

Proyecto Final  
Estudio sobre la Aplicación del Método SIFT en Imágenes SAR

Agopian Michel, Colloca Tomás, Di Nucci Nicolás Santiago

22 de julio de 2017

Departamento de Ingeniería Informática  
ITBA

# Índice

|   |           |
|---|-----------|
| <b>1. Introducción</b>  | <b>3</b>  |
| <b>2. Estado del Arte</b>   | <b>4</b>  |
| 2.1. Motivación . . . . .   | 4         |
| 2.2. Método SIFT . . . . .  | 4         |
| 2.3. Variantes del algoritmo SIFT . . . . .                                       | 5         |
| 2.3.1. Separación de puntos característicos en máximos y mínimos . . . . .        | 5         |
| 2.3.2. Extensión de SIFT agregando un nuevo elemento a los descriptores . . . . . | 6         |
| 2.3.3. Variación de SIFT a partir de un nuevo descriptor . . . . .                | 6         |
| 2.3.4. Algoritmo SAR-SIFT . . . . .   | 6         |
| <b>3. Imágenes SAR</b>  | <b>7</b>  |
| 3.1. Proceso Multilook . . . . .  | 8         |
| 3.2. Despliegue de imágenes SAR . . . . .   | 8         |
| 3.2.1. Distribución $\mathcal{G}_I^0$ . . . . .                                   | 9         |
| 3.2.2. Generación de imágenes SAR sintéticas . . . . .                            | 10        |
| 3.2.3. Ecualización de imágenes SAR . . . . .                                     | 10        |
| <b>4. Métodos utilizados en este trabajo</b>                                      | <b>13</b> |
| 4.1. SIFT . . . . .   | 13        |
| 4.2. Método SAR-SIFT . . . . .  | 17        |
| <b>5. Experimentos realizados</b>   | <b>21</b> |
| 5.1. Mapa de estimaciones del parámetro $\alpha$ . . . . .                        | 21        |
| 5.2. Comparación entre SIFT y SAR-SIFT . . . . .                                  | 24        |
| 5.2.1. Imágenes sintéticas . . . . .  | 24        |
| 5.2.2. Imágenes reales . . . . .  | 25        |
| <b>6. Resultados y conclusiones</b>   | <b>28</b> |
| 6.1. Mapa de estimaciones del parámetro $\alpha$ . . . . .                        | 28        |
| 6.2. Comparación entre SIFT y SAR-SIFT . . . . .                                  | 29        |
| 6.2.1. Imágenes sintéticas . . . . .  | 29        |
| 6.2.2. Imágenes reales . . . . .  | 33        |
| <b>7. Conclusiones Finales</b>  | <b>40</b> |

# Proyecto Final

## Estudio sobre la Aplicación del Método SIFT en Imágenes SAR

Agopian Michel, Colloca Tomás, Di Nucci Nicolás Santiago

Tutor: Dra. María Juliana Gambini

Instituto Tecnológico Buenos Aires (ITBA)

### Resumen

El método SIFT (Scale Invariant Feature Transform) se utiliza para reconocer un objeto dentro de una imagen o decidir si dos imágenes son iguales. Posee la ventaja de que puede realizar el reconocimiento aún si alguna de las dos imágenes involucradas está bajo el efecto de transformaciones afines, como traslación, rotación o cambios de escala; mientras que también es resistente a cambios de iluminación y borrono. Por otro lado, las imágenes SAR (Synthetic Aperture Radar), que son ampliamente utilizadas en aplicaciones de cuidado ambiental, son de naturaleza muy diferente a las imágenes ópticas como son las imágenes en formato RGB (Red, Green, Blue), satelitales o infrarrojas. Las imágenes SAR poseen ruido *speckle*, el cual es muy difícil de eliminar por ser inherente al método de captura. Por esta razón, los métodos que fueron desarrollados para imágenes ópticas, fallan cuando son aplicados en imágenes SAR.

En este trabajo se estudia la aplicación del método SIFT y su variante: el algoritmo SAR-SIFT desarrollado para aplicarse en imágenes SAR. Se introduce primero el motivo de interés del uso de este tipo de imágenes, mencionando que el ruido *speckle* que presentan dificulta la aplicación de los métodos de procesamiento de imágenes como el método SIFT. Se explica además cómo se puede desplegar una imagen SAR a partir de un archivo y cómo es posible generar imágenes sintéticas utilizando la distribución  $\mathcal{G}_I^0$ . Esta distribución permite modelar los datos provenientes de imágenes SAR. Está indexada por tres parámetros:  $\alpha$  relacionado con la textura de la zona,  $\gamma$  correspondiente a la luminosidad y  $L$  que describe cantidad de *looks*.

Debido a la distribución de los valores de los píxeles de las imágenes SAR se explica la importancia de ecualizar las mismas para poder visualizarlas. Luego se explican en detalle los algoritmos SIFT y SAR-SIFT, y sus principales diferencias. Finalmente se propone una alternativa a la aplicación del método SIFT a partir de la generación de un mapa de valores estimados del parámetro  $\alpha$ . Se comparan los métodos SIFT y SAR-SIFT tras ser aplicados tanto a imágenes sintéticas generadas con distintos valores de sus parámetros, como a reales, variando la resolución y la rotación de las mismas.

### Abstract

The SIFT (Scale Invariant Feature Transform) is used to recognize an object within an image or to decide whether two images are the same. It presents the advantage of being able to recognize the images even if one of them is affected by transformations such as, translation, rotation, or scale changes; while it is also resistant to light changes and blurring.

On the other side, SAR (Synthetic Aperture Radar) images, that are broadly used on environmental care application, are from a very different nature than those from optic images

as the ones in RGB (Red, Green, Blue) format, satellite or infrared. SAR images possess *speckle* noise, which is very difficult to eliminate for being inherent to the capture method. For this reason, the methods that were developed for optic images, fail when applied on SAR images.

In this document, the application of the SIFT method and its variant SAR-SIFT, developed for the application on SAR images, are studied. First, it is introduced the main interest in the use of this type of images while mentioning the *speckle* noise they present hinders the application of image processing methods like the SIFT method. Also, it is explained how to deploy SAR images from a file and how it is possible to generate synthetic images with the use of the  $\mathcal{G}_I^0$  distribution. This distribution allows to model data coming from SAR images. It is determined by three parameters:  $\alpha$  that describes the texture of the zone,  $\gamma$  corresponding to the luminosity and  $L$  which is the amount of *looks*.

The importance of equalizing SAR images for their visualization is explained due to the pixel values distribution of these. Then, the algorithms SIFT and SAR-SIFT are explained on detail along with their main differences. Finally, an alternative for the application of SIFT based on the generation of a map of the estimated value of the texture parameter  $\alpha$  is proposed. The application of the methods SIFT and SAR-SIFT are compared by using synthetic images generated with diverse parameters, and by using real images with rotation transformations and lowering their resolution.

## 1. Introducción

Las imágenes SAR (Synthetic Aperture Radar) son de suma importancia porque pueden ser utilizadas para la evaluación de catástrofes, terremotos, tsunamis, detección de talas ilegales, detección temprana de incendios y monitoreos ambientales, entre otras aplicaciones. Para crear una imagen SAR, se emiten ondas electromagnéticas, y se reciben y registran los ecos de cada una de ellas. La onda electromagnética atraviesa las nubes, y por esa razón este tipo de imágenes puede tomarse independientemente del clima. Además, el radar tiene iluminación propia, y por lo tanto, no necesita luz solar para tomar las imágenes. Estas son dos ventajas importantes que poseen las imágenes SAR con respecto a otros sistemas de captura.

Sin embargo, estas imágenes tienen la dificultad de que presentan un tipo de ruido llamado ruido *speckle*, que provoca la apariencia granulada como muestra la Figura 1 .

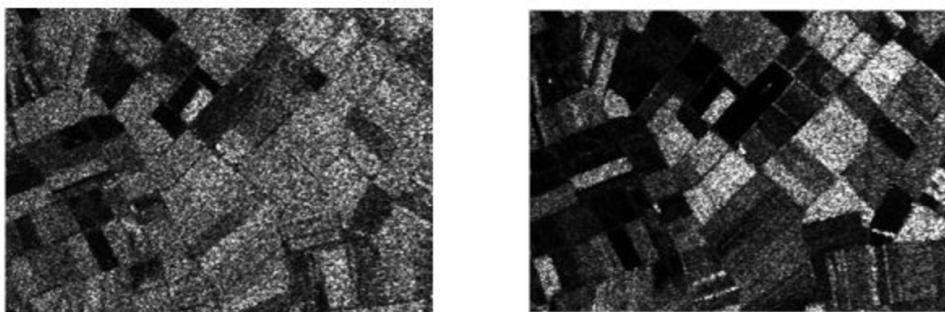


Figura 1: Ejemplo de imágenes SAR de una zona de cultivos.

Este ruido es inherente al método de captura, y es la razón por la cual dificulta el funcionamiento de los métodos clásicos de procesamiento de imágenes.

Por otro lado, el algoritmo SIFT (Scale-Invariant Feature Transform)[7], desarrollado en el



año 2004, es un método que sirve para identificar objetos dentro de una imagen, y es resistente a transformaciones de rotación, traslación, escalamiento, iluminación y borrono. Este método posee múltiples aplicaciones en reconocimiento de objetos en imágenes, reconocimiento de personas, reconocimiento de gestos, unión de imágenes, reconstrucciones 3D, y más. El algoritmo se divide en varias etapas: detección de extremos en el espacio escala, localización de puntos de interés, asignación de orientaciones, elaboración de descriptores, y correspondencia de puntos de interés.

Este método fue desarrollado para reconocimiento de objetos en imágenes ópticas y falla cuando es aplicado a imágenes SAR. En este trabajo estudiaremos una variante llamada SAR-SIFT que busca mejorar el funcionamiento de este algoritmo sobre imágenes SAR.

Este último tiene utilidad sobre las imágenes SAR para clasificar terrenos, detectar cambios, encontrar correspondencias entre zonas del mismo, por ejemplo, luego de un desastre natural.

Este trabajo está compuesto de la siguiente manera: En la Sección 2 se explica el estado del arte. Luego, en la Sección 3, se explica la naturaleza de las imágenes SAR y todo lo que se debe tener en cuenta para su despliegue y utilización. En la Sección 4 se explican los métodos utilizados en el trabajo (SIFT y SAR-SIFT, Secciones 4.1 y 4.2 respectivamente). Finalmente, en la Sección 5 se describen los experimentos realizados para estudiar los métodos, seguidos de los resultados y conclusiones obtenidas en las Secciones 6 y 7.

## 2. Estado del Arte

### 2.1. Motivación

Las imágenes SAR (Synthetic Aperture Radar) han tomado mayor importancia en el último tiempo ya que tienen un amplio campo de aplicación y, por la forma en la que se efectúa su captura, esta última no se ve afectada por el estado del clima ni la iluminación. No obstante, el ruido *speckle* inherente a ellas dificulta la aplicación de algoritmos de procesamiento de imágenes. El método SIFT (Scale-Invariant Feature Transform) sirve para hallar puntos característicos en dos imágenes, y luego buscar correspondencias entre los mismos, para así poder determinar si una imagen o parte de esta está incluida en la otra. La ventaja que provee dicho algoritmo frente a otros es que es invariante a cambios de escala, rotación, traslación e iluminación.

El objetivo de esta sección es exhibir distintas variantes de SIFT, entre ellas una de particular interés para imágenes SAR, y analizar sus ventajas y desventajas.

### 2.2. Método SIFT

El método SIFT (Scale Invariant Feature Transform) , desarrollado por Lowe [7], consiste en reconocer un objeto o una imagen dentro de otra imagen.

El método transforma una imagen en un conjunto de descriptores, originalmente arreglos de 128 elementos que describen a un punto característico. Cada uno de estos descriptores es invariante con respecto a traslaciones, rotaciones, cambios de escala o iluminación de la imagen. Además, estos descriptores son robustos a otras transformaciones como cambios de perspectiva, ruido, cambio de contraste, deformaciones y borronos.

El método consiste en tres operaciones sucesivas: Encontrar los puntos característicos, elaborar un descriptor para cada uno de ellos, y finalmente hallar las correspondencias entre descriptores de dos imágenes. Para esto último, se compara cada descriptor de una imagen con todos los de la otra, y se considera que existe una correspondencia si la distancia euclídea entre ambos es menor a un umbral.

Un ejemplo de aplicación del método es la unión de sucesivas imágenes para crear una imagen panorámica, lo que se conoce como *panorama stitching*[2]. Se puede ver un ejemplo de esto en la Figura 2 .

