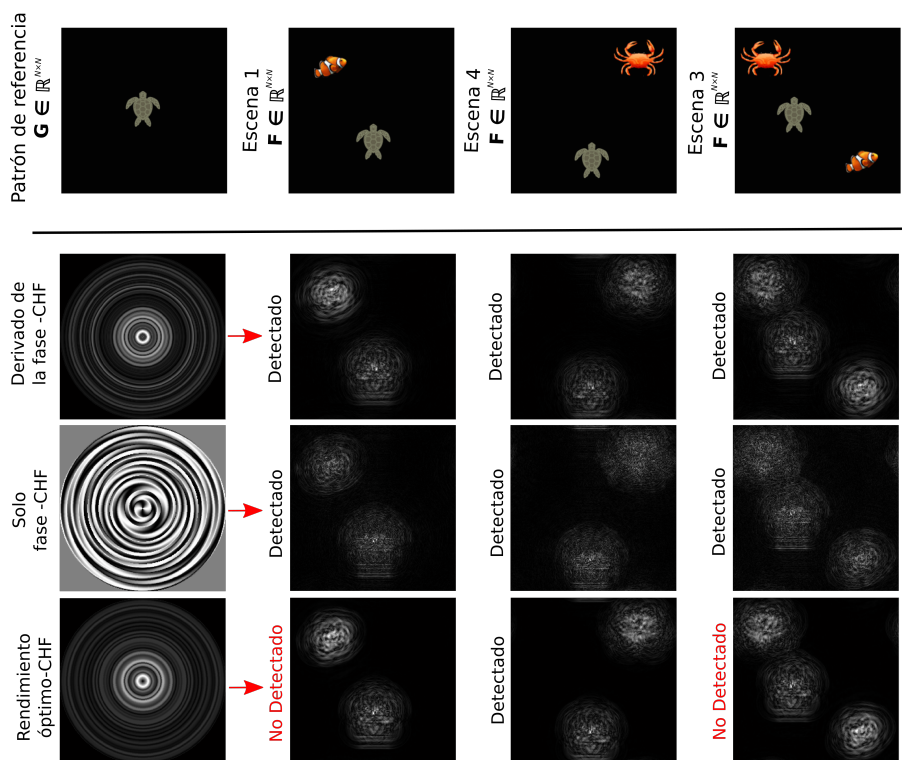
de cada filtro calculado. Las salida de correlación cuando el objetivo es rotado a un ángulo de  $0^\circ$  y  $62^\circ$  son mostradas en las Figs. 5(b) y 5(c) respectivamente. Note que, el objeto pez es detectado apropiadamente por todos los filtros cuando el objetivo no se encuentra rotado. Mientras que, un ángulo de rotación de  $62^\circ$  sobre el objetivo es suficiente para ocasionar que los filtros convencionales solo fase y acoplado no detecten el objeto pez. Sin embargo, el pico de correlación permanece constante cuando los CHF's son incluidos en la estimación de la matriz de correlación. Con base en los resultados de invarianza a la rotación fueron descartados los filtros convencionales 1 y 2 para el seguimiento de un objetivo, ya que el cambio de orientación es una característica común en aquellos objetos que describen un movimiento.

El siguiente experimento pretende evaluar el rendimiento de los filtros de correlación en términos de discriminación, refiriéndose a la discriminación como la capacidad para detectar el objeto deseado. Para cada patrón de referencia se obtuvo la correlación cruzada con tres imágenes de entrada diferentes usando los CHF's 3, 4 y 5 calculados del patrón. Las imágenes de entrada utilizadas en el experimento permitieron evaluar la detección de los patrones de referencia en presencia de hasta dos objetos diferentes. Específicamente, cada patrón de referencia se evaluó contra dos imágenes de entrada que exhiben a dos de los objetos evaluados y una tercera imagen que incluye a los tres objetos de prueba suponiendo en cada imagen la presencia del objetivo. Las Figs. 6, 7 y 8 muestra los resultados de correlación y la capacidad que obtuvo cada CHF para detectar los tres patrones de referencias sobre las imágenes de entrada.

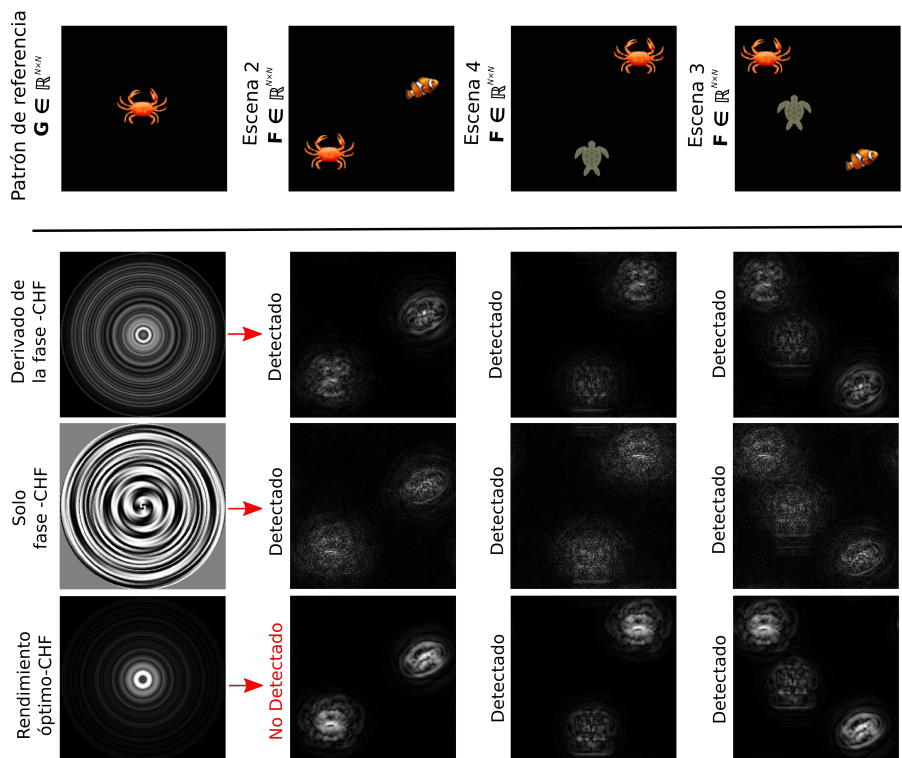
**Figura 6:** Estimación de la correlación para la detección del objetivo pez.



**Figura 7:** Estimación de la correlación para la detección del objetivo tortuga.



**Figura 8:** Estimación de la correlación para la detección del objetivo cangrejo.



La discriminación se calculó como un porcentaje de la diferenciación entre el

pico objetivo y el pico de ruido con respecto a la máxima intensidad presente en la matriz de correlación

$$D = \left( \frac{P_o - P_r}{\max(\mathbf{R})} \right) 100, \quad (2.7)$$

donde  $P_o$  y  $P_r$  son los valores de pico de correlación producidos al aplicar el filtro en la posición del objetivo y el ruido respectivamente. Para efectos del cálculo, el valor de pico de ruido  $P_r$  fue tomado como la intensidad máxima de correlación excluyendo el valor del pico objetivo  $P_o$ . Los valores de discriminación obtenidos por los CHF para los diferentes objetivos se muestran en la tabla 1

### Objeto tortuga

**Tabla 1:** Valores de discriminación obtenidos por los CHF en la detección de los objetivos pez, tortuga y cangrejo

CHF	Pez			Tortuga			Cangrejo		
	E1	E2	E3	E1	E4	E3	E2	E4	E3
#3	<b>78.6</b> %	27.9 %	30.5 %	-1.3 %	37.8 %	-7.9 %	-9.2 %	75.8 %	3.2 %
#4	13.9 %	-12.9 %	-3.5 %	<b>59.3</b> %	<b>58</b> %	<b>54.9</b> %	<b>40.7</b> %	26.1 %	<b>9.3</b> %
#5	77.4 %	<b>56.5</b> %	<b>61.7</b> %	1.4 %	55.2 %	10.2 %	8.8 %	<b>77.7</b> %	4.3 %

Los porcentajes negativos indican la discriminación que obtuvo el pico de ruido con referencia al pico objetivo, lo que supone una detección errónea del objetivo. Observe que el CHF derivado de la fase representó el filtro óptimo en el seguimiento de un objetivo, ya que mostró ser invariante a distorsiones de rotación y además logró detectar exitosamente y en presencia de ruido todos los objetivos evaluados. Con base en la simulaciones presentadas se optó por usar el CHF derivado de la fase como el filtro de correlación  $H_s$  en el desarrollo del método propuesto.

## 2.3. ALGORITMO PROPUESTO

El algoritmo de PR propuesto fue implementado para estimar la matriz de correlación cruzada a partir de un conjunto de medidas comprimidas. La captura de éstas medidas fue simulada por un diseño teórico de la matriz de adquisición que permitió probar la exactitud del método en el contexto de reconstrucción. Con el objetivo de evaluar la matriz de correlación como una herramienta para la detección, el algoritmo se implementó para estimar la ruta que sigue un objetivo a lo largo de un video. Los parámetros de entrada al algoritmo corresponden a: el video de prueba con  $L$  cuadros  $\mathcal{F} = \{\mathbf{F}_1, \mathbf{F}_2, \dots, \mathbf{F}_L\}$ , la matriz que modela el sistema óptico compresivo  $\Phi$ , la base de representación propuesta  $\Psi_1$ , el umbral para un proceso de binarización  $\varepsilon$ , y el conjunto  $\tau$  de parámetros de regularización en el problema de optimización convexo referentes a cada cuadro del video. Aunque no se mencione explícitamente el patrón de referencia, su respectivo CHF se encuentra incluido en la base de representación  $\Psi_1$ . Una vez el algoritmo termina su ejecución se obtienen la posiciones estimadas del objetivo a lo largo del video ( $\chi = \chi_1, \chi_2, \dots, \chi_L, \gamma = \gamma_1, \gamma_2, \dots, \gamma_L$ ). El pseudocódigo en el Algoritmo

1 presenta la secuencia de pasos mas relevantes del algoritmo propuesto para obtener las reconstrucciones de correlación y evaluarlas sobre el seguimiento de un objetivo.