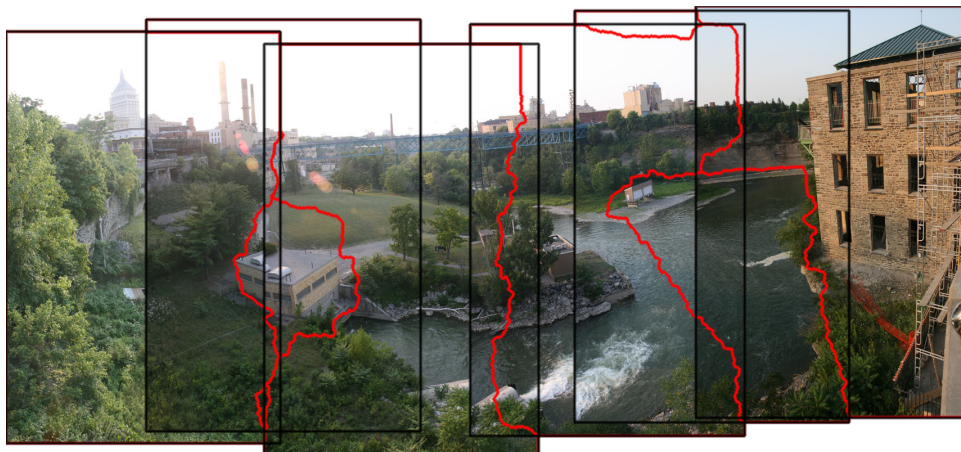


Figura 2: Ejemplo de *panorama stitching* .

## 2.3. Variantes del algoritmo SIFT

Una de las desventajas del algoritmo SIFT es su alto costo computacional. Es por ello, que existen diversas variantes al método para agilizar el mismo a costa de precisión. Algunas de estas se explican en las siguientes subsecciones.

Por otro lado, debido a la presencia de ruido de tipo *speckle* el cual es inherente a las imágenes SAR, que impide el correcto funcionamiento del algoritmo SIFT, se han desarrollado variantes del último que buscan mejorar su precisión. En la últimas dos subsecciones se explican dos variaciones del algoritmo, entre ellas el método SAR-SIFT.

### 2.3.1. Separación de puntos característicos en máximos y mínimos

La forma original de detección de puntos extremos propuesta por Lowe, consiste en tomar los extremos tras aplicar Diferencias de Gaussianas sobre la imagen. Es por ello, que los mismos pueden ser tanto mínimos o máximos.

Como se menciona en [1], la cantidad de puntos característicos mínimos y máximos es prácticamente la misma en una imagen, y además, las correspondencias suelen ser dentro de una misma categoría. Es por eso que en [1] se propone comparar únicamente los puntos característicos del mismo tipo entre sí, lo cual conlleva a una reducción del 50 % del tiempo en hallar correspondencias entre dos imágenes.

### 2.3.2. Extensión de SIFT agregando un nuevo elemento a los descriptores

La comparación entre los distintos descriptores de dos imágenes es una de las causas del alto costo computacional del algoritmo SIFT. En [1] se propone agregar un ángulo al arreglo de elementos que conforman a un descriptor. Este último, que se puede obtener sin un costo computacional adicional, es invariante a la rotación y cambios de iluminación. Es por esto que puede ser utilizado en el algoritmo SIFT sin comprometer la invarianza del mismo frente a distintos cambios. Esta variación propone dividir a los descriptores en subconjuntos disjuntos en función del nuevo ángulo. Finalmente, lo que sugiere en [1] es comparar únicamente aquellos descriptores que se encuentran en el mismo subconjunto, de esta forma es posible acelerar el método SIFT hasta 18 veces sin perder demasiadas correspondencias correctas.

### 2.3.3. Variación de SIFT a partir de un nuevo descriptor

En [10] se introduce el parámetro LGRPH (Local Gradient Ratio Pattern Histogram), que es tomado del vecindario de puntos distinguidos. Este propone obtener los puntos característicos mediante el método de Diferencias de Gaussianas (como varios algoritmos SIFT modificados) y luego calcula el LGRPH para cada uno de ellos. Está probado experimentalmente que este enfoque es robusto frente al ruido *speckle* hallado en las imágenes SAR. Existen varias modificaciones del método SIFT, como SIFT-OCT[9], SIFT+ISEF-OCT[9], SAR-SIFT[5], ISIFT[8]. Todas las mencionadas anteriormente se enfocan en sobreponerse a los disturbios en el cálculo de los puntos característicos debido al ruido *speckle*, pero poco hacen para la mejora de los descriptores.

Además, en el mismo artículo se propone un método en el que se utiliza el parámetro LGRPH y se compara el mismo con SIFT y SIFT-OCT, obteniendo para el nuevo método un porcentaje de aciertos más alto y encontrando igual o mayor cantidad de correspondencias correctas. Sin embargo, este método conlleva mayor costo computacional que los otros.

### 2.3.4. Algoritmo SAR-SIFT

El algoritmo SIFT no funciona correctamente en imágenes SAR debido a su fuerte intensidad y la naturaleza multiplicativa del ruido que presenta. Se propone el método SAR-SIFT[5] como adaptación de SIFT para imágenes SAR.

El mismo utiliza un nuevo método de detección de puntos característicos, un cálculo de la orientación más adecuado para el ruido *speckle*, y un nuevo descriptor que provee un mejor rendimiento que el enfoque usual de SIFT en imágenes SAR. Se validó este nuevo algoritmo con distintos experimentos.

Con esta adaptación del algoritmo SIFT para imágenes SAR, se espera poder usar todas las aplicaciones del algoritmo, como reconocimiento de objetos, *tracking*, *stitching*. Un ejemplo de esto puede observarse en la Figura 3, donde se comparan los resultados de utilizar el algoritmo SAR-SIFT contra los de utilizar el algoritmo SIFT.

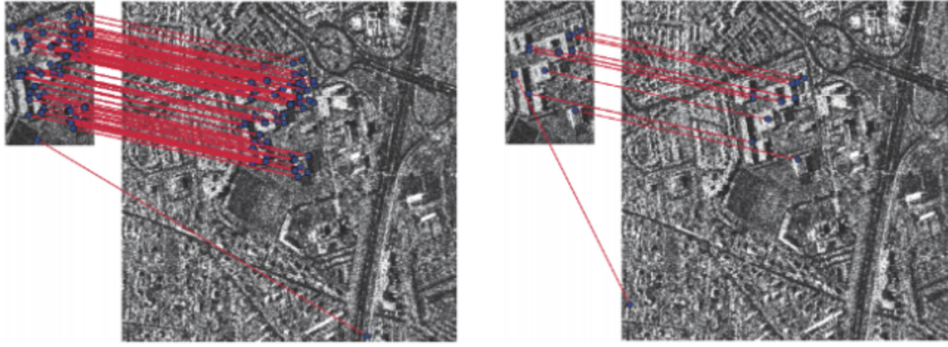


Figura 3: Resultados de aplicar el método SIFT a una imagen SAR (derecha) y el método SAR-SIFT a la misma imagen (izquierda).

### 3. Imágenes SAR

El sistema SAR (Synthetic Aperture Radar) captura imágenes utilizando iluminación coherente por medio de una antena instalada en una aeronave. La antena de radar emite ondas electromagnéticas con diferentes longitudes de onda, las cuales refractan en la región objetivo y la respuesta es capturada por el radar. Dado que un haz más estrecho permite obtener una mejor resolución, y para esto se requiere una antena de mayor tamaño, SAR emula una longitud más grande de la antena utilizando el movimiento de la aeronave, captando y combinando los ecos de las ondas emitidas para las distintas posiciones. De esta forma, comparado con otros métodos de captura, el sistema SAR logra crear imágenes de alta resolución con antenas comparativamente más pequeñas.

Los radares polarimétricos, como lo es SAR, emiten radiaciones electromagnéticas con dos tipos de polarización: horizontal ( $h$ ) y vertical ( $v$ ). A su vez, de la respuesta de ambas polarizaciones emitidas se captan las componentes horizontal y vertical. De esta forma, por ejemplo, la componente del retorno  $Z_{hv}$  representa la componente vertical de la respuesta de emitir la radiación con polarización horizontal. Por lo tanto, el retorno  $Z$  queda definido como:

$$Z = [Z_{hh}, Z_{hv}, Z_{vh}, Z_{vv}] \quad (1)$$

A las componentes del retorno  $Z$  se las llama respectivamente banda  $HH$ , banda  $HV$ , banda  $VH$ , y banda  $VV$ .

Las imágenes que utilizan una combinación de estas bandas para los niveles de RGB son llamadas imágenes polarimétricas, mientras que las que solo utilizan una banda son llamadas monopolarimétricas. En este trabajo, las imágenes SAR reales utilizadas son monopolarimétricas, por lo tanto, las mismas se visualizan a partir de una única banda.

Las imágenes SAR pueden ser utilizadas para la evaluación de daños en catástrofes naturales, terremotos, tsunamis, detección de talas ilegales, detección temprana de incendios y monitoreos ambientales, entre otras aplicaciones. La ventaja más grande que presentan es que su captura no se ve afectada por el estado del clima ni la iluminación. Sin embargo, las mismas presentan un tipo de ruido, denominado ruido *speckle*, el cual es muy complejo de eliminar y dificulta la aplicación de los algoritmos clásicos para procesamiento de imágenes.

En la Figura 4 se puede observar el resultado de aplicar el método de Canny de detección de bordes a una imagen SAR. Puede observarse que se obtienen muchos puntos de borde falsos. En la Figura 5 podemos ver el mismo método aplicado a una imagen óptica.

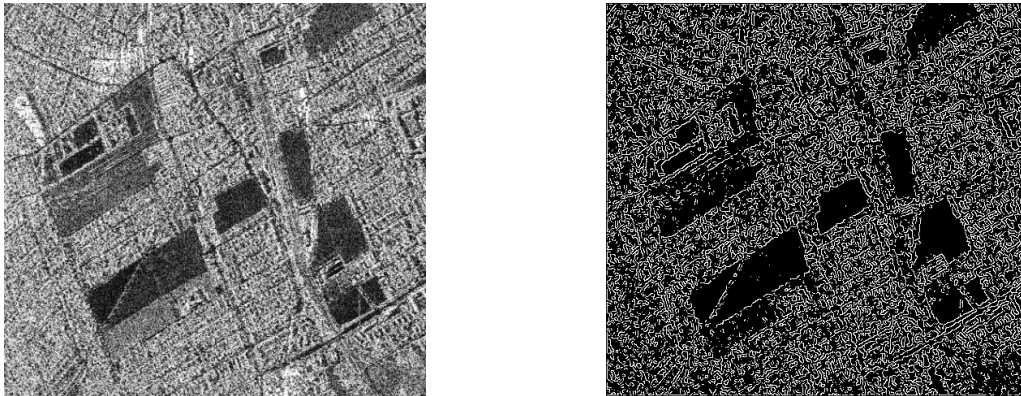


Figura 4: Resultados de aplicar el método de Canny de detección de bordes a una imagen SAR utilizando un filtro Gaussiano de tamaño  $7 \times 7$  y  $\sigma = 1$ .



Figura 5: Resultados de aplicar el método de Canny de detección de bordes a una imagen óptica utilizando un filtro Gaussiano de tamaño  $7 \times 7$  y  $\sigma = 1$ .

### 3.1. Proceso Multilook

Para reducir el ruido *speckle*, una de las técnicas utilizadas es generar varias "vistas", o *looks*, a partir del mismo conjunto de datos durante el proceso de generación de la imagen. Esta técnica se denomina Proceso Multilook. Cada *look* es una observación independiente de la misma escena. El promediado píxel a píxel de cada uno de ellos genera una imagen *multilook* que posee menor ruido *speckle*.

### 3.2. Despliegue de imágenes SAR

Las imágenes SAR utilizadas para los experimentos del presente trabajo se encuentran en formato ENVI, acompañadas de un archivo denominado ENVI header que cuenta con información esencial para la lectura de la misma. Dentro de este último, se tienen en cuenta las siguientes especificaciones para proceder con la lectura:

- *samples*: Ancho de la imagen SAR.
- *lines*: Alto de la imagen SAR.
- *data type*: Formato en el que se expresan los números dentro del archivo que contiene a la imagen (integer, float, double, complex, entre otros).
- *byte order*: Orden en el que se encuentran los números (little endian o big endian).
- *header offset*: Número de bytes del header que se encuentran embebidos dentro del archivo que contiene a la imagen.
- *interleave*: Formato en el que se encuentran intercalados los datos dentro del archivo que contiene a la imagen, pudiendo ser BSQ, BIL o BIP.

No se tienen en cuenta otros datos opcionales como el tipo de sensor utilizado para tomar la imagen o la fecha de adquisición.

A partir de dicha información, se procede a la lectura del archivo que cuenta con los datos de la imagen teniendo en consideración la información que se obtiene del header para poder lograr una correcta visualización de la imagen y también un procesamiento óptimo de la misma.

### 3.2.1. Distribución $\mathcal{G}_I^0$

La distribución  $\mathcal{G}_I^0$  se utiliza para modelar datos provenientes de imágenes SAR porque es capaz de modelar zonas extremadamente texturadas pero también zonas poco texturadas[4].

En el modelo multiplicativo se explica que una variable aleatoria  $Z$  puede formarse a partir de otras dos variables aleatorias  $X$  e  $Y$ , como se muestra a continuación:

$$Z = X \cdot Y, \quad (2)$$

donde  $X$  e  $Y$  corresponden al *backscatter* y al ruido *speckle*, respectivamente.

La variable aleatoria  $Y$  se modela con la distribución  $\Gamma(L, L)$ , donde  $L \geq 1$  es el número equivalente de looks. Si  $X$  sigue una distribución Gaussiana Inversa Generalizada, denotada por  $\mathcal{N}^{-1}(\alpha, \lambda, \gamma)$ ,  $Z$  entonces obedece una distribución  $\mathcal{G}_I$ . Para valores particulares de los parámetros de la distribución  $\mathcal{N}^{-1}$  se obtienen las distribuciones  $\Gamma(\alpha, \lambda)$ ,  $\Gamma^{-1}(\alpha, \gamma)$  e  $IG(\gamma, \lambda)$  (Gaussiana Inversa), que llevan a las distribuciones  $K$ ,  $\mathcal{G}_I^0$  y  $\mathcal{G}^H$  para  $Z$ , respectivamente.

En este trabajo utilizamos la distribución para datos de intensidad  $\mathcal{G}^0$ , denotada por  $\mathcal{G}_I^0$ . Su función de densidad está dada por:

$$f_{\mathcal{G}_I^0}(z) = \frac{L^L \Gamma(L - \alpha)}{\gamma^\alpha \Gamma(-\alpha) \Gamma(L)} \cdot \frac{z^{L-1}}{(\gamma + zL)^{L-\alpha}}, \quad (3)$$

donde  $-\alpha, \gamma, z > 0$  y  $L \geq 1$ .

Los momentos de orden  $r$  son:

$$E(z^r) = \left(\frac{\gamma}{L}\right)^r \frac{\Gamma(-\alpha - r)}{\Gamma(-\alpha)} \cdot \frac{\Gamma(L + r)}{\Gamma(L)} \quad (4)$$

Notar que si la muestra de píxeles tiene media igual a 1, entonces:

$$1 = E(z) = \left(\frac{\gamma}{L}\right) \frac{\Gamma(-\alpha - 1)}{(-\alpha - 1) * \Gamma(-\alpha - 1)} \cdot \frac{L * \Gamma(L)}{\Gamma(L)} = \frac{\gamma}{-\alpha - 1} \implies \gamma = -\alpha - 1 \quad (5)$$

Una de las características más importantes de la distribución  $\mathcal{G}^0$ , tanto para datos de intensidad o amplitud es que puede interpretarse el parámetro  $\alpha$  en términos de la rugosidad de la zona observada. Para valores de  $\alpha$  cercanos a cero (típicamente en el intervalo  $(-3, 0)$ ), la zona de la imagen corresponde a una región muy heterogénea, como es el caso de las zonas urbanas en las imágenes SAR. A medida que el valor del parámetro  $\alpha$  disminuye, corresponde a zonas cada vez más homogéneas, como son las regiones de forestación (usualmente  $(-6, -3]$ ) y pastura (en  $(-\infty, -6)$ ). Esta es la razón por la cual a este parámetro se lo llama parámetro de rugosidad y también la razón por la cual su estimación precisa es de suma importancia en el análisis de imágenes con ruido *speckle*.

Por otra parte, el parámetro  $\gamma$  está relacionado con la luminosidad de la imagen.

### 3.2.2. Generación de imágenes SAR sintéticas

Las imágenes sintéticas son imágenes generadas por el hombre a través de algoritmos con el fin de poder estudiar como se comportan los métodos de interpretación automática de imágenes. Poder generar imágenes sintéticas permite la creación de las mismas variando parámetros, hecho que no puede ocurrir cuando se tiene una imagen real. Esto permite realizar un estudio más profundo de los distintos métodos en cuestión.

Para generar imágenes sintéticas con ruido *speckle*, se toma una imagen en blanco y negro, y se reemplazan los píxeles negros por un valor  $z_1$  y los blancos por un valor  $z_2$ , donde  $z_1$  y  $z_2$  son variables aleatorias generadas con distribución  $\mathcal{G}_I^0(\alpha_1, \gamma_1, L)$  y  $\mathcal{G}_I^0(\alpha_2, \gamma_2, L)$ , respectivamente.

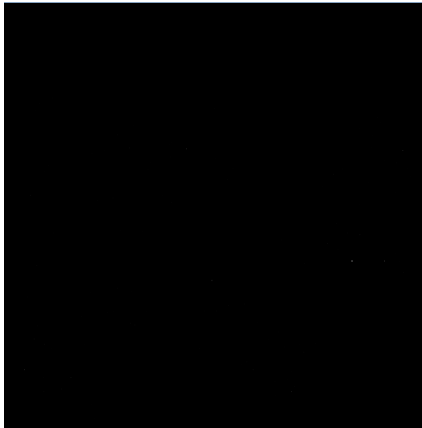
Esto se puede realizar teniendo en cuenta que la distribución  $\mathcal{G}_I^0$  se puede modelar a partir del cociente entre dos variables aleatorias con distribuciones  $\Gamma$ . Sean  $X \sim \Gamma(-\alpha, \frac{1}{\gamma})$  e  $Y \sim \Gamma(L, \frac{1}{L})$ , entonces  $Z = \frac{Y}{X}$  tiene distribución de probabilidad  $\mathcal{G}_I^0(\alpha, \gamma, L)$ .

### 3.2.3. Ecualización de imágenes SAR

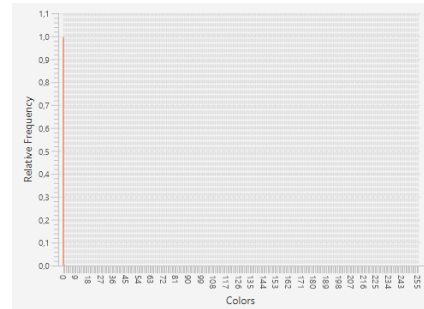
La ecualización de una imagen es una transformación que busca obtener una distribución uniforme para el histograma de la misma. En otras palabras, se busca que haya la misma cantidad de píxeles para los distintos niveles de gris.

Las imágenes SAR suelen tener una amplitud muy grande en los valores de los píxeles. Además, la mayoría de los valores son bajos y muy cercanos entre sí, mientras que otros pocos son varios órdenes de magnitud mayores. Por esta razón, es necesario ecualizar las imágenes para poder visualizarlas.

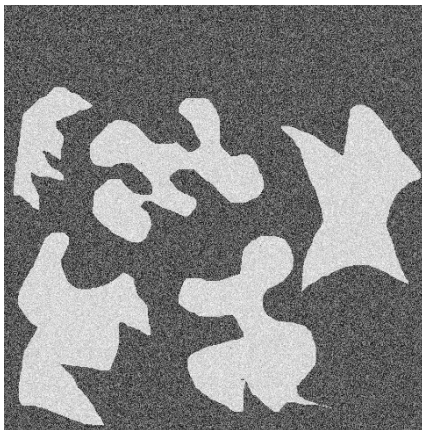
En las Figuras 6 y 7 se puede ver el resultado de aplicar la ecualización a imágenes SAR sintéticas y reales respectivamente, y cómo se visualizan sin la misma.



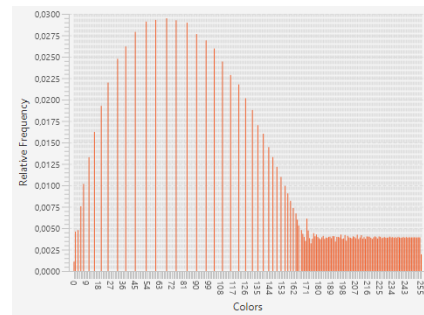
(a) Imagen SAR sintética antes de ecualizar.



(b) Histograma de imagen SAR sintética antes de ecualizar.



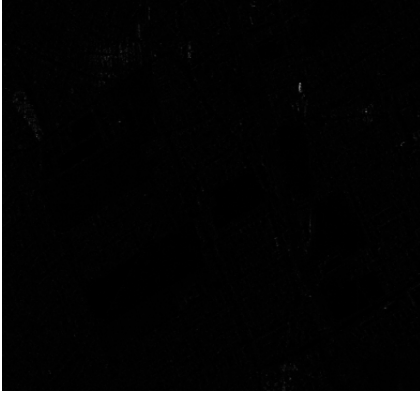
(c) Imagen SAR sintética luego de ecualizar.



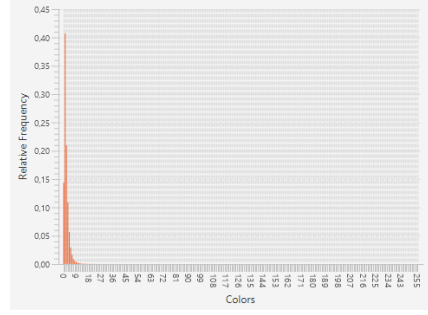
(d) Histograma de imagen SAR sintética luego de ecualizar.

Figura 6: Comparación entre imágenes SAR sintéticas con y sin ecualización, y sus respectivos histogramas.





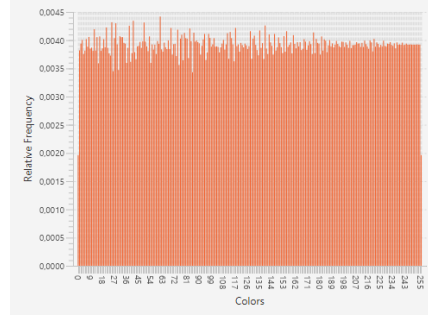
(a) Imagen SAR real antes de ecualizar.



(b) Histograma de imagen SAR real antes de ecualizar.



(c) Imagen SAR real luego de ecualizar.



(d) Histograma de imagen SAR real luego de ecualizar.

Figura 7: Comparación entre imágenes SAR reales con y sin ecualización, y sus respectivos histogramas.

Para realizar la ecualización sobre una imagen  $\mathcal{I}$  de ancho  $\mathcal{W}$  y alto  $\mathcal{H}$ , y donde  $\mathcal{P}_{x,y}$  es el valor del píxel en la posición  $(x, y)$ , se define:

$$\begin{aligned} \min &= \min\{\mathcal{P}_{x,y}, \forall_{x,y}\} \\ \max &= \max\{\mathcal{P}_{x,y}, \forall_{x,y}\} \\ \text{casillero}(\mathcal{P}_{x,y}) &= \frac{\mathcal{P}_{x,y} - \min}{\mathcal{S}} \end{aligned} \quad (6)$$

donde:

$$\begin{aligned} \mathcal{N}(\text{cantidad de casilleros}) &= \mathcal{W} * \mathcal{H} \\ \mathcal{S}(\text{amplitud del casillero}) &= \frac{\max - \min}{\mathcal{N} - 1} \end{aligned} \quad (7)$$

En primer lugar, se inicializa un histograma con  $\mathcal{N}$  casilleros con un valor inicial 0, donde el valor de cada casillero representa el valor que tiene la barra correspondiente del gráfico del histograma. Luego, para cada píxel  $\mathcal{P}_{x,y}$  se incrementa en uno el valor en el casillero dado por  $\text{casillero}(\mathcal{P}_{x,y})$ .

Luego se calculan las frecuencias relativas  $f$  para cada nivel de gris (normalizando el histograma por  $\mathcal{W} * \mathcal{H}$ ), y las frecuencias relativas acumuladas  $F$ , donde:

$$\begin{aligned} F_0 &= f_0 \\ F_i &= F_{i-1} + f_i \\ F_{N-1} &= 1 \end{aligned} \tag{8}$$

Finalmente, teniendo en cuenta que se redefine la frecuencia relativa acumulada para el píxel  $(x, y)$ ,  $F_{casillero(\mathcal{P}_{x,y})}$ , como  $F_{x,y}$ , se reemplaza el valor de cada píxel por un nuevo valor  $P'_{x,y}$ , definido por:

$$\mathcal{P}'_{x,y} = \frac{(max - min) * (F_{x,y} - F_0)}{1 - F_0} + min, \forall_{x,y} \tag{9}$$

## 4. Métodos utilizados en este trabajo

En este trabajo realizamos una comparación entre el método SIFT y el método SAR-SIFT aplicado a imágenes con ruido *speckle*. En esta sección describimos detalladamente estos dos métodos. Para más información ver [7] [5].

### 4.1. SIFT

El método SIFT (Scale-Invariant Feature Transform) es un algoritmo que sirve para hallar puntos característicos en imágenes y luego buscar correspondencias entre los mismos, para así poder determinar si una imagen, o parte de esta, está incluida en la otra. Para cada punto característico se elabora un descriptor, que es un arreglo de valores. Dichos descriptores son invariantes a cambios de escala, rotación, iluminación y traslación, permitiendo así hallar correspondencias entre dos imágenes distintas de una misma escena. En la Figura 8 se puede ver que dicho algoritmo es invariante a cambios de escala, rotación, traslación e iluminación, siendo esta una ventaja que tiene sobre otros algoritmos de reconocimiento de objetos dentro de una imagen.



Figura 8: Ejemplo de correspondencia de puntos característicos correspondientes al mismo cuadro a pesar del cambio de escala, rotación, traslación e iluminación.