---

### Algoritmo 1

---

```

1: Procedimiento TRACKINGFROMCOMPRESSIVECORRELATION( $\mathfrak{F} \in \mathbb{R}^{N \times N \times L}, \Phi \in \mathbb{R}^{m \times n}, \Psi \in \mathbb{R}^{n \times n}, \varepsilon, \tau$ )
2:    $k \leftarrow 0$ 
3:   mientras  $k < L$  hacer ▷ Iterar sobre cada cuadro del video
4:      $\mathbf{y}_k \leftarrow \Phi \mathcal{J}(\mathbf{F}_k)$  ▷ Adquisición comprimida
5:      $\hat{\boldsymbol{\theta}}_k \leftarrow \underset{\boldsymbol{\theta}_k \in \mathbb{R}^{N \times N}}{\operatorname{argmin}} \|\mathbf{y}_k - \Phi \Psi \boldsymbol{\theta}_k\|_2^2 + \tau \|\boldsymbol{\theta}_k\|_1$  ▷ Paso de reconstrucción
6:      $\hat{\mathbf{R}}_k \leftarrow \mathcal{J}^{-1}(\hat{\boldsymbol{\theta}}_k)$  ▷ Matriz de correlación cruzada
7:      $\mathbf{W}_k \leftarrow \delta_\varepsilon(\hat{\mathbf{R}}_k, \varepsilon)$ 
8:      $\ell_k \leftarrow \zeta(\mathbf{W}_k)$ 
9:      $\chi_k \leftarrow \frac{\sum_z (1,0) \ell_k}{z}$  ▷ Localización
10:     $\gamma_k \leftarrow \frac{\sum_z (0,1) \ell_k}{z}$  ▷ Localización
11:     $k \leftarrow k + 1$ 
12:   fin mientras
13:   return  $\chi, \gamma$  ▷ Resultados de seguimiento y correlaciones estimadas
14: fin Procedimiento

```

---

El paso a paso del Algoritmo 1 se describe como

1. Instrucción 3, Permite recorrer cada cuadro del video de forma iterativa.
2. Instrucción 4, adquisición compresiva de un cuadro de video emparejado en intensidades. las  $m$  medidas  $\mathbf{y}_k$  son capturadas a partir del cuadro  $\mathbf{F}_k$  modelando al sistema compresivo como la matriz de muestreo  $\Phi$ .
3. Instrucción 5, estimación de la correlación cruzada mediante un algoritmo de reconstrucción. Usando el vector de mediciones  $\mathbf{y}_k$  se soluciona el problema de optimización convexa para obtener una aproximación de la matriz de correlación-cruzada.
4. Instrucción 6, Reordenamiento de los coeficiente en  $\hat{\boldsymbol{\theta}}_k$  para obtener la estructura matricial  $\hat{\mathbf{R}}_k$ .
5. Instrucción 7, binarización de la matriz de correlación cruzada. Para encontrar la ubicación del pico de correlación más alto, la matriz de correlación obtenida  $\hat{\mathbf{R}}_k$  es binarizada con respecto al umbral  $\varepsilon$  usando el operador  $\delta_\varepsilon$ . si la intensidad de un pixel en  $\hat{\mathbf{R}}_k$  es menor a  $\varepsilon$ , la intensidad del pixel correspondiente en  $\mathbf{W}_k$  es establecida como 0. De lo contrario, tomara intensidad 1.  $\delta_\varepsilon$  denota el operador:

$$\mathbf{W}_{k_{ij}} = \delta_\varepsilon(\hat{\mathbf{R}}_k) = \begin{cases} 1, & \hat{\mathbf{R}}_{k_{ij}} > \varepsilon \\ 0, & otherwise \end{cases} \quad (2.8)$$

- 
6. Instrucción 8, búsqueda de los píxeles con una intensidad mayor al umbral de acuerdo a la matriz binaria  $\mathbf{W}_k$ . El conjunto  $\ell_k$  contiene las tuplas de índices que representan las posiciones horizontales y verticales del objetivo. El operador  $\zeta$  con el parámetro  $\mathbf{W}_k$  obtiene el conjunto:

$$\ell_k : \{(i_s, j_s) | \mathbf{W}_{k_{i_s j_s}} = 1\} \quad s = 1, \dots, z \quad (2.9)$$

donde  $z$  es el número de coeficientes uno en  $\mathbf{W}_k$ .

7. Instrucción 9, Calcula la posición horizontal del centroide  $\chi_k$  como una relación entre la sumatoria de las posiciones horizontales en  $\ell_k$  y el número de posiciones  $z$ .
8. Instrucción 10, Calcula la posición vertical del centroide  $\gamma_k$  como una relación entre la sumatoria de las posiciones verticales en  $\ell_k$  y el número de posiciones  $z$ .
9. Instrucción 13, Devuelve las posiciones del objetivo  $(\chi, \gamma)$ .

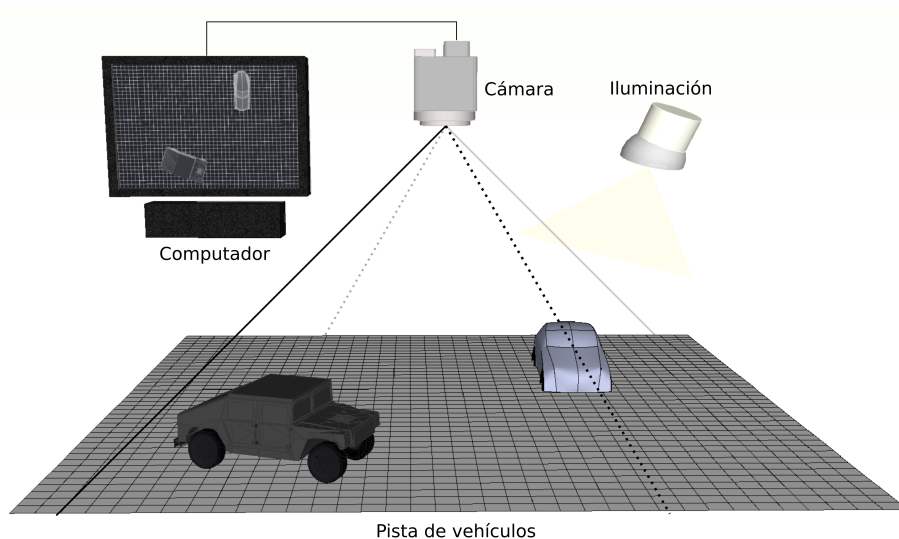
### 3. MATERIALES Y MÉTODOS.

Esta sección presenta el montaje experimental construido para simular una arquitectura óptica compresiva que captura los datos en un ambiente real y controlado. Luego, una descripción detallada del conjunto de datos de prueba para evidenciar su estructura y la forma en la cual fueron capturados es presentada. Finalmente, las métricas de calidad en la literatura fueron introducidas para evaluar la exactitud de la aproximación propuesta.

#### 3.1. MONTAJE EXPERIMENTAL PARA LA ADQUISICIÓN DE UN CONJUNTO DE DATOS REAL

Una aplicación de PR fue construida para verificar la calidad de las reconstrucciones de correlación y su utilidad para tareas de localización de objetivos. La configuración consiste de una arquitectura óptica tradicional, una escena de prueba y un algoritmo de PR. La Fig. 9 ilustra el montaje experimental realizado para soportar la aplicación descrita. El procedimiento consiste en la adquisición completa de una escena de prueba mediante la arquitectura óptica que describe una cámara digital con  $640 \times 480$  píxeles de resolución espacial y dirigida a  $-90$  grados acimut. Para evaluar la localización de un objeto, la escena de prueba corresponde a una pista negra cuadrada con dimensiones  $110 \times 110 [cm^2]$  sobre la cual diferentes vehículos cambian su posición a lo largo del tiempo. Por lo tanto, una secuencia de video de la escena registra el movimiento de los vehículos sobre la pista.

**Figura 9:** Montaje experimental para obtener los videos de prueba.



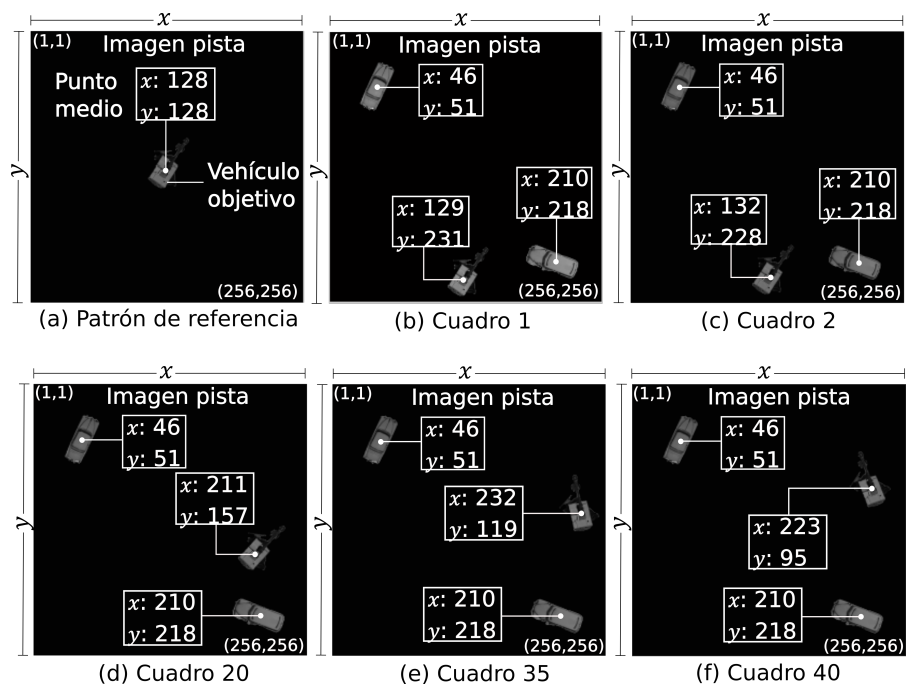
Luego, los datos obtenidos son enviados a una computadora como entrada al algoritmo de PR. Este algoritmo consta de tres partes: el cálculo de las medidas

comprimidas, la estimación de la correlación cruzada mediante un método de reconstrucción y la localización del objetivo. Primero, las medidas comprimidas son calculadas por (2.1) para simular un sistema óptico compresivo que adquiere cada escena o cuadro de video  $F_k$ . Una vez se obtienen las medidas, el algoritmo de reconstrucción basado en escasez GPSR la procesa para estimar la correlación cruzada entre la escena y un patrón referencia previamente almacenado. Posteriormente, un conjunto de operaciones ya conocidas estiman la ubicación del objetivo con base en los picos de correlación.