Este método se divide en varias etapas: detección de extremos en el espacio escala, localización de puntos de interés, asignación de orientaciones, elaboración de un descriptor para cada punto de interés, y correspondencia entre puntos de interés. En esta sección explicamos cada uno de estos pasos:

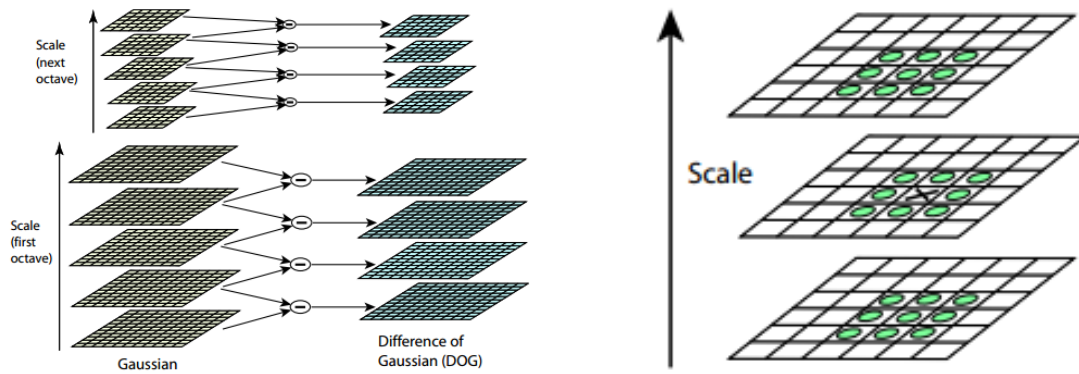
- Detección de extremos en el espacio escala

El primer paso del método consiste en construir el espacio escala gaussiano. Para conseguir esto, se aplican sucesivos filtros gaussianos con valores incrementales de  $\sigma$ , obteniéndose un conjunto de imágenes que se denomina octava. Por otro lado, se obtienen varias octavas al repetir este proceso para imágenes resultantes de reducir la escala de la imagen original. En la Figura 9 se puede ver el espacio escala generado para una imagen.



Figura 9: Ejemplo del espacio escala generado a partir de una imagen para buscar extremos, donde cada fila es una octava para una escala distinta.

Para cada escala, se restan dos imágenes adyacentes, obteniendo lo que se conoce como Diferencia de Gaussianas (DoG). Luego para cada imagen DoG se buscan extremos locales comparando con los 8 vecinos de la misma imagen, y con los 9 vecinos de la imagen DoG anterior, y los otros 9 de la imagen posterior. En la Figura 10 se puede ver cómo se obtienen las Diferencias de Gaussianas y cómo se buscan los extremos locales al comparar con los vecinos en las tres dimensiones.



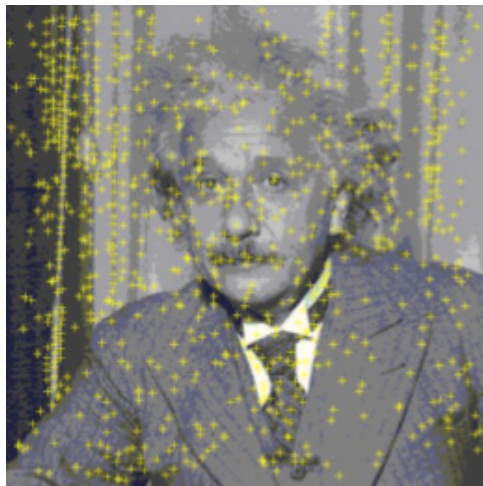
(a) Cálculo de las Diferencias de Gaussianas (DoG) para las primeras dos octavas.

(b) Búsqueda de extremos locales comparando con los 8 vecinos de la misma imagen, los 9 vecinos de la DoG anterior, y los otros 9 de la posterior.

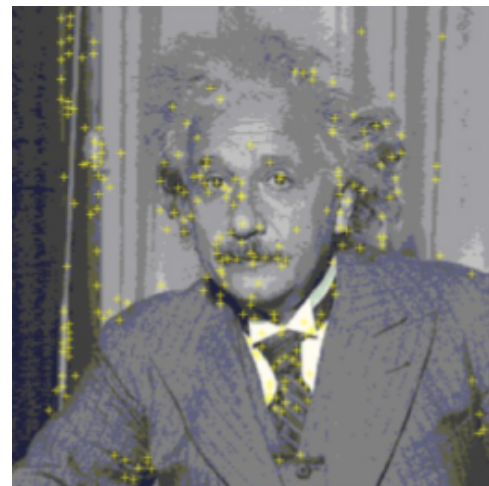
Figura 10: Construcción de las Diferencias de Gaussianas para las distintas octavas y búsqueda de extremos locales comparando con los vecinos en las tres dimensiones.

#### ■ Refinamiento

Sin embargo, el método descrito anteriormente para buscar extremos puede arrojar varios falsos positivos. Por esta razón, se realiza un refinamiento donde se eliminan aquellos con poco contraste o puntos característicos de borde, puesto que estos últimos no son resistentes a cambios de escala. En la Figura 11 se puede ver como se reduce la cantidad de puntos característicos luego de aplicar el refinamiento.



(a) Extremos hallados en el espacio escala gaussiano antes del refinamiento.



(b) Puntos característicos conservados luego del refinamiento.

Figura 11: Proceso de refinamiento de los extremos hallados en una imagen descartando aquellos con bajo contraste y puntos de borde.

- Asignación de orientaciones

Una vez encontrados los puntos de interés, se busca garantizar la invarianza respecto a la rotación de las imágenes. Para esto, para cada punto de interés  $L(x, y)$ , se toman sus vecinos y se calcula la magnitud y la dirección del gradiente a partir de las siguientes ecuaciones:

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2}$$

$$\theta(x, y) = \arctan\left(\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)}\right) \quad (10)$$

Con esto, se realiza un histograma de direcciones, ponderado por la magnitud del gradiente, donde el mayor pico indica la orientación del punto en cuestión. Si existe otro pico significativo (encima del 80 % del valor del más importante) se crea otro punto de interés en la misma posición y escala, pero con distinta orientación.

- Descriptor de punto de interés

Luego, se crean los descriptores de los puntos de interés. Para esto, para cada punto, se toma una máscara de 16x16 puntos alrededor del mismo que, a su vez, se divide en 16 submáscaras de 4x4 puntos. Para cada máscara de 4x4 se crea un histograma de orientaciones aunque de sólo 8 casilleros o barras, que luego se concatenan en un vector que constituye su descriptor. De esta forma, el descriptor es un arreglo de 128 elementos, número que resulta de multiplicar 16, la cantidad de submáscaras, por 8, la cantidad de valores del histograma de orientaciones. En la Figura 12 se puede ver cómo se construye un descriptor.

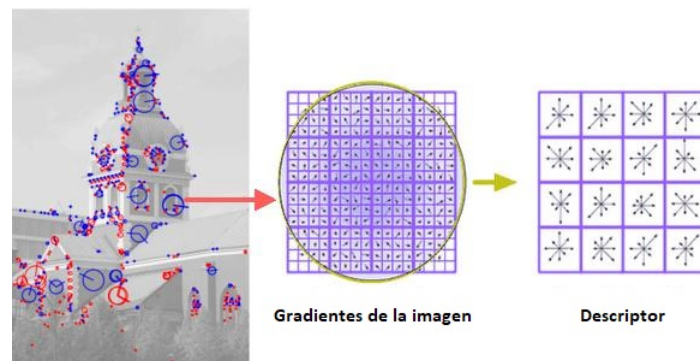
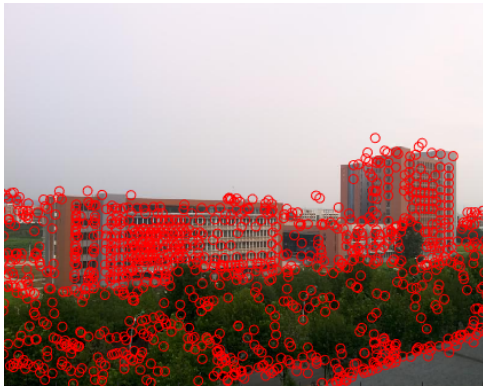


Figura 12: Ejemplo ilustrativo de la asignación de la orientación para un punto característico, y la posterior construcción de su descriptor.

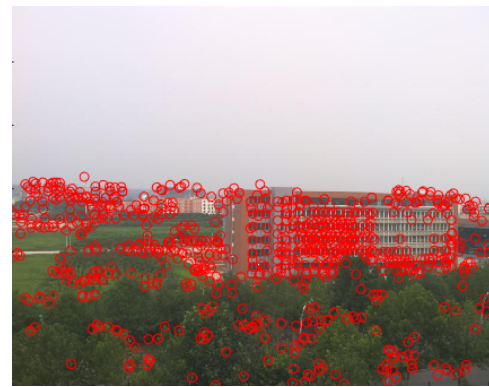
- Correspondencia de puntos de interés

Por último, la correspondencia de puntos característicos entre imágenes se obtiene a través de la búsqueda del punto más próximo en el espacio de descriptores. Puede ocurrir que el punto esté muy cerca del primero a causa del ruido, por lo que en lugar de tomar sólo el punto más próximo, se toma la razón entre la distancia de los primeros dos puntos más cercanos. Si dicho valor es menor a cierto umbral, entonces referencia a un mismo punto

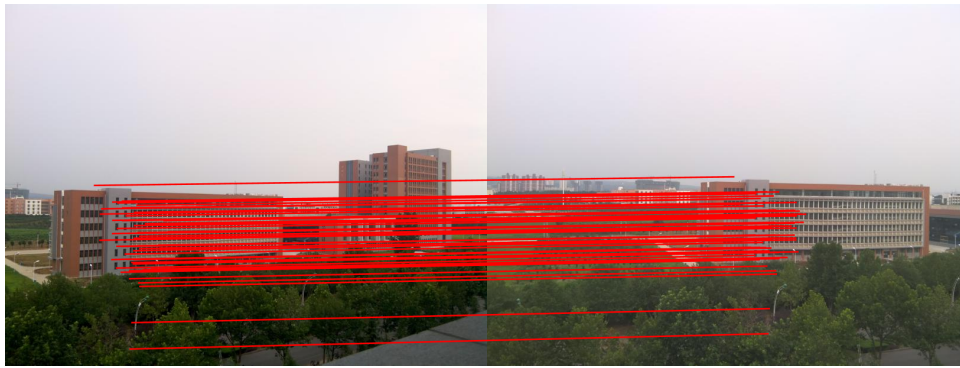
en dos imágenes. En la Figura 13 se puede ver un ejemplo de las correspondencias halladas entre los puntos característicos de dos imágenes que corresponden al mismo edificio.



(a) Puntos característicos hallados (1458) para una imagen de un edificio.



(b) Puntos característicos hallados (1057) para una imagen distinta del mismo edificio que el de la Figura 13(a) .



(c) Correspondencias halladas (71) entre los puntos característicos de las imágenes que se pueden ver en las Figuras 13(a) y 13(b) .

Figura 13: Ejemplo de las correspondencias halladas entre los puntos característicos de dos imágenes que corresponden al mismo edificio.

## 4.2. Método SAR-SIFT

El método SAR-SIFT[5] es un algoritmo basado en SIFT y con el mismo objetivo, diseñado específicamente para imágenes SAR para así poder soportar la fuerte intensidad de estas, y la naturaleza multiplicativa del ruido que presentan. El mismo propone un nuevo método de detección de puntos característicos, un cálculo de la orientación más adecuado para el ruido *speckle*, y un nuevo descriptor que genera una mejor performance que el enfoque usual de SIFT en imágenes SAR, como se muestra en la Figura 3 .

### ■ Detección de puntos característicos

Para la detección de puntos característicos se decide en primer lugar reemplazar las Diferencias de Gaussianas, puesto que utilizar las mismas para buscar extremos tiene el problema en imágenes SAR que en zonas más texturizadas la diferencia del gradiente es mayor que en zonas menos rugosas, y por lo tanto se obtienen más falsos positivos. Por

esta razón, se decide utilizar un método llamado ROEWA (Ratio of Exponentially Weighted Averages), un método que se suele utilizar para calcular bordes en imágenes SAR, al ser más resistente a la naturaleza multiplicativa del ruido *speckle*. Este calcula la media ponderada exponencialmente de las dos regiones delimitadas por la recta que pasa por el píxel en cuestión y una dirección dada, y luego calcula la ración entre ellas. En la Figura 14 se puede ver una ilustración que representa las dos regiones delimitadas a partir de un píxel y la dirección horizontal.

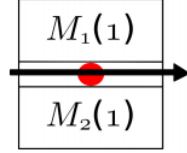


Figura 14: Ejemplo ilustrativo de la división en dos regiones de una ventana sobre un píxel a partir de la dirección horizontal.