### 3.2. CONJUNTO DE DATOS SINTÉTICO

La escena de prueba fue inicialmente reflejada sobre datos sintéticos. El diseño de estos datos en la computadora consistió en crear un conjunto de imágenes escena e imágenes patrón de referencia para obtener una aproximación a priori del comportamiento del correlador respecto a la localización del objetivo. Inicialmente, imágenes en negro denominadas imágenes pista fueron creadas para representar la pista de vehículos o el fondo de la imagen. De este modo, ubicando vehículos en posiciones fijas sobre la imagen pista fueron obtenidas las escenas, también vistas como los cuadros de un video sintético. De acuerdo a lo anterior, en cada video sintético elaborado el vehículo objetivo fue desplazado exhaustivamente cuadro por cuadro como se muestra en la Fig. 10 para simular su movimiento sobre la pista y obtener un video de prueba apropiado para el seguimiento de objetivos. De manera similar, los patrones referencia fueron obtenidos fijando el vehículo objetivo sin distorsiones geométricas justo en el punto medio de la imagen pista.

Figura 10: Diseño de los datos sintéticos.



---

El patrón de referencia en la Fig. 10(a) fue representado por el vehículo objetivo sobre un fondo negro uniforme. Las Figs. 10(b), 10(c), 10(d), 10(e) y 10(f) son cuadros de video vistos como imágenes en negro con tres vehículos ubicados en diferentes posiciones. En detalle, el vehículo objetivo cambia su posición en cada cuadro para simular movimiento, mientras que los vehículos restantes mantienen una posición fija a lo largo del video

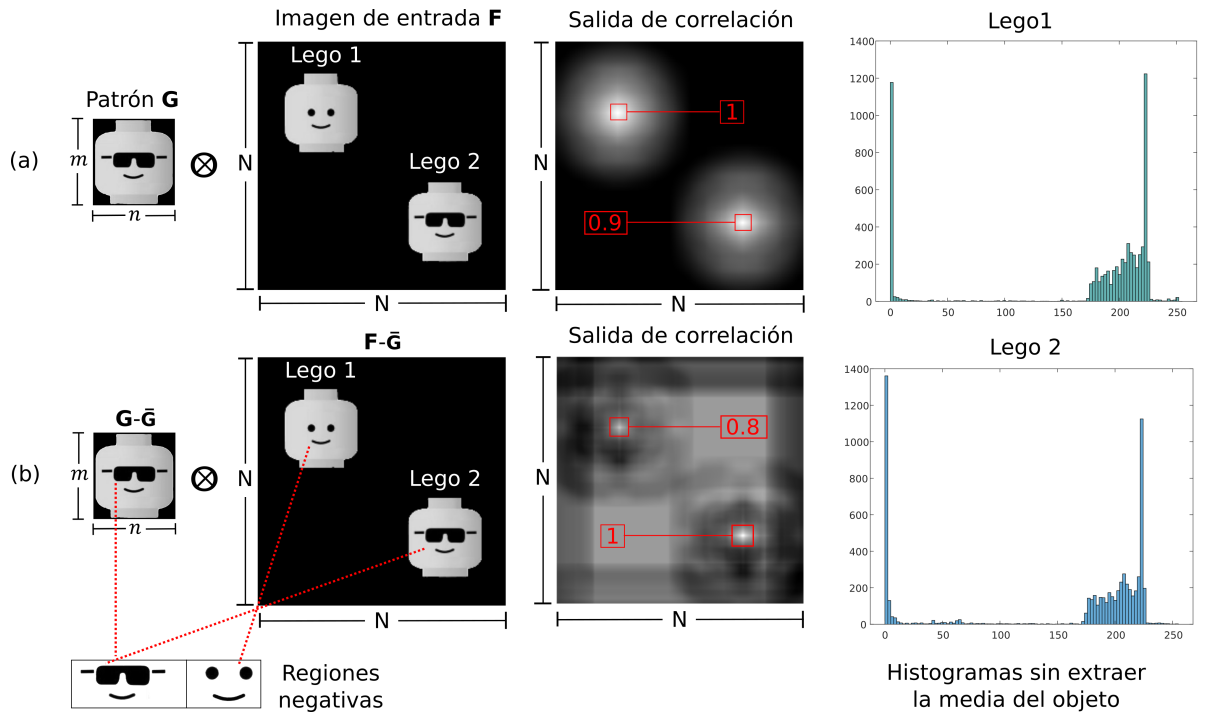
### 3.3. EMPAREJAMIENTO DE HISTOGRAMAS

En la correlación cruzada, los objetos pueden ser confundidos cuando uno de ellos predomina en intensidad. Esto se debe a que la sumatoria de multiplicaciones punto a punto será mayor entre el patrón de referencia y el objeto de mayor intensidad independientemente si corresponde o no al objetivo. Por lo tanto, la correlación cruzada para la detección de objetivos en presencia de ruido por intensidad puede calcularse restando la media del patrón de referencia a cada uno de sus píxeles y a cada píxel de la imagen de entrada. El resultado de esta operación retorna las imágenes con valores de intensidad positivos y negativos que ayudan a caracterizar la forma del objetivo. Un patrón con regiones de intensidad positivas y negativas obtendrá un pico de correlación con un objeto que presente la misma distribución de signos sobre sus píxeles. Así, los valores de intensidad en el patrón de referencia y en la imagen de entrada se multiplicarán con respecto al signo, positivos con positivos y negativos con negativos, respectivamente, produciendo un alto índice de correlación debido a una sumatoria de términos positivos. De lo contrario, si los objetos correlacionados son diferentes y no siguen la misma distribución de signos se multiplicarán valores positivos con negativos produciendo términos negativos en la sumatoria que reducirían el índice de correlación. El valor medio de un patrón de referencia  $G$  es calculado como el promedio de los valores de intensidad en sus píxeles

$$\tilde{G} = \frac{\sum_{i=0}^{m-1} \sum_{j=0}^{n-1} G(i, j)}{mn} \quad (3.1)$$

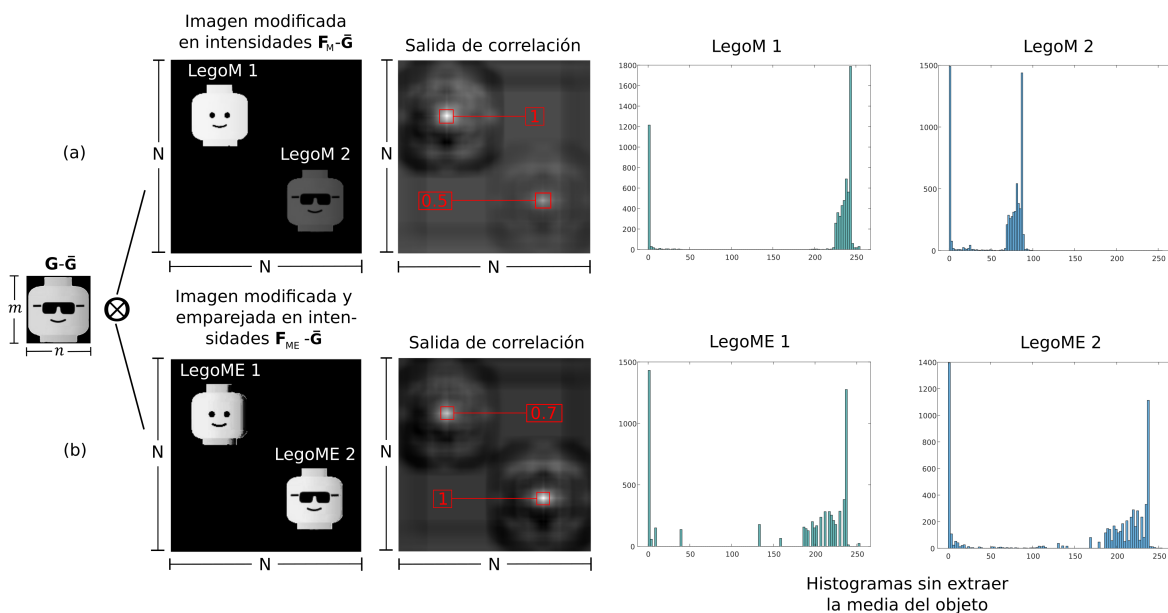
La Fig. 11 ilustra la importancia de sustraer la media del patrón a las imágenes a correlacionar para la detección del objetivo correcto. El patrón de referencia Lego 2 es localizado erróneamente cuando se trabaja con intensidades positivas, mientras que, sustrayendo la media del patrón se logra caracterizar al objetivo obteniendo su ubicación correcta. La caracterización se da por una distribución de signos positivos y negativos sobre el patrón que sugiere una alta similitud cuando se compara contra otros objetos con la misma distribución de signos. Esto sugiere sustraer la media del patrón sobre la imagen de entrada para obtener valores positivos y negativos que permitan una comparación justa. Recuerde que la ubicación estimada del objetivo es obtenida como la posición del pico de correlación más alto.

**Figura 11:** Sustracción de la media para la detección de objetivos.



En la Fig. 11(a) se muestra el resultado de correlación entre el patrón Lego 2 y la imagen de entrada  $F$ . El pico de correlación más alto detecta al objeto equivocado Lego 1 ya que este exhibe mayor intensidad alrededor de los ojos, en donde Lego 2 contiene píxeles de baja intensidad que corresponden a las gafas oscuras del objeto. El problema es solucionado sustrayendo la media  $\bar{G}$  tanto al patrón de referencia como a la imagen de entrada. La Fig 11(b) presenta esta aproximación sobre dos objetos con histogramas similares. El resultado de correlación permitió detectar al objeto correcto Lego 2 convirtiendo la desventaja de los píxeles con menor intensidad sobre la región de las gafas en la clave para discriminarlo del objeto Lego 1. Ahora, estos píxeles serán negativos y su multipliquen punto a punto con los píxeles correspondientes en Lego 1 producirán términos negativos que reducirán la magnitud del pico de correlación. Sin embargo, en presencia de objetos con histogramas diferentes la sustracción de la media puede no ser suficiente para la detección exitosa del objetivo. Un ejemplo de detección errónea del objetivo aún cuando se ha sustraído la media es presentado en la Fig. 12. Por tal motivo, el conjunto de datos sintético y el conjunto de datos real fueron sometidos a un emparejamiento de histogramas (HM, del inglés, Histogram Matching) para solucionar el problema descrito. El método consiste en la transformación de una imagen para que su histograma coincida con un histograma especificado [37]. En este caso, la transformación se llevará a cabo entre cada uno de los objetos presentes en la imagen de entrada para que logren obtener un histograma similar al patrón de referencia.