Para el cálculo de las medias, se aplica una máscara elaborada a partir de la dirección y la región sobre la cual se la quiere calcular. Por ejemplo, las medias para la dirección horizontal y escala  $\theta$  están dadas por:

$$\begin{aligned} M_{1,\theta}(x, y) &= \exp\left(-\frac{|x| + |y|}{\theta}\right) \quad \forall x, \forall y > 0 \\ M_{2,\theta}(x, y) &= \exp\left(-\frac{|x| + |y|}{\theta}\right) \quad \forall x, \forall y < 0 \end{aligned} \quad (11)$$

Y luego, el cociente entre ambas está dado por:

$$R_{H,\theta}(x, y) = \frac{M_{1,\theta}}{M_{2,\theta}} \quad (12)$$

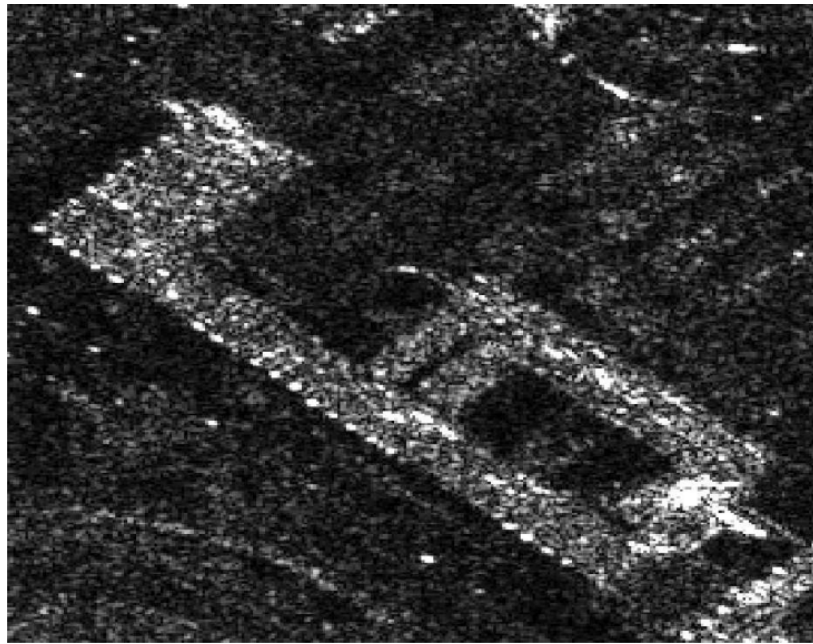
A partir de los cocientes para las direcciones horizontal y vertical, se define el gradiente horizontal y vertical respectivamente como:

$$\begin{aligned} G_{H,\theta}(x, y) &= \log(R_{H,\theta}(x, y)) \\ G_{V,\theta}(x, y) &= \log(R_{V,\theta}(x, y)) \end{aligned} \quad (13)$$

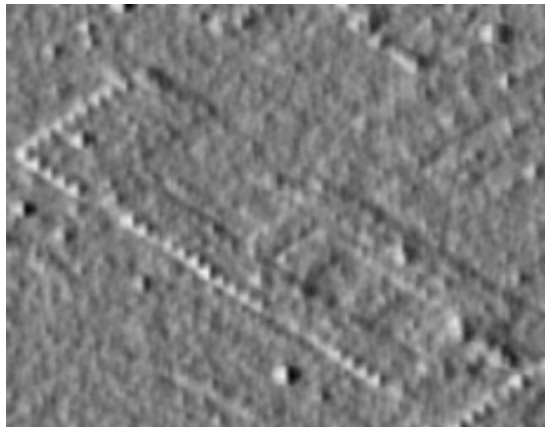
y luego se calcula la magnitud del gradiente y la orientación de manera análoga a los métodos clásicos:

$$\begin{aligned} G_{\theta}^M(x, y) &= \sqrt{G_{H,\theta}(x, y)^2 + G_{V,\theta}(x, y)^2} \\ G_{\theta}^A(x, y) &= \arctan\left(\frac{G_{V,\theta}(x, y)}{G_{H,\theta}(x, y)}\right) \end{aligned} \quad (14)$$

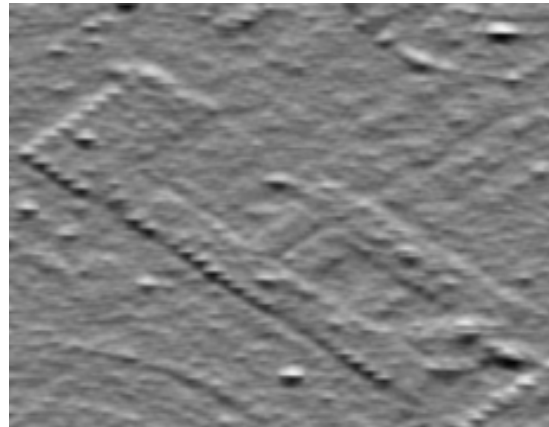




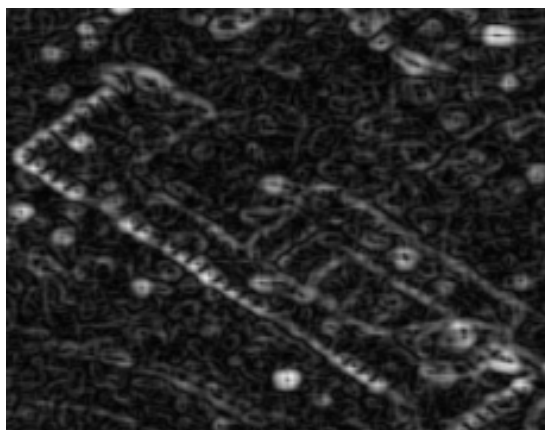
(a) Imagen SAR real.



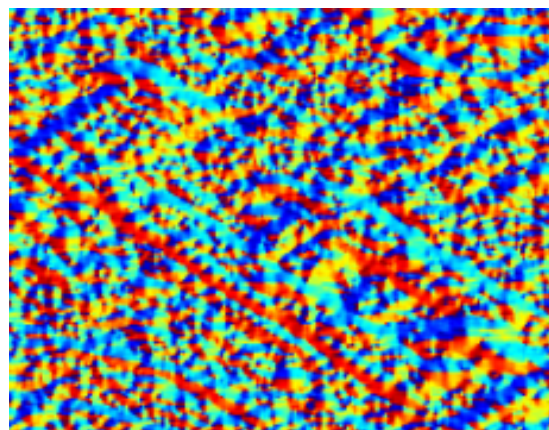
(b) Componente horizontal.



(c) Componente vertical.



(d) Magnitud del gradiente.



(e) Orientación del gradiente, donde cada color representa una orientación en función de una rueda de colores siendo el rojo la dirección noreste, azul noroeste, celeste sudoeste, y amarillo sudeste.

Figura 15: Cálculo del gradiente con  $\theta = 2$ .



A esta nueva forma de calcular el gradiente se la llama GR (Gradient Ratio). En las Figuras 15(b) y 15(c) se pueden ver las componentes horizontal y vertical respectivamente de la imagen SAR real de la Figura 15(a) calculadas utilizando (13). Por otro lado, para la misma imagen SAR se pueden observar en las Figuras 15(d) y 15(e) la magnitud y la orientación del gradiente calculadas a partir de (14).

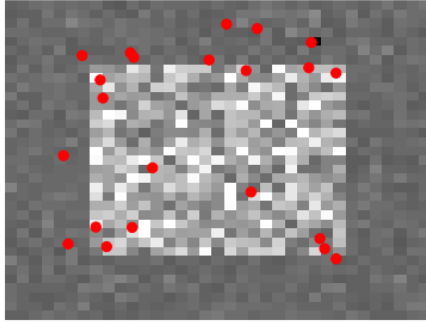
Luego, se utiliza un esquema similar al criterio multiescala de Harris, definiendo la función multiescala SAR-Harris:

$$C_{SH,\theta}(x, y) = \theta^2 * \mathcal{G}_{\sqrt{2}*\theta} * \begin{bmatrix} (G_{H,\theta})^2 & G_{H,\theta} \cdot G_{V,\theta} \\ G_{H,\theta} \cdot G_{V,\theta} & (G_{V,\theta})^2 \end{bmatrix} \quad (15)$$

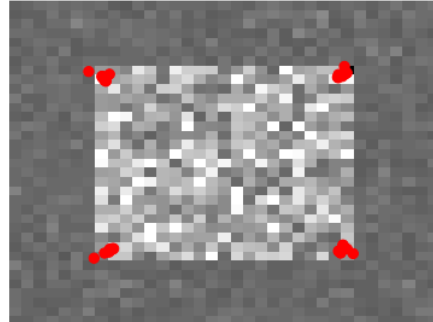
$$R_{SH,\theta}(x, y) = \det(C_{SH,\theta}(x, y)) - d * \text{tr}(C_{SH,\theta}(x, y))$$

Se recomienda utilizar como parámetro  $d$  el mismo valor recomendado para el criterio de Harris [6] que es 0,04.

Finalmente, para el cálculo de puntos característicos se utiliza la función multiescala SAR-Harris (15) con distintas escalas  $\theta_k = \theta_0 * r^{m-1}$  con  $1 < m < m_{max}$ , siendo  $m_{max}$  la cantidad de escalas a computar. A partir de cada escala se buscan los máximos locales comparando con los vecinos y aplicando un umbral  $d_{SH}$  al valor  $R_{SH,\theta}$ . En las Figuras 16(a) y 16(b) se comparan los puntos característicos hallados en una misma imagen sintética utilizando el método de Diferencias de Gaussianas y el método SAR-HARRIS respectivamente. Por otro lado, en las Figuras 17(a) y 17(b) se pueden observar los puntos característicos hallados con los mismos dos métodos pero aplicados en una misma imagen SAR real. Para ambos tipos de imagen con la función multiescala SAR-Harris se obtiene un número similar de puntos característicos que con el método DoG, pero menos distribuidos y con mayor precisión.



(a) Método DoG (22 puntos característicos).



(b) Método SAR-Harris (35 puntos característicos).

Figura 16: Comparación de métodos para hallar puntos característicos en una imagen sintética.

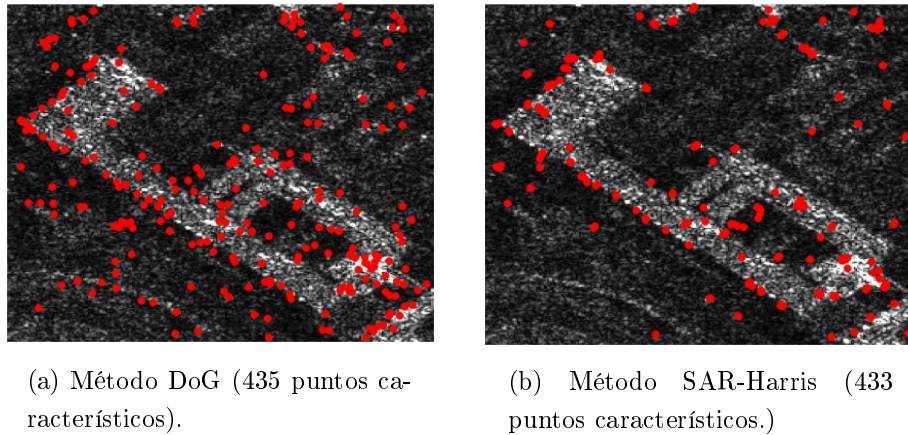


Figura 17: Comparación de métodos para hallar puntos característicos en una imagen SAR real.

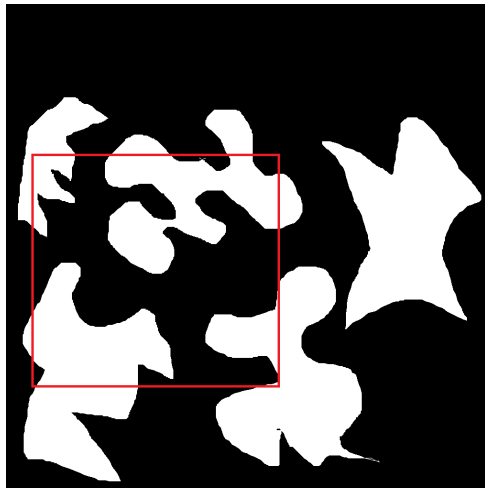
- Cálculo de la orientación de un punto característico y su nuevo descriptor

En el algoritmo SIFT original, tanto la asignación de la orientación a un punto característico como la extracción de su descriptor se basan en histogramas de las orientaciones del gradiente en el vecindario del punto en cuestión. En [5] se propone utilizar la orientación  $G_\theta^A$  calculada con GR (14). Además, se decide reemplazar el vecindario cuadrado de dimensión  $16 \times 16$  y sectores cuadrados de  $4 \times 4$ , por un vecindario circular, cuyo tamaño depende de la escala, y sectores con coordenadas polares logarítmicas.

## 5. Experimentos realizados

### 5.1. Mapa de estimaciones del parámetro $\alpha$

Los siguientes experimentos se realizan modelando los datos provenientes de dos imágenes sintéticas utilizando la distribución  $\mathcal{G}_I^0$  y se estima el parámetro  $\alpha$  para caracterizar la textura de la región. El parámetro  $\alpha$  se estima para toda la imagen utilizando una ventana deslizante, y así se genera lo que se denomina un mapa de alfas. Una vez que se obtiene el mapa de alfas para ambas imágenes, se aplica SIFT para analizar si el método funciona para este nuevo tipo de imagen obtenida, ya que la última no presenta el ruido *speckle* que tiene la imagen SAR original. En la Figura 18 se pueden ver las dos imágenes binarias originales con las cuales se construyen las imágenes SAR sintéticas. En la Figura 19 se pueden ver las mismas imágenes generadas con ruido *speckle*.