**Figura 12:** Emparejamiento de histogramas para la detección de objetivos.



La Fig. 12(a) muestra la salida de correlación para detectar a Lego 2 cuando los histogramas de los objetos sobre la imagen de entrada  $F_M$  han sido modificados. El objeto LegoM 1 se obtuvo aumentando la intensidad de algunos píxeles sobre el objeto Lego 1 y el objeto LegoM 2 se obtuvo reduciendo el rango de intensidades en el objeto Lego 2. Estas operaciones pueden observarse sobre el histograma de LegoM 1 que exhibe una mayor concentración en las intensidades más altas y el histograma LegoM 2 que muestra una reducción en el rango de intensidades. Observe que a pesar de sustraer la media del patrón, la diferencia obtenida en intensidades entre LegoM 1 y LegoM 2 produce términos de suma-toria con valores muy altos que compensan la pérdida en las regiones negativas. En consecuencia se obtiene el pico de correlación sobre el objeto equivocado LegoM 1. De acuerdo a lo anterior, se observó que la operación de sustraer la media tiene un mayor efecto cuando los objetos sobre las escena tienen intensidades similares. Por lo tanto, se optó por estimar la correlación cruzada emparejando los histogramas de los objetos sobre la imagen de entrada como se muestra en la Fig. 12(b). Los histogramas de los objetos LegoME 1 y LegoME 2 corresponden al resultado del proceso de emparejamiento de histogramas entre los objetos LegoM 1 y LegoM 2 con el patrón de referencia  $G$ , respectivamente. El resultado de correlación luego de las operaciones de emparejamiento de los histogramas y sustracción de la media mostró su utilidad para detectar al objetivo exitosamente.

### 3.4. MÉTRICAS DE RENDIMIENTO

La calidad de las reconstrucción de correlación fue evaluada en términos de las métricas: relación pico señal a ruido (PSNR, del inglés Peak Signal-to-Noise Ratio), Error de seguimiento y tasa de éxito. Las métricas fueron analizadas calcu-

lando un promedio a lo largo del video para analizar el resultado final sobre el seguimiento de un objeto.

**3.4.1 Relación pico señal a ruido** El PSNR esta definido como la relación entre la máxima energía posible de una señal y la energía de ruido corrupto que afecta la fidelidad de su representación. El PSNR aquí es usado para medir la calidad de las reconstrucciones de correlación en cada cuadro de video. Por lo tanto, El PSNR en un video con un número de cuadros  $L$  es estimado como el promedio de valores  $PSNR_k$  en  $[dB]$  para  $k = 1, 2, \dots, L$ ,

$$PS\tilde{N}R = \frac{\sum_{k=0}^{L-1} PSNR_k}{L},$$

$$PSNR_k = 10 \log_{10} \left( \frac{\max(\mathbf{R}_k)^2}{MSE(\mathbf{R}_k, \hat{\mathbf{R}}_k)} \right), \quad (3.2)$$

donde  $PS\tilde{N}R$  es el valor de PSNR promedio para un un video,  $PSNR_k$  es el valor de PSNR obtenido en el cuadro  $k$ ,  $\mathbf{R}_k$  es la matriz de correlación real obtenida en el cuadro  $k$ ,  $\max(\mathbf{R}_k)$  es la máxima energía posible de  $\mathbf{R}_k$  and  $MSE$  es el error cuadrático medio entre  $\mathbf{R}_k$  y su aproximación  $\hat{\mathbf{R}}_k$ .

**3.4.2 Error de seguimiento** La métrica tradicional para verificar la exactitud en pixeles del seguimiento de objetos es el error de seguimiento. Esta métrica es estimada en cada cuadro de video  $\mathbf{F}_k$  por la distancia euclidiana entre la ubicación estimada  $\hat{C}_k$  y la ubicación real  $C_k$  del objetivo a partir de los picos de correlación.

$$\epsilon_k = \|\hat{C}_k - C_k\|_2, \quad (3.3)$$

donde  $\epsilon_k$  es el error de seguimiento obtenido en el cuadro  $k$  y las tuplas  $C_k$  y  $\hat{C}_k$  contienen la posición real y estimada del objetivo respectivamente. Cada una de estas tupas contiene dos índices que corresponden a las posiciones vertical y horizontal del objetivo sobre la imagen de entrada. El error de seguimiento de seguimiento  $\hat{\epsilon}$  en un video es calculado como el promedio de errores  $\epsilon_k$  obtenidos en cada uno de sus cuadros

$$\hat{\epsilon} = \frac{\sum_{k=0}^{L-1} \epsilon_k}{L}. \quad (3.4)$$

La métrica de error de seguimiento es una medida de exactitud que evalúa la habilidad del CHF derivado de la fase para discriminar al objetivo y obtener el pico de correlación lo mas cercano a su centroide. Además, evalúa el método propuesto para recuperar el pico objetivo en su posición real y con la misma discriminación respecto a otros picos de correlación. Las posiciones reales fueron calculadas como el centroide del objetivo, lo que requirió su segmentación en cada uno de los cuadros del video. Generalmente, se aceptan errores de seguimiento que favorezcan a aquellas posiciones estimadas que se encuentran dentro del objetivo ya que en estos casos se considera exitosa su localización.

---

**3.4.3 Tasa de éxito** La tasa de éxito representa el porcentaje de ubicaciones correctas a lo largo del video. Por lo tanto, su valor para un número de éxitos  $B$  es calculado como

$$\varsigma = \frac{B}{L} (100) \% \quad (3.5)$$

En términos de análisis acerca de la utilidad de la aplicación para tareas de seguimiento de objetivos, esta métrica podría verse como la más relevante ya que provee una medida de cuántas veces fue posible la localización del objetivo aceptando una franja de error. Esto puede ser interpretado pensando en que la ubicación estimada  $\hat{C}_j$  no debe ser estrictamente exacta a la ubicación real  $C_j$ , si no que es suficiente con obtener una posición estimada que se encuentre justo sobre el objetivo. Para verificar si la posición estimada cumple o no esta condición se optó por aproximar el área del objetivo a una región cuadrada de píxeles de mayor tamaño concéntrica a la ubicación real del objetivo. Esta región aproximada permitió clasificar las posiciones en exitosas y no exitosas para calcular la métrica de tasa de éxito. Además, califico como exitosas las posiciones cercanas al objetivo con un error de seguimiento ocasionado por el choque entre vehículos.

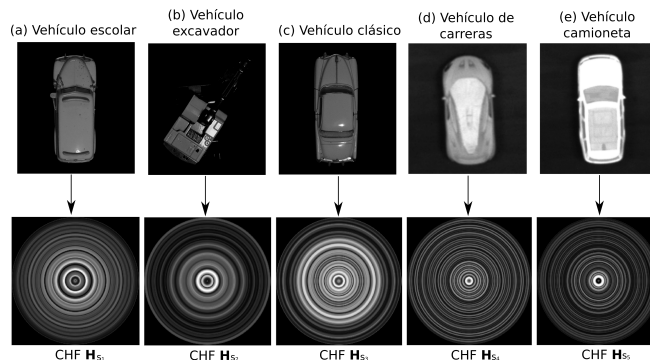
## 4. SIMULACIONES Y RESULTADOS

En el trabajo desarrollado, se llevaron a cabo experimentos con datos reales y sintéticos para evaluar la eficacia de la aproximación de recuperación propuesta sobre la matriz de correlación cruzada. Los picos de correlación fueron usados para obtener la ubicación del objetivo en cada cuadro de un video y así lograr como salida su ruta en el tiempo. Los resultados de seguimiento fueron obtenidos a partir de tres secuencias de video sintéticos y dos secuencias de video reales con  $256 \times 256$  pixeles de resolución espacial. El objetivo de las métricas calculadas es verificar la posibilidad de estimar la matriz de correlación con una calidad comparable a la correlación tradicional para tareas de localización y detección de objetos reduciendo el número de medidas adquiridas. Las reconstrucciones de correlación a partir de estas capturas usando el algoritmo GPSR fueron evaluadas en términos de las métricas PSNR, error de seguimiento y tasa de éxito. Además, los resultados de correlación obtenidos fueron comparados contra el método tradicional de correlación para observar el impacto en la calidad cuando se adquiere un menor número de medidas.