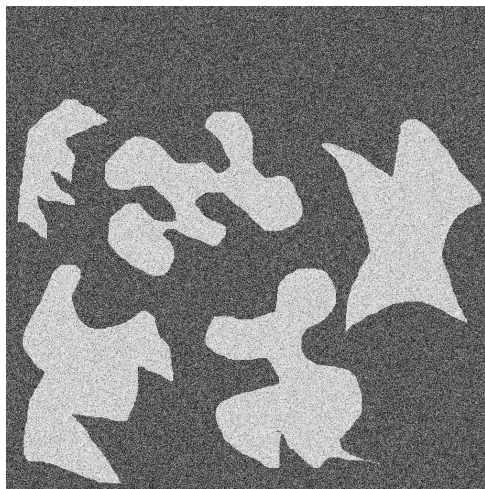


(a) Imagen binaria original. El cuadrado rojo es el fragmento que se utiliza como segunda imagen.

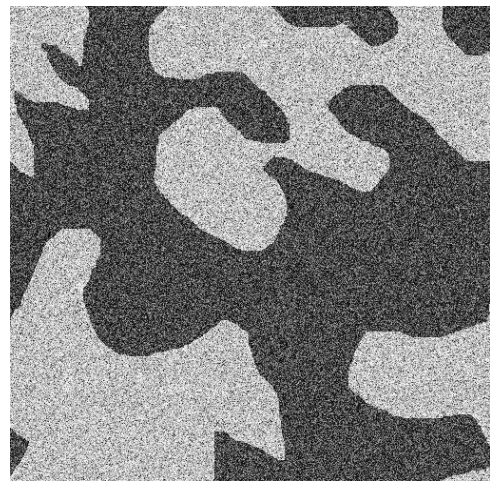


(b) Fragmento de la imagen binaria original.

Figura 18: Imágenes binarias originales con las cuales se generan imágenes SAR sintéticas con ruido *speckle* para luego obtener sus mapas de alfas y poder aplicar SIFT.



(a) Imagen SAR sintética generada a partir de 18(a) .



(b) Imagen SAR sintética generada a partir de 18(b) .

Figura 19: Imágenes SAR sintéticas generadas a partir de las imágenes que se pueden ver en la Figura 18 .

Para obtener el mapa de alfas a partir de una imagen SAR, se toman píxeles avanzando de  $a$  columnas y  $b$  filas, y por cada uno de ellos se toma una ventana de  $N \times N$  píxeles con los cuáles se estima el parámetro  $\alpha$ . En la Figura 20 se puede ver una ilustración de este proceso.

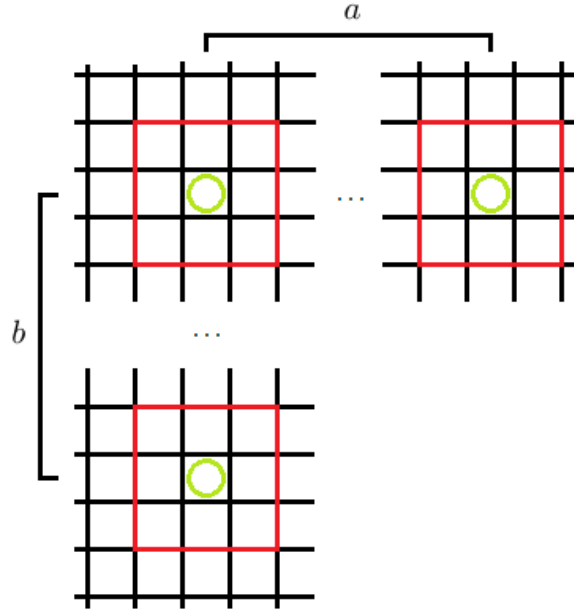


Figura 20: Ilustración del proceso de deslizamiento de la ventana de  $N \times N$  píxeles, representadas en rojo, y de cómo se toman los píxeles, representados con círculos verdes, cada  $a$  columnas y  $b$  filas.

A los píxeles  $p_V(x, y)$  de cada ventana  $V$  se los divide por la media  $\bar{M}_V$ . De esta forma, utilizando ( 5 ), se obtiene  $\gamma^* = -1 - \alpha$ :

$$\bar{M}_V = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N p_v(i, j) \quad (16)$$

$$z_V = \frac{p_V}{\bar{M}_V}$$

Luego se estima  $\alpha_V$  para la ventana  $V$  por el método de máxima verosimilitud utilizando la función de densidad de la distribución  $\mathcal{G}_I^0$  (3) y los valores de los píxeles normalizados  $z_V$ :

$$A = \{-20, -19,5, \dots, -2,5, -2\}$$

$$\alpha_V = \underset{\alpha \in A}{\operatorname{argmax}} \prod_{i=1}^{N \times N} f_{\mathcal{G}_I^0}(z_V(i), \alpha, -1 - \alpha, L) \quad (17)$$

Finalmente, por cada ventana  $V$  se crea una ventana  $V'$  de dimensión  $\frac{a}{2} \times \frac{b}{2}$  centrada en el mismo píxel que  $V$ . A cada píxel de la ventana  $V'$  se lo reemplaza por el valor estimado  $\alpha_V$  de la ventana  $V$  correspondiente. De esta forma, se obtiene una nueva imagen la cual se denomina mapa de alfas. En la Figura 21 se pueden ver los mapas de alfas de las imágenes de la Figura 19 utiizando  $a = b = 10$  y  $N = 51$ .

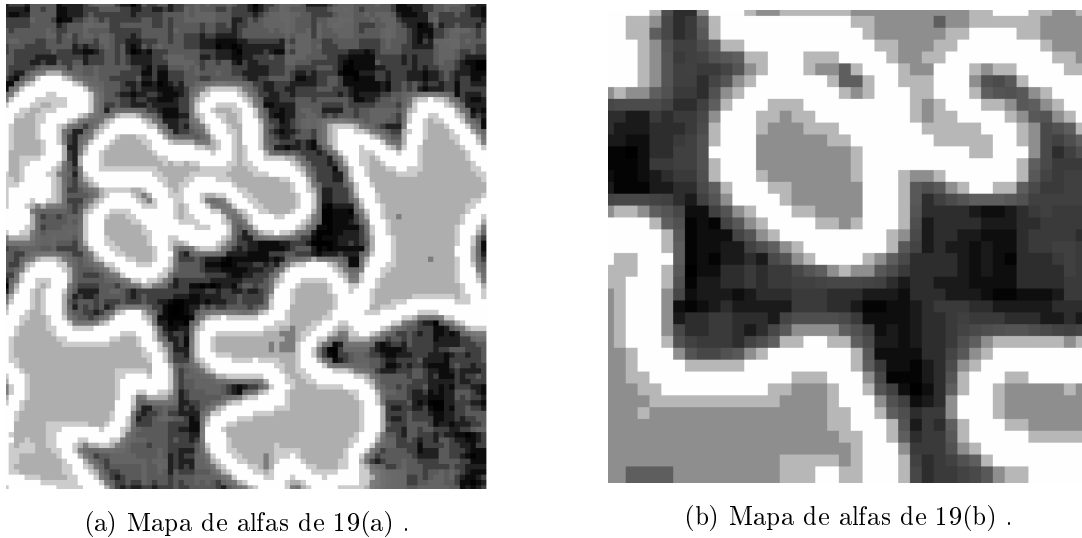


Figura 21: Mapas de alfas que se obtienen a partir de las imágenes que se pueden ver en la Figura 19 .

## 5.2. Comparación entre SIFT y SAR-SIFT

Se realizan distintos experimentos para comparar los métodos SIFT y SAR-SIFT los cuales se efectúan tanto en imágenes sintéticas como reales.

### 5.2.1. Imágenes sintéticas

Para los experimentos con imágenes sintéticas se generan las mismas como se explica en la Sección 3.2.2 . Así se genera una imagen con parámetros  $\alpha_1 = -1,5$ ,  $\alpha_2 = -20$ ,  $\gamma_1 = \gamma_2 = 1$  y  $L = 8$  a la cual se la llama imagen de referencia. Se experimenta variando los parámetros  $\alpha_2$  y  $\gamma_2$ , la resolución de la imagen dada por  $k$ , y la rotación de la imagen dada por  $\theta$ . El cambio de resolución se lleva a cabo tomando ventanas de  $k \times k$  y reemplazando todos los píxeles de la misma por el promedio de sus valores.

Todos los experimentos se realizan para distintos valores de  $L$  con  $L \in \{1, 3, 5, 8\}$ , variando sólo una de las variables mencionadas. Además, se generan 50 imágenes sintéticas con cada combinación de parámetros, y para cada una de ellas, se aplica SIFT y SAR-SIFT utilizando como segunda imagen la imagen de referencia. Finalmente, para cada combinación de parámetros, se toma la media y el desvío estándar de la cantidad de correspondencias obtenidas con las 50 imágenes. En la Figura 22 se puede ver la imagen de referencia.

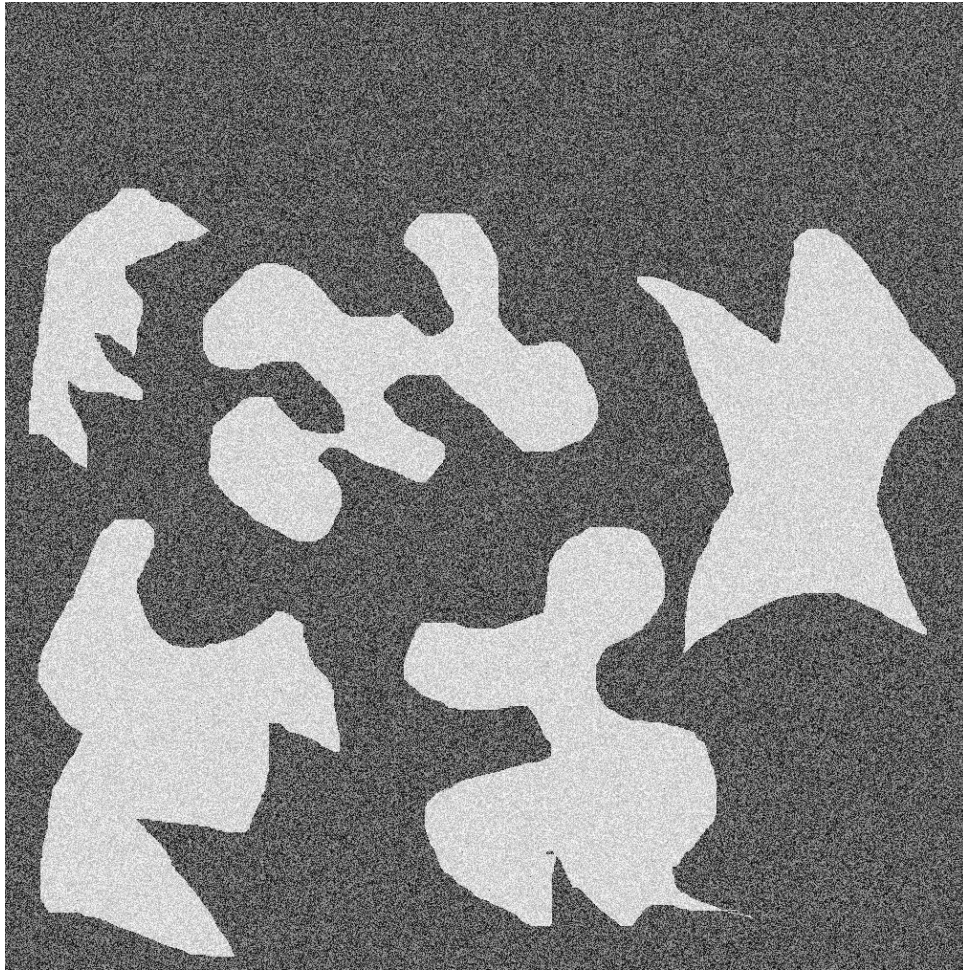


Figura 22: Imagen de referencia con parámetros  $\alpha_1 = -1,5$ ,  $\alpha_2 = -20$ ,  $\gamma_1 = \gamma_2 = 1$  y  $L = 8$  utilizada para los experimentos con imágenes SAR sintéticas.

El primer experimento consiste en variar el parámetro  $\alpha_2$  con  $\alpha_2 \in \{-2, -5, -8, -11, -14, -17, -20\}$ . El mismo se realiza fijando  $\alpha_1 = -1,5$ ,  $\gamma_1 = \gamma_2 = 1$  y variando  $\alpha_2$  y  $L$ .

El segundo experimento consiste en variar el parámetro  $\gamma_2$  con  $\gamma_2 \in \{0,1,1,10,100,1000\}$ . El mismo se realiza fijando  $\alpha_1 = -1,5$ ,  $\alpha_2 = -3$ ,  $\gamma_1 = 1$  y variando  $\gamma_2$  y  $L$ .

El tercer experimento consiste en variar la resolución de las imágenes con escalas  $k$  donde  $k \in \{2,4,8,16\}$ . Este experimento se realiza fijando  $\alpha_1 = -1,5$ ,  $\alpha_2 = -20$ ,  $\gamma_1 = \gamma_2 = 1$  y variando  $k$  y  $L$ .

El cuarto y último experimento que se lleva a cabo con imágenes sintéticas consiste en rotar las imágenes con un ángulo  $\theta$  donde  $\theta \in \{0, 15, 30, 45, 60, 75, 90, 180\}$ . El mismo se realiza fijando  $\alpha_1 = -1,5$ ,  $\alpha_2 = -20$ ,  $\gamma_1 = \gamma_2 = 1$  y variando  $\theta$  y  $L$ .