### 4.1. DESCRIPCIÓN DEL CONJUNTO DE DATOS DE PRUEBA

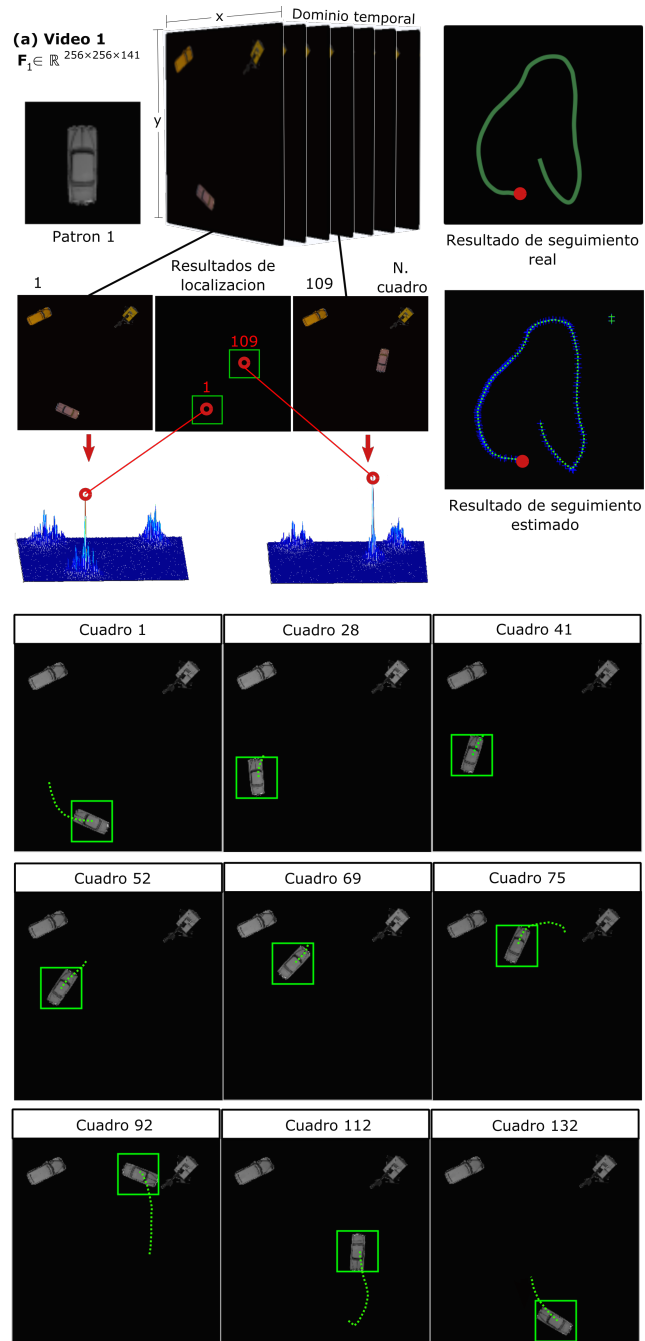
La base de datos de prueba consiste de cinco secuencias de videos a escala escala de grises, cinco patrones de referencia y sus CHF's respectivos. Los datos sintéticos provenientes a partir del diseño propuesto en Fig. 10 exhiben tres videos  $F_1 \in \mathbb{R}^{256 \times 256 \times 141}$ ,  $F_2 \in \mathbb{R}^{256 \times 256 \times 135}$  y  $F_3 \in \mathbb{R}^{256 \times 256 \times 135}$  creados para realizar seguimiento de los vehículos escolar, retroexcavadora y clásico, respectivamente. Los otros dos videos  $F_4 \in \mathbb{R}^{256 \times 256 \times 100}$  y  $F_5 \in \mathbb{R}^{256 \times 256 \times 100}$  con los vehículos camioneta y de carreras corresponden a la data real adquirida por el montaje experimental en Fig. 9. Los patrones de referencia y los cuadros del video fueron redimensionados al mismo tamaño cuadrado, ya que el diseño de CHF's retorna solo filtros de tamaño cuadrado. En cada video, un patrón  $i$  es seguido en presencia de ruido usando un CHF derivado de la fase  $H_{s_i}$  con  $i = 1, 2, 3, 4$  o  $5$ . La Fig. 13 ilustra los patrones de referencia usados en las simulaciones y sus CHF's con dimensiones espaciales de  $256 \times 256$ .

**Figura 13:** Vehículos modelo usados como patrón de referencia a lo largo de sus CHF's.



## 4.2. RESULTADOS VISUALES

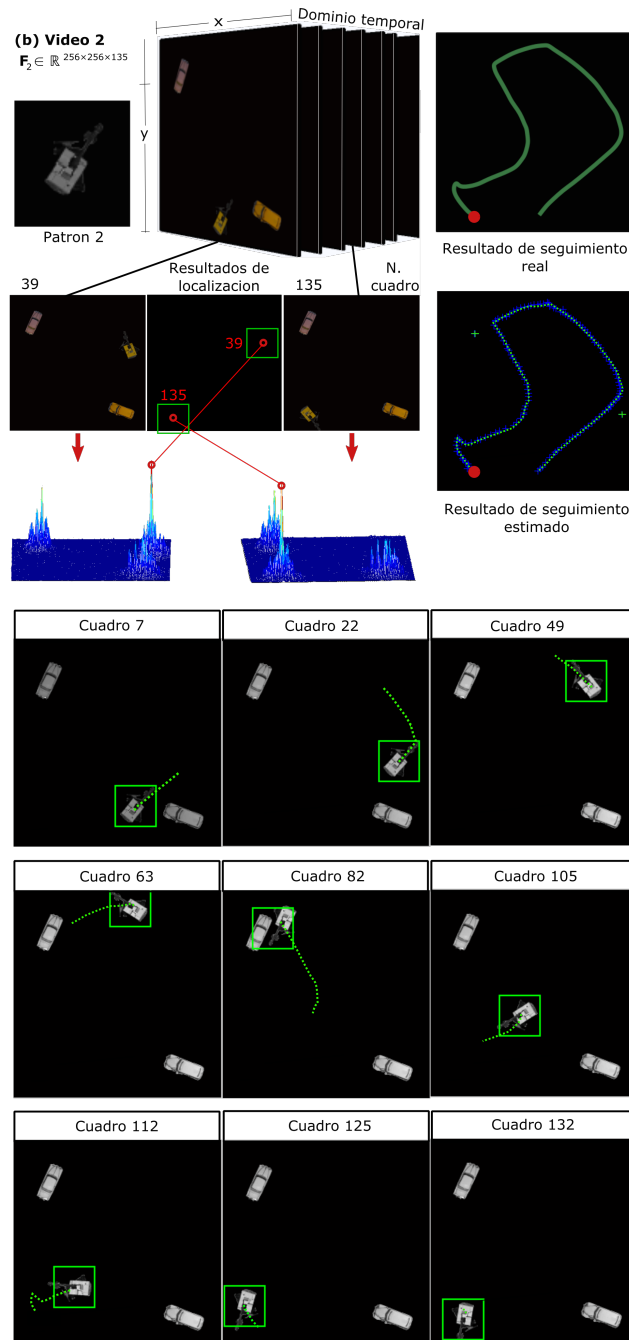
Figura 14: Conjunto de datos de prueba video 1.



La base de datos de prueba video 1 en la Fig. 14 consiste del video  $F_1$  con 141 cuadros y el CHF  $H_1$  calculado a partir de el patrón 1. Este filtro es usado para localizar el objetivo en cada cuadro de video sin importar su orientación. Los resultados de localización a partir de la matriz de correlación cruzada en los cuadros 1 y 109 son mostrados para evidenciar la precisión del pico de correlación. Además, la ubicación del objetivo puede ser visualizada en los cuadros 1, 28,

41, 52, 69, 75, 92, 112 y 132. Finalmente, las ubicaciones obtenidas a lo largo del video fueron registradas una por una en el resultado de seguimiento estimado, el cual puede ser comparado visualmente con el resultado de seguimiento real. Observe que, los errores pueden ser distinguidos como una marca apartada de la secuencia que sigue la ruta real. Específicamente, el seguimiento del vehículo escolar sobre el video 1 exhibió dos errores de detección en los cuadros 15 y 83, debido a una detección equivocada del objetivo ocasionada por el CHF.

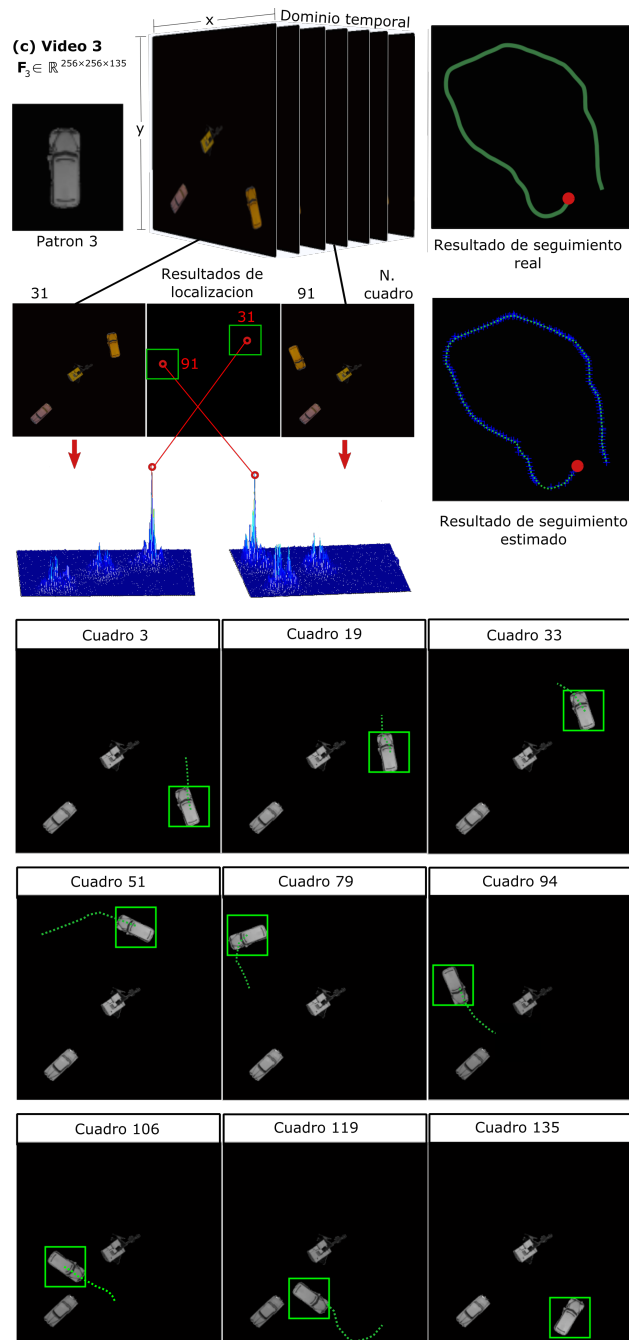
**Figura 15:** Conjunto de datos de prueba video 2.



La base de datos de prueba video 2 en Fig. 15 contiene el video  $F_2$  con 135

cuadros y el CHF  $H_2$  usado para estimar la ubicación del patrón 2 a lo largo del video. Los resultados de localización a partir de la matriz de correlación cruzada son mostrados para los cuadros 39 y 135. Además, la ubicación del objetivo es visualizada en los cuadros 7, 22, 49, 63, 82, 105, 112, 125 y 132. Observe que el resultado de seguimiento estimado sobre  $F_2$  exhibe dos errores evidentes en la ubicación del objetivo en los cuadros 9 y 120, debido a una detección equivocada del objetivo ocasionada por el CHF.