### 5.2.2. Imágenes reales

Los experimentos que se realizan con imágenes reales se dividen en dos grupos: aquellos que se realizan sobre dos imágenes distintas de la misma región, y aquellos que se realizan utilizando dos bandas distintas correspondientes a la misma imagen SAR. En el caso del primer grupo se toma una de las dos imágenes y se le aplican cambios de rotación y de escala, mientras que la otra se la conserva sin modificar. Luego se aplican los métodos SIFT y SAR-SIFT entre la imagen



modificada y aquella sin modificar. Para el segundo grupo se realiza un experimento similar donde a partir de las cuatro bandas de una imagen SAR se toma una de ellas como imagen de referencia, y a las otras tres se les aplican cambios de rotación y de resolución. Una vez realizado esto se aplican los métodos SIFT y SAR-SIFT entre cada banda modificada y aquella sin modificar. Finalmente se toma el promedio de la cantidad de correspondencias obtenidas con cada una de las tres bandas. En la Figura 23 se muestran los pares de imágenes reales distintas entre sí que corresponden a una misma región. Las imágenes de un mismo par se obtienen a partir de tomas diferentes y por lo tanto presentan cambios de iluminación, resolución, traslación, y levemente de rotación. En la Figura 24 se muestran las tres bandas a las cuales se aplican cambios de resolución y rotación, y la cuarta banda que se la conserva sin modificar, donde las cuatro corresponden a una misma imagen SAR.

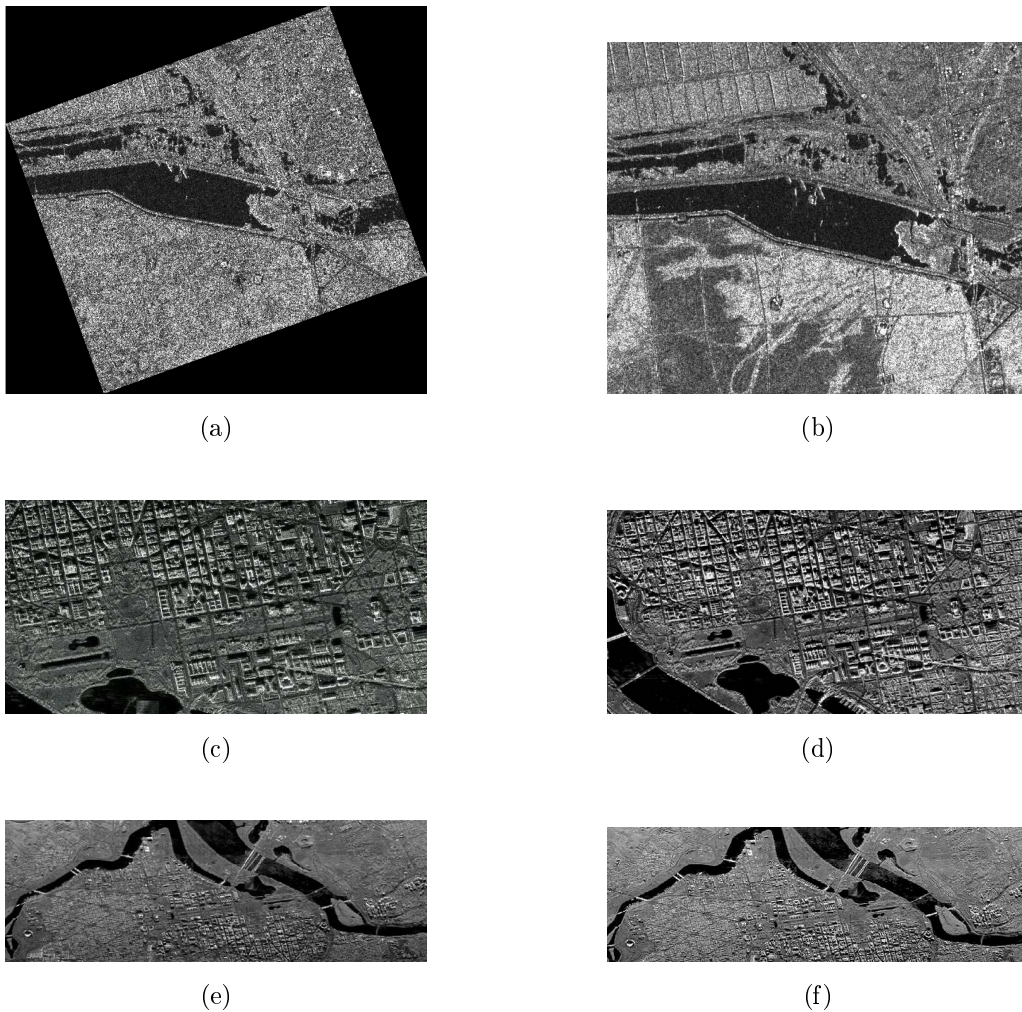
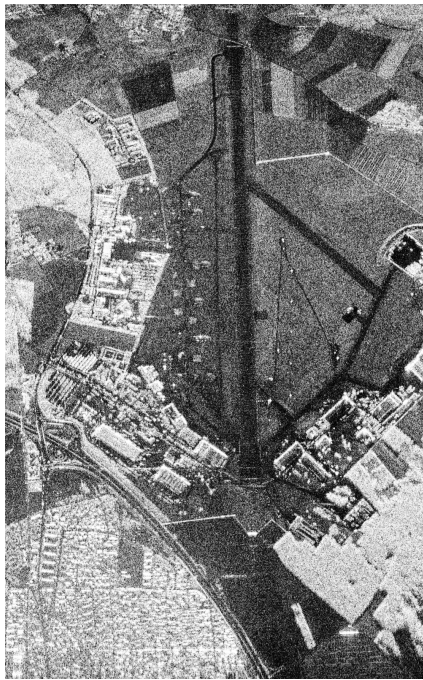
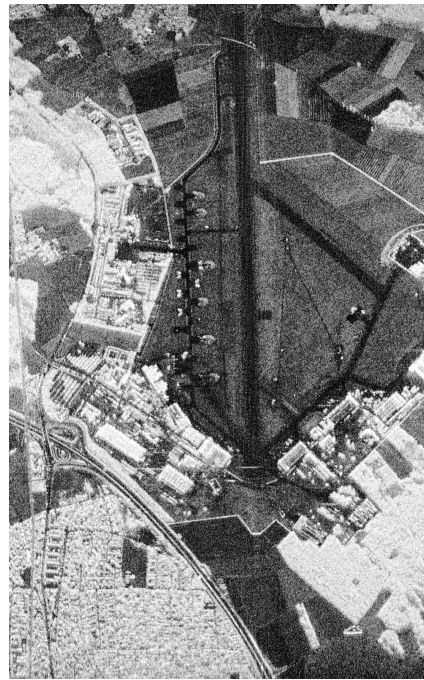


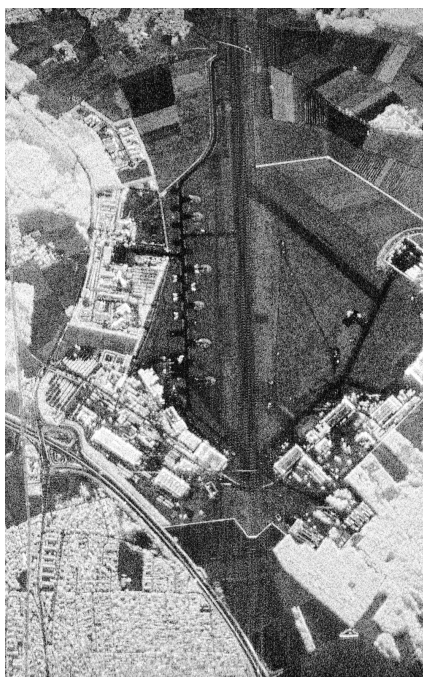
Figura 23: Pares de imágenes SAR reales distintas entre sí ((a) - (b), (c) - (d), (e) - (f)), correspondientes a una misma región.



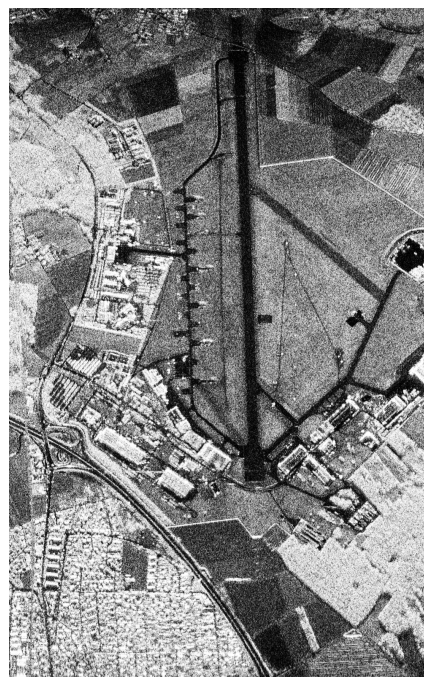
(a) Banda HH.



(b) Banda HV.



(c) Banda VH.



(d) Banda VV.

Figura 24: Cuatro bandas correspondientes a una misma imagen SAR. Se toma (a) como la banda de referencia.

Para ambos grupos el cambio de la resolución de las imágenes se lleva a cabo con escalas  $k$  donde  $k \in \{2, 4, 8, 16\}$ . Para este experimento no se rotan las imágenes.

Por otro lado, la rotación que se realiza para ambos grupos es con un ángulo  $\theta$  donde



$\theta \in \{0, 15, 30, 45, 60, 75, 90, 180\}$ . Dicho experimento se lleva a cabo sin modificar la resolución de las imágenes.

## 6. Resultados y conclusiones

### 6.1. Mapa de estimaciones del parámetro $\alpha$

En la Figura 25 se puede ver el resultado de aplicar el algoritmo SIFT entre dos mapas de alfas que corresponden a imágenes SAR sintéticas de una imagen binaria y un fragmento de la última. La cantidad de correspondencias que encuentra el algoritmo inicialmente es 116, sin embargo, la gran mayoría de esas son falsos positivos y sólo 13 realmente refieren al mismo punto característico hallado en ambas imágenes. Es por esta razón que se considera que la aplicación de SIFT a los mapas de alfas no es una solución al problema de reconocimiento de imágenes en imágenes SAR.

El filtrado de falsos positivos se logra utilizando una variante más performante del algoritmo RANSAC ( Random Sample Consensus), llamada FSC (Fast Sample Consensus), donde el primero consiste en calcular de manera iterativa los parámetros que permiten excluir los *outliers* de un conjunto de datos dado. En este caso particular, se entiende por *outliers* a las correspondencias que resultan falsos positivos.

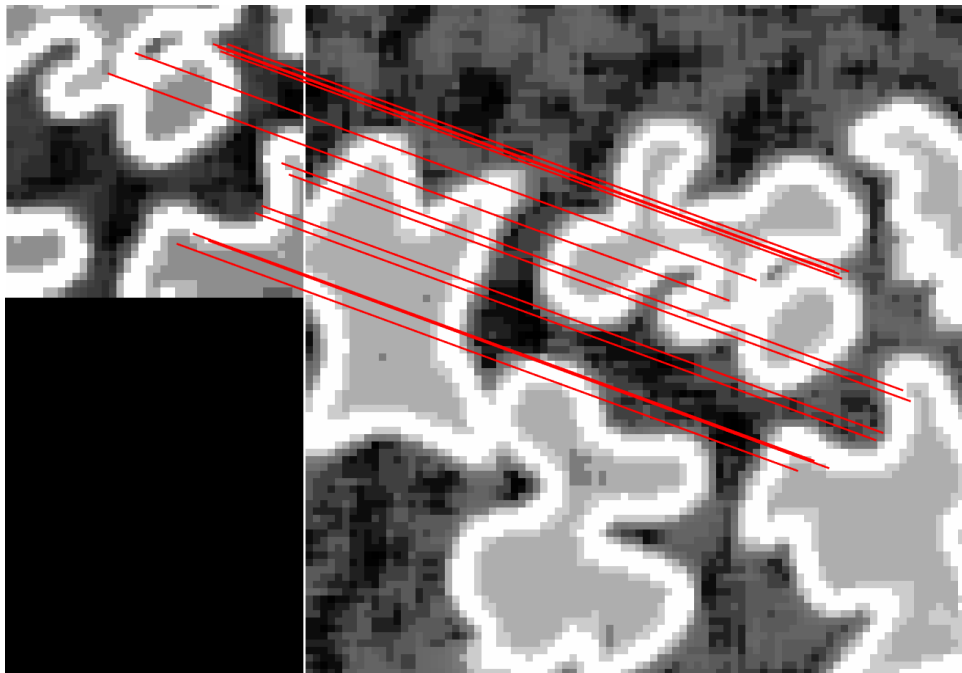


Figura 25: Aplicación del algoritmo SIFT a los mapas de alfas que se obtienen como se describe en la Sección 5.1 y que se pueden ver en la Figura 21 . Solamente se visualizan aquellas 13 correspondencias que se consideran que refieren al mismo punto característico hallado en ambas imágenes luego de aplicar el algoritmo FSC.

## 6.2. Comparación entre SIFT y SAR-SIFT

### 6.2.1. Imágenes sintéticas

En la Figura 26 se pueden ver algunos ejemplos de las imágenes generadas para el primer experimento sobre imágenes SAR sintéticas que consiste en variar los parámetros  $\alpha_2$  y  $L$ .