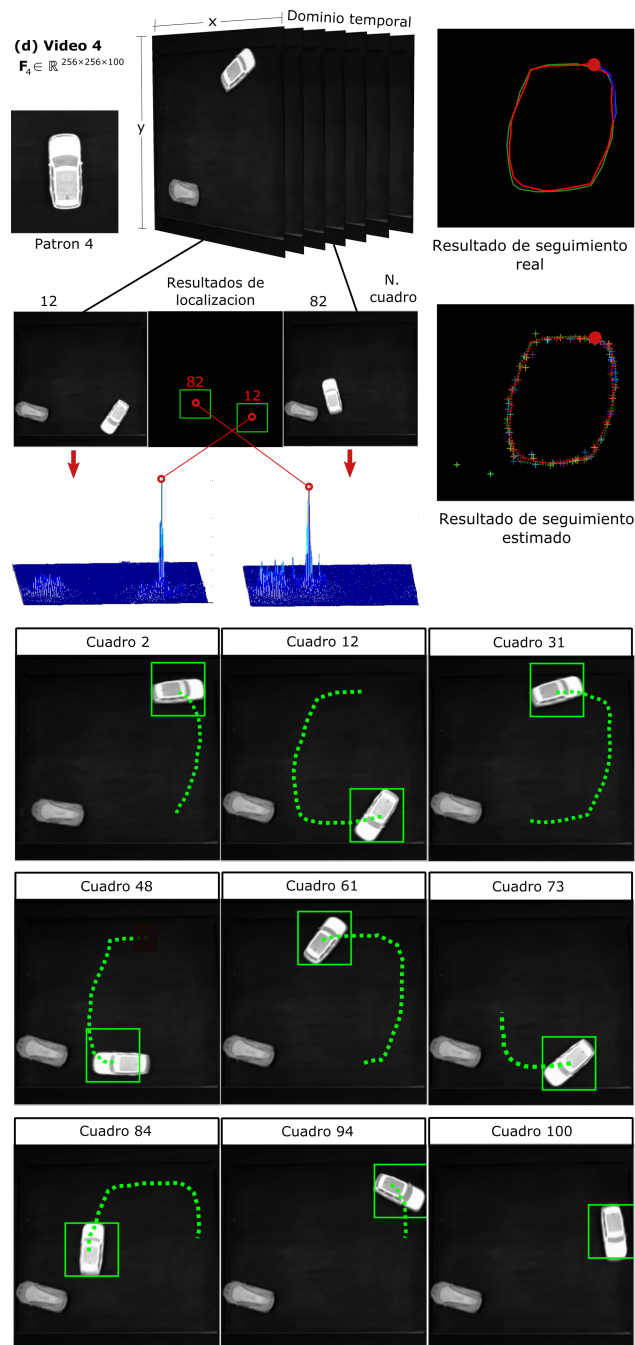
**Figura 16:** Conjunto de datos de prueba video 3.



La base de datos de prueba video 3 en Fig. 16 consiste del video  $F_3$  con 135 cua-

dros y el CHF  $H_3$  usado para estimar la ubicación del patrón 3 en cada cuadro de video. Los resultados de localización a partir de la matriz de correlación cruzada son mostrados para los cuadros 31 y 91. Además, la ubicación del objetivo es visualizada en los cuadros 3, 19, 33, 51, 79, 94, 106, 119 y 135. Observe que, el resultado de seguimiento estimado sobre  $F_3$  no exhibe errores de localización.

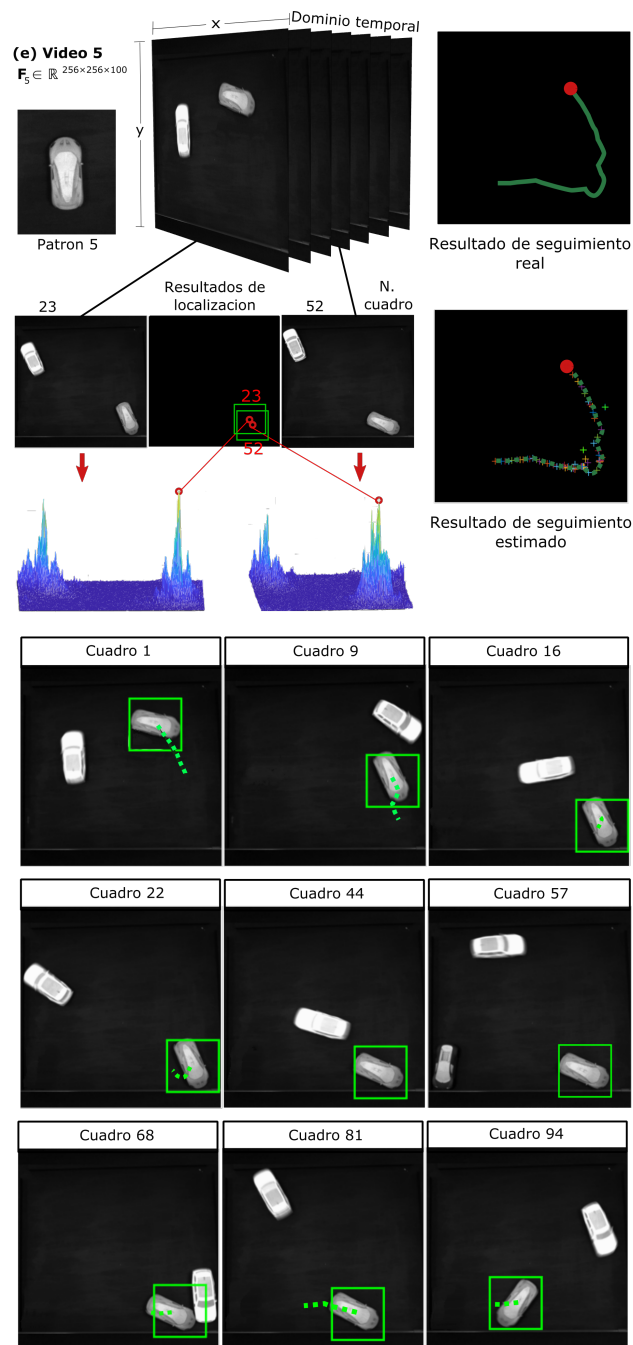
**Figura 17:** Conjunto de datos de prueba video 4.



Los videos  $F_4$  y  $F_5$  con 100 cuadros cada uno y almacenados en las bases de datos de prueba video 4 y video 5, respectivamente fueron obtenidos por el montaje

experimental mostrado en la Fig. 9. De forma similar, sus CHF<sub>4</sub> y CHF<sub>5</sub> fueron calculados a partir del patrón 4 y el patrón 5 capturados como datos reales respectivamente. En la Fig. 17 el patrón 4 es seguido a lo largo del video  $F_4$  y resultados de localización a partir de la matriz de correlación son mostrados para los cuadros 12 y 82. Además, la ubicación del objetivo es visualizada en los cuadros 2, 12, 31, 48, 61, 73, 84, 94 y 100. Observe que el resultado de seguimiento estimado sobre  $F_4$  exhibe dos errores evidentes en la ubicación del objetivo en los cuadros 25 y 99, debido a la reducción en magnitud del pico de correlación objetivo en el proceso de reconstrucción.

**Figura 18:** Conjunto de datos de prueba video 5.



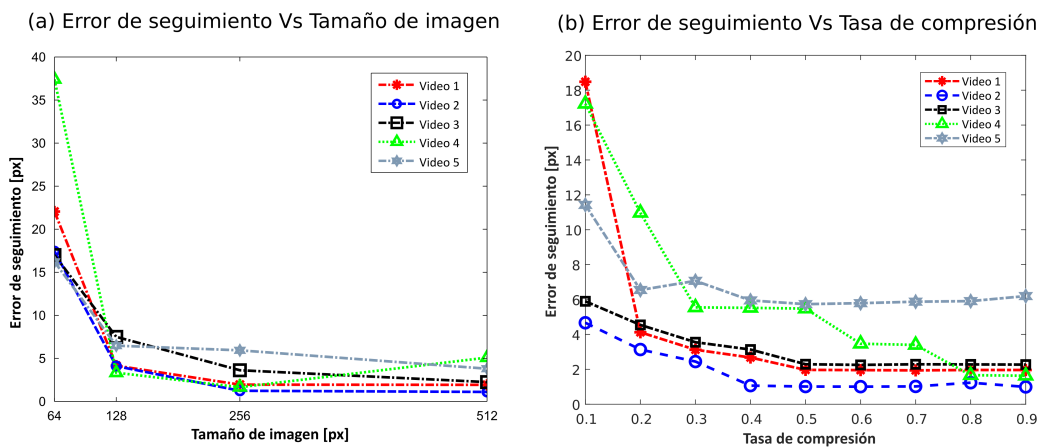
En la Fig. 18 el patrón 5 es seguido a lo largo del vídeo  $F_5$  y resultados de localización usando la matriz de correlación cruzada son mostrados para los cuadros 23 y 52. Además, la ubicación del objetivo es visualizada en los cuadros 1, 9, 16, 22, 44, 57, 68, 81 y 94. Observe que el resultado de seguimiento estimado sobre  $F_5$  exhibe tres errores evidentes en la ubicación del objetivo en los cuadros 12, 16, y 64, debido a errores en el CHF y a la reducción en magnitud del pico objetivo cuando se recupera la matriz de correlación cruzada. Las simulaciones llevadas a cabo sobre los cinco videos en las Figs. 14, 15, 16, 17 y 18 corresponden a el seguimiento del objetivo usando la estimación de correlación propuesta con una tasa de compresión del 40 %, lo cual es equivalente a capturar y procesar menos de la mitad de la información comparativamente a la imagen de entrada.

### 4.3. ANÁLISIS DE LAS MÉTRICAS DE CALIDAD

Las métricas de calidad fueron evaluadas contra los parámetros de tasa de compresión y tamaño de la imagen. Las pruebas respecto al tamaño de la imagen fueron realizadas variando su valor desde  $64 \times 64$  hasta  $512 \times 512$  con una tasa de compresión fija del 40 % y las pruebas respecto a la tasa de compresión fueron llevadas a cabo variando su valor desde 0,1 hasta 0,9, que corresponde a usar desde el 10 % hasta el 90 % de toda la información en la imagen de entrada respectivamente.

**4.3.1 Resultados de error de seguimiento** Las ubicaciones reales de los objetivos fueron rigurosamente calculadas cuadro por cuadro para estimar el error de seguimiento producido por los métodos tradicional y propuesto a lo largo de cada video. La Fig. 19 muestra el error de seguimiento promedio por video definido en (3.3) variando los parámetros de la tasa de compresión y el tamaño de la imagen.