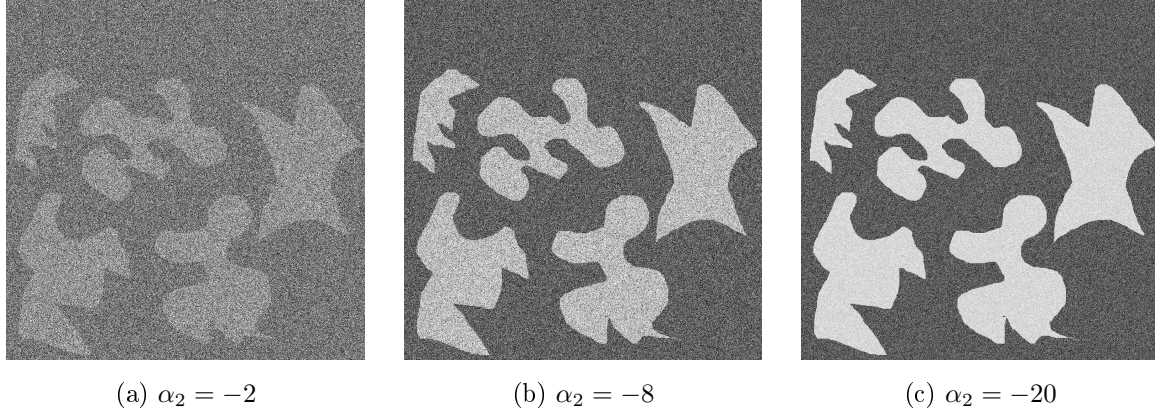


Figura 26: Imágenes SAR sintéticas con parámetros fijos  $\alpha_1 = -1,5$ ,  $\gamma_1 = \gamma_2 = 1$ , y  $L = 8$ , y variando  $\alpha_2$ . La imagen de la Figura 26(c) es la imagen de referencia.

En la Figura 27 se pueden ver los resultados del primer experimento descrito para imágenes sintéticas. Por lo tanto, se muestra la cantidad de correspondencias en función del valor de  $\alpha_2$  para cada valor de  $L$  y donde el intervalo que se muestra en cada punto es la desviación estándar. Como se puede ver, la cantidad de correspondencias disminuye en ambos métodos a medida que  $\alpha_2$  aumenta y  $L$  disminuye. Por otro lado, en todos los casos SAR-SIFT encuentra una mayor cantidad de correspondencias que SIFT.

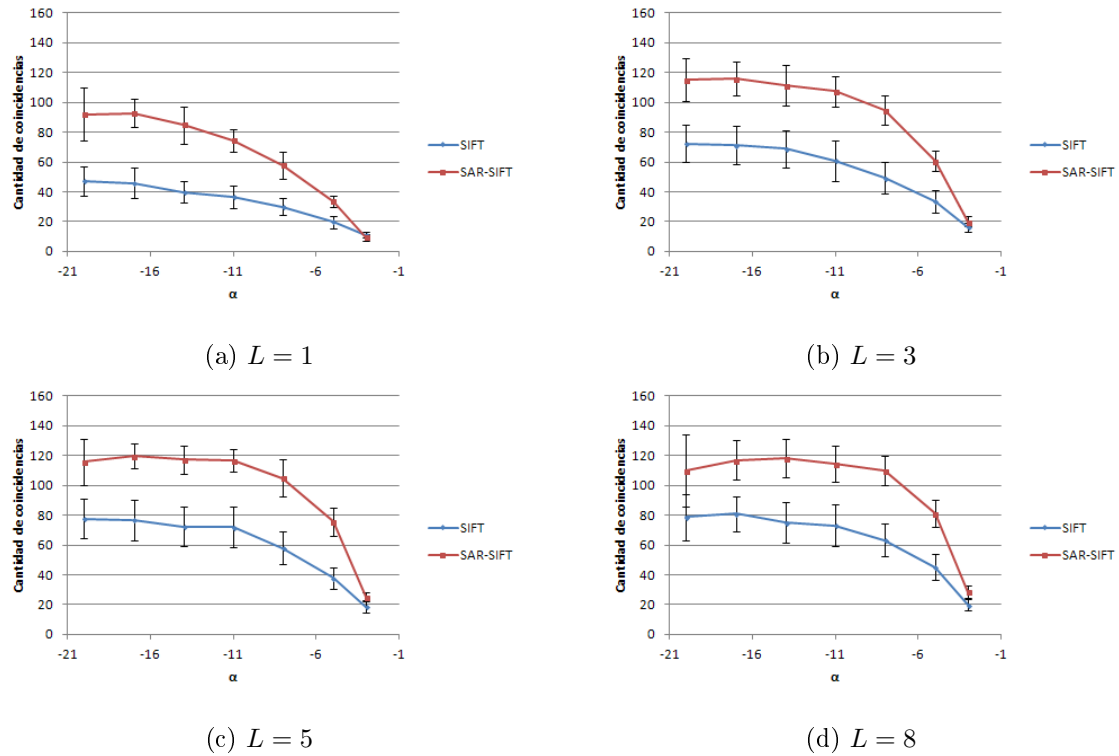


Figura 27: Comparación de cantidad de correspondencias obtenidas con SIFT y con SAR-SIFT entre la imagen de referencia e imágenes sintéticas generadas a partir de la misma imagen binaria con parámetros fijos  $\alpha_1 = -1,5$ ,  $\gamma_1 = \gamma_2 = 1$  y variando  $\alpha_2$  y  $L$ .

En la Figura 28 se pueden ver algunos ejemplos de las imágenes generadas para el segundo experimento sobre imágenes SAR sintéticas que consiste en variar los parámetros  $\gamma_2$  y  $L$ .

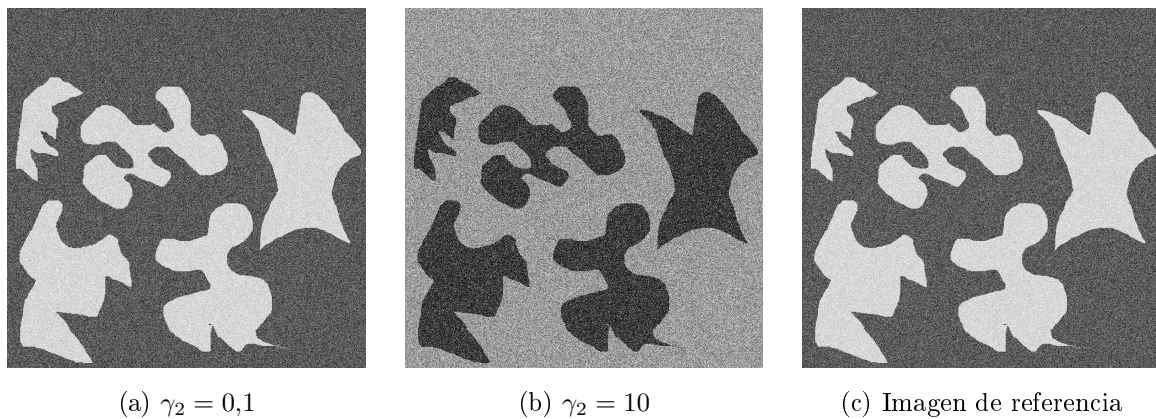


Figura 28: Imágenes SAR sintéticas con parámetros fijos  $\alpha_1 = -1,5$ ,  $\alpha_2 = -3$ ,  $\gamma_1 = 1$ , y  $L = 8$ , y variando  $\gamma_2$  (Figuras 28(a) y 28(b)) junto con la imagen de referencia (Figura 28(c)).

En la Figura 29 se pueden ver los resultados del segundo experimento descrito para imágenes sintéticas. Por lo tanto, se muestra la cantidad de correspondencias en función del valor de  $\gamma_2$  para cada valor de  $L$  y donde el intervalo que se muestra en cada punto es la desviación estándar. Como se puede ver, la cantidad de correspondencias disminuye en ambos métodos a

medida que  $\gamma_2$  aumenta y  $L$  disminuye. Por otro lado, SAR-SIFT sólo encuentra una cantidad significativamente mayor que SIFT para  $\gamma_2 = 0,1$  para los distintos valores de  $L$ , y para  $\gamma_2 = 1$  siempre y cuando  $L$  sea mayor a 1.

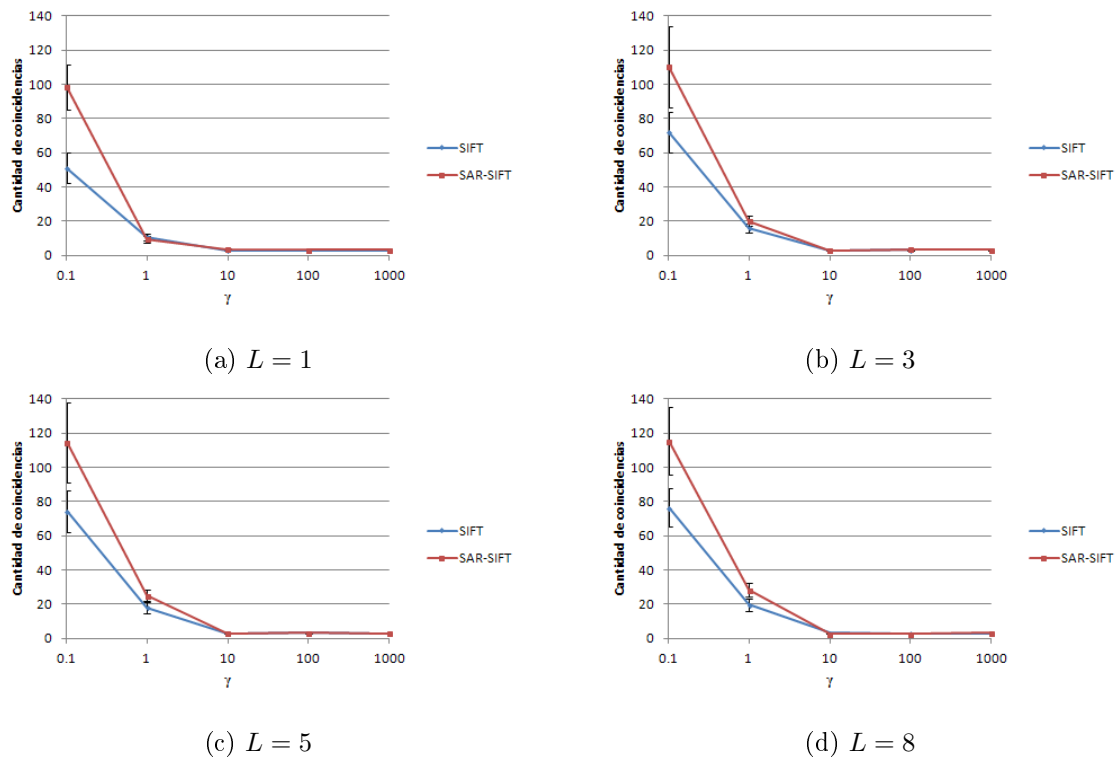


Figura 29: Comparación de cantidad de correspondencias obtenidas con SIFT y con SAR-SIFT entre la imagen de referencia e imágenes sintéticas generadas a partir de la misma imagen binaria con parámetros fijos  $\alpha_1 = -1,5$ ,  $\alpha_2 = -3$ ,  $\gamma_1 = 1$  y variando  $\gamma_2$  y  $L$ .

En la Figura 30 se pueden ver algunos ejemplos de las imágenes generadas para el tercer experimento sobre imágenes SAR sintéticas que consiste en variar el parámetro de escala  $k$ .

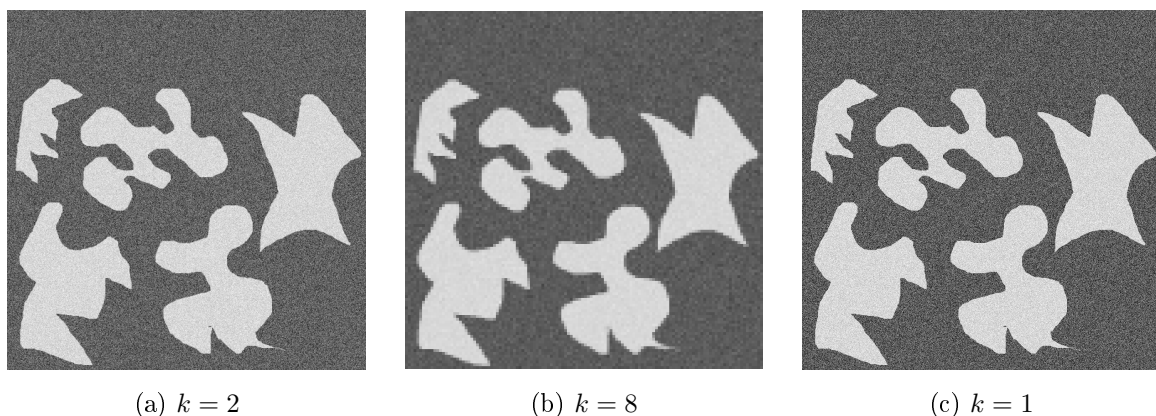