**Figura 19:** Evaluación de la métrica error de seguimiento usando el método propuesto.

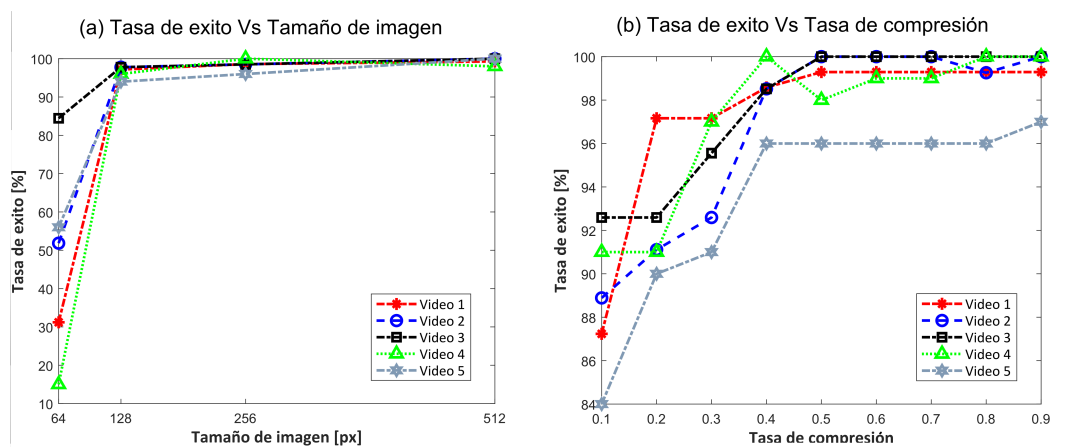


Las líneas constantes a lo largo de la variable independiente pueden obedecer

a un error introducido por la expresión de correlación en (1.12) y no por la aproximación de reconstrucción obtenida. En contraste, los errores de reconstrucción son muy evidentes cuando la cantidad de información es lo suficientemente reducida. La fig. 19(a) muestra la exactitud para ubicar el objetivo a diferentes tamaños de resolución espacial. Trabajar con imágenes de baja resolución aumentó considerablemente el error de seguimiento e incluso ocasionó la detección errónea del objetivo. la Fig. 19(b) refleja el impacto en la calidad al reducir la tasa de compresión hasta un 10 % en donde la pérdida de información ocasiona la equivocación en el seguimiento del objetivo.

**4.3.2 Resultados de tasa de éxito** Las regiones de tolerancia a error fueron estimadas de las ubicaciones reales para delimitar un área cuadrada sobre la cual se considera está ubicado el objetivo. Estas regiones están descritas por las coordenadas superior-izquierda e inferior-derecha del cuadrado. Si la ubicación del objetivo estimada se encuentra adentro de la región de tolerancia a error, la localización es considerada exitosa. Por lo tanto, buenos resultados en esta métrica indican la habilidad para localizar al objetivo con una exactitud aceptable. La Fig. 20 ilustra la tasa de éxito por video calculada en (3.5) como una función de la tasa de compresión y el tamaño de la imagen.

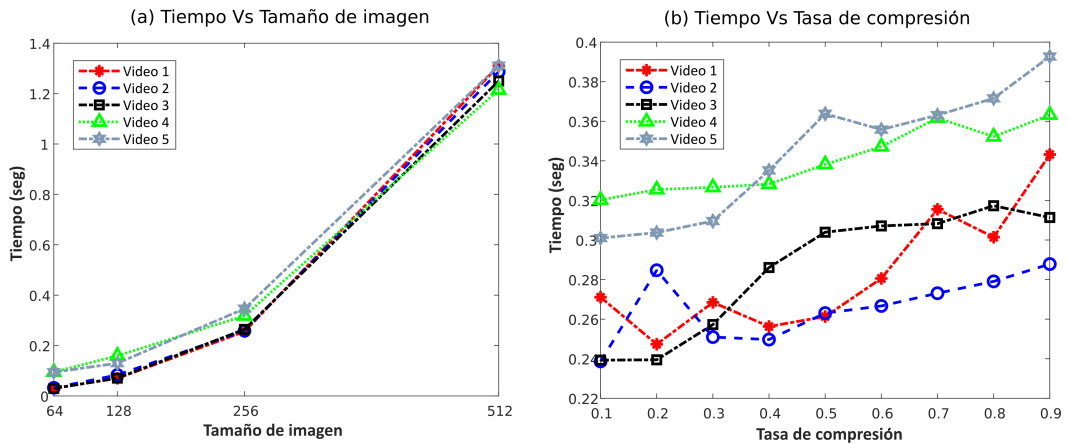
**Figura 20:** Análisis de la métrica tasa de éxito usando el método propuesto.



Los resultados obtenidos en la Fig. 20(a) muestran valores de tasa de éxito por arriba de un 95 % para un tamaño de la imagen de  $128 \times 128$ ,  $256 \times 256$  y  $512 \times 512$ . El valor de esta métrica disminuye para un tamaño de la imagen de  $64 \times 64$ , surgiendo la necesidad de trabajar con imágenes a un mínimo de resolución espacial donde los detalles básicos de los objeto estén presentes. En la Fig. 20(b) los resultados logrados sobre las imágenes con una resolución espacial aceptable exhibieron una correcta localización en la mayoría de la veces capturando a una tasa de compresión de hasta el 20 %.

**4.3.3 Resultados de tiempo de ejecución** Los tiempos de ejecución registrados en la Fig. 21 corresponden al tiempo que toma: calcular las medidas comprimidas, estimar la correlación cruzada, y localizar el objetivo, donde las reconstrucciones fueron obtenidas usando cien iteraciones del algoritmo GPSR.

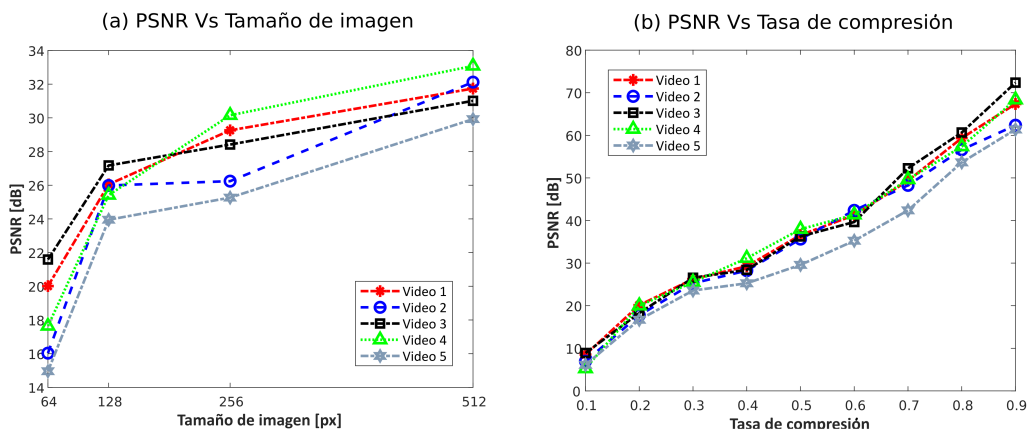
**Figura 21:** Complejidad temporal del algoritmo de reconocimiento de patrones basado en compresión.



La Fig. 21 muestra los promedios de tiempo de ejecución por video variando los parámetros de tasa de compresión y tamaño de la imagen.

**4.3.4 Resultados de PSNR** Los valores de PSNR fueron calculados como el valor promedio por video especificado en la ecuación 3.2. Esta métrica fue estimada para analizar la calidad de las reconstrucciones de correlación de una manera general. Por lo tanto, un valor alto en esta métrica indica que la matriz de correlación fue recuperada con una buena aproximación pero no garantiza eficacia en la detección de un objetivo.

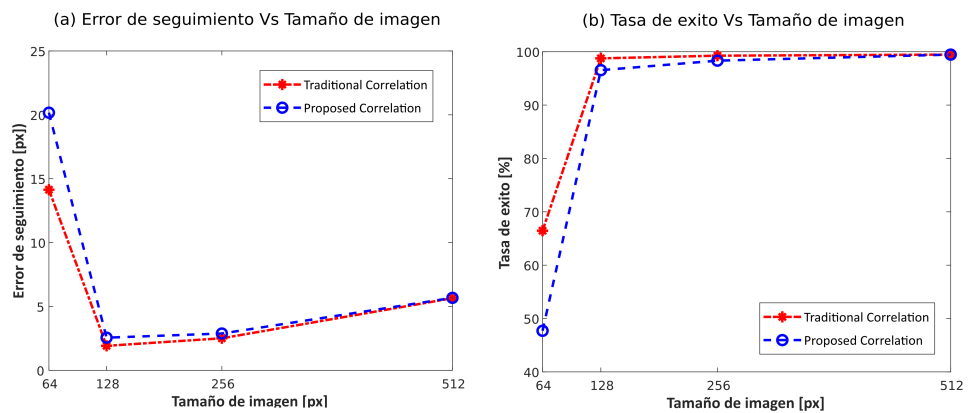
**Figura 22:** Análisis de la métrica PSNR sobre el método propuesto.



La figura 22 muestra la calidad de las reconstrucciones de correlación obtenidas adquiriendo la imagen a diferentes tasas de captura y variando su tamaño. Note en la figura 22(a) que para un tamaño de imagen de  $256 \times 256$  se logra reconstruir la correlación con una calidad de hasta 30 y 27[ $dB$ ] en términos de PSNR capturando un 40 % y un 30 % de la información respectivamente. Además, se observa en la figura 22(b) una reducción considerable en la calidad adquiriendo la información a una tasa de captura menor al 30 %.

**4.3.5 Resultados de comparación** Para evaluar el impacto de capturar menos medidas sobre la escena, los métodos de correlación tradicional y propuesto fueron comparados en términos de las métricas error de seguimiento y tasa de éxito variando el tamaño de la imagen de entrada. Los valores de error de seguimiento y tasa de éxito en la Fig. 23 fueron calculados como el promedio de resultados obtenidos en los cinco videos. Una observación crucial, es la calidad comparable entre la correlación propuesta capturando un 40 % de mediciones y la correlación tradicional que captura la imagen completa. La mayor diferencia se evidencia sobre las imágenes con un tamaño de  $64 \times 64$ , donde la baja resolución espacial no permite diferenciar los objetos debido a la pérdida del detalle.

**Figura 23:** Comparación entre las métricas de calidad usando el método de correlación tradicional y propuesto.



## 5. CONCLUSIONES

En este trabajo ha sido introducido un método para estimar la matriz de correlación cruzada a partir de medidas capturadas por un sistema de adquisición compresivo. El procedimiento clave consistió en incluir un conjunto de operadores lineales en el modelo de adquisición compresiva para lograr establecer la representación escasa de la imagen de entrada como la correlación cruzada. De este modo, sin afectar el modelo de adquisición compresiva fue posible teóricamente adquirir información acerca de la correlación. Luego, el problema inverso fue solucionado minimizando las normas  $l_1$  y  $l_2$  mediante el algoritmo GPSR, el cual mostró recuperar la correlación con una calidad aceptable.

Las simulaciones sobre tareas de localización y detección de objetos pudieron ser llevadas a cabo con una reducción en el número de capturas de hasta un 40 %, lo que sugiere el uso de información innecesaria por métodos que capturan la imagen de tamaño completo. Además, el método propuesto muestra una forma elaborada de evitar el cálculo de la correlación en el dominio comprimido, lo que suele ser una tarea difícil.

En el trabajo futuro se incluye el diseño de una arquitectura de adquisición compresiva real para comprobar lo planteado matemáticamente con las medidas comprimidas simuladas. Aprovechando la implementación de la arquitectura también se desea calcular la correlación en diferentes bandas del espectro para incluir un segundo criterio en la localización y detección de objetivos. Por otra parte, la correlación es también variante ante cambios significativos de intensidad en la imagen, lo que requirió realizar una comparación de intensidades entre los objetos en ella. Sin embargo, esta operación debe ser redefinida para realizarse en el dominio comprimido o bien ser aplicada directamente en el filtro de correlación.

## Referencias

- [1] J. Tyckowski, C. Kyrtos, T. Davies, C. Hopson, F. Breynaert, and P. Bonduel. Object detection by pattern recognition, November 28 2000. US Patent 6,154,149.
- [2] Min-Seok Choi and Whoi-Yul Kim. A novel two stage template matching method for rotation and illumination invariance. *Pattern recognition*, 35(1):119–129, 2002.
- [3] Du-Ming Tsai and Chien-Ta Lin. Fast normalized cross correlation for defect detection. *Pattern Recognition Letters*, 24(15):2625–2631, 2003.
- [4] Josué Alvarez Borrego. Estadística no lineal de superficies marinas y reconocimiento de objetos y estructuras mediante el uso de correlación no lineal e índices espectrales, 2011.
- [5] BVK Vijaya Kumar, Joseph A Fernandez, Andres Rodriguez, and Vishnu Nares Boddeti. Recent advances in correlation filter theory and application. In *Optical Pattern Recognition XXV*, volume 9094, page 909404. International Society for Optics and Photonics, 2014.
- [6] E Oran Brigham and EO Brigham. *The fast Fourier transform and its applications*, volume 1. prentice Hall Englewood Cliffs, NJ, 1988.
- [7] David Casasent and Demetri Psaltis. Position, rotation, and scale invariant optical correlation. *Applied optics*, 15(7):1795–1799, 1976.
- [8] Bahram Javidi. *Image recognition and classification: algorithms, systems, and applications*. CRC Press, 2002.
- [9] Qiong Wang and Jingyu Yang. Eye detection in facial images with unconstrained background. *Journal of Pattern Recognition Research*, 1(1):55–62, 2006.
- [10] Daniel Y Ts'o, Charles D Gilbert, and Torsten N Wiesel. Relationships between horizontal interactions and functional architecture in cat striate cortex as revealed by cross-correlation analysis. *Journal of neuroscience*, 6(4):1160–1170, 1986.
- [11] Kay Gr̃unewald, Prashant Desai, Dennis C Winkler, J Bernard Heymann, David M Belnap, Wolfgang Baumeister, and Alasdair C Steven. Three-dimensional structure of herpes simplex virus from cryo-electron tomography. *Science*, 302(5649):1396–1398, 2003.
- [12] Emmanuel J Candès and Michael B Wakin. An introduction to compressive sampling. *IEEE signal processing magazine*, 25(2):21–30, 2008.

- 
- [13] Jia Yu. The analysis of compressive sensing theory. In *P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC), 2014 Ninth International Conference on*, pages 157–160. IEEE, 2014.
- [14] Mário AT Figueiredo, Robert D Nowak, and Stephen J Wright. Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems. *IEEE Journal of selected topics in signal processing*, 1(4):586–597, 2007.
- [15] Keinosuke Fukunaga. *Introduction to statistical pattern recognition*. Academic press, 2013.
- [16] Nasser M Nasrabadi. Pattern recognition and machine learning. *Journal of electronic imaging*, 16(4):049901, 2007.
- [17] J Alba and Jesús Cid. Reconocimiento de patrones. *Carlos III, Madrid, Universidad Carlos III, Madrid. Retrived from <http://www.gts.tsc.uvigo.es/pi/Reconocimiento.pdf>*, 2006.
- [18] Brijesh Verma and John Zakos. A computer-aided diagnosis system for digital mammograms based on fuzzy-neural and feature extraction techniques. *IEEE transactions on information technology in biomedicine*, 5(1):46–54, 2001.
- [19] Ganesh K Venayagamoorthy, Viresh Moonasar, and Kumbes Sandrasegaran. Voice recognition using neural networks. In *Communications and Signal Processing, 1998. COMSIG'98. Proceedings of the 1998 South African Symposium on*, pages 29–32. IEEE, 1998.
- [20] Kannappan Palaniappan, Filiz Bunyak, Praveen Kumar, Ilker Ersoy, Stefan Jaeger, Koyeli Ganguli, Anoop Haridas, Joshua Fraser, Raghuveer M Rao, and Guna Seetharaman. Efficient feature extraction and likelihood fusion for vehicle tracking in low frame rate airborne video. In *Information fusion (FUSION), 2010 13th Conference on*, pages 1–8. IEEE, 2010.
- [21] Youguo Pi, Wenzhi Liao, Mingyou Liu, and Jianping Lu. Theory of cognitive pattern recognition. In *Pattern Recognition Techniques, Technology and Applications*. InTech, 2008.
- [22] Yongbum Lee, Takeshi Hara, Hiroshi Fujita, Shigeki Itoh, and Takeo Ishigaki. Automated detection of pulmonary nodules in helical ct images based on an improved template-matching technique. *IEEE Transactions on medical imaging*, 20(7):595–604, 2001.
- [23] Jochen Böhm, Achilleas S Frangakis, Reiner Hegerl, Stephan Nickell, Dieter Typke, and Wolfgang Baumeister. Toward detecting and identifying macromolecules in a cellular context: template matching applied to electron tomograms. *Proceedings of the National Academy of Sciences*, 97(26):14245–14250, 2000.

- 
- [24] Abhishek Nagar, Karthik Nandakumar, and Anil K Jain. Biometric template transformation: a security analysis. In *Media Forensics and Security II*, volume 7541, page 754100. International Society for Optics and Photonics, 2010.
- [25] Jonathan Y Stein. *A Computer Science Perspective*, chapter Convolution and Correlation. Wiley, 2000.
- [26] John P Lewis. Fast normalized cross-correlation. In *Vision interface*, volume 10, pages 120–123, 1995.
- [27] A Vander Lugt, FB Rotz, and A Klooster Jr. Character reading by optical spatial filtering. *Optical and Electro-Optical Information Processing*, pages 125–141, 1965.
- [28] H Fuentes and O Gonzalez. Implementation of a 2d real-time correlator based on a dsp. Technical report, Internal Report, Universidad Industrial de Santander, 2003.
- [29] Ping Zi-Liang and Erik Dalsgaard. Synthetic circular-harmonic phase-only filter for shift, rotation, and scaling-invariant correlation. *Applied optics*, 34(32):7527–7531, 1995.
- [30] Ph Réfrégier. Optical pattern recognition: optimal trade-off circular harmonic filters. *Optics communications*, 86(2):113–118, 1991.
- [31] Oscar Gualdron and Henri H Arsenault. Phase derived circular harmonic filter. *Optics communications*, 104(1-3):32–34, 1993.
- [32] Emmanuel J Candès et al. Compressive sampling. In *Proceedings of the international congress of mathematicians*, volume 3, pages 1433–1452. Madrid, Spain, 2006.
- [33] Stephen J Wright, Robert D Nowak, and Mário AT Figueiredo. Sparse reconstruction by separable approximation. *IEEE Transactions on Signal Processing*, 57(7):2479–2493, 2009.
- [34] Manya V Afonso, José M Bioucas-Dias, and Mário AT Figueiredo. An augmented lagrangian approach to the constrained optimization formulation of imaging inverse problems. *IEEE Transactions on Image Processing*, 20(3):681–695, 2011.
- [35] Joel A Tropp and Anna C Gilbert. Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Transactions on information theory*, 53(12):4655–4666, 2007.
- [36] BVK Vijaya Kumar and Laurence Hasebrook. Performance measures for correlation filters. *Applied optics*, 29(20):2997–3006, 1990.
- [37] Rafael C Gonzalez and Richard E Woods. Object recognition. *Digital image processing, 3rd ed. Pearson*, pages 861–909, 2008.

## BIBLIOGRAFÍA

ARSENAULT, Henri y SHENG, Yunlong. Properties of the circular harmonic expansion for rotation-invariant pattern recognition. *Applied Optics*, 1986, vol. 25, no 18, p. 3225-3229.

BARANIUK, Richard, et al. An introduction to compressive sensing. *Connexions e-textbook*, 2011.

CASASENT, David y PSALTIS, Demetri. Position, rotation, and scale invariant optical correlation. *Applied optics*, 1976, vol. 15, no 7, p. 1795-1799.

CHOI, Min-Seok; KIM y Whoi-Yul. A novel two stage template matching method for rotation and illumination invariance. *Pattern recognition*, 2002, vol. 35, no 1, p. 119-129.

COOLEY, James; LEWIS, Peter y WELCH, Peter. The fast Fourier transform and its applications. *IEEE Transactions on Education*, 1969, vol. 12, no 1, p. 27-34.

FIGUEIREDO, Mário; NOWAK, Robert y WRIGHT, Stephen. Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems. *IEEE Journal of selected topics in signal processing*, 2007, vol. 1, no 4, p. 586-597.

FOUCART, Simon y RAUHUT, Holger. A mathematical introduction to compressive sensing. *Basel: Birkhäuser*, 2013.

FUENTES, Henry y GONZALEZ, Oscar. Implementation of a 2D Real-Time Correlator Based on a DSP. *Internal Report, Universidad Industrial de Santander*, 2003.

FUKUNAGA, Keinosuke. *Introduction to statistical pattern recognition*. Academic press, 2013.

JAVIDI, Bahram. *Image recognition and classification*. New York: Marcel Dekker, 2002.

SRINIVASAN, Nirmala. *Cross-correlation of biomedical images using two dimensional discrete Hermite Functions*. Tesis Doctoral, University of Akron, 2012.

TSAI, Du-Ming y LIN, Chien-Ta. Fast normalized cross correlation for defect detection. *Pattern Recognition Letters*, 2003, vol. 24, no 15, p. 2625-2631.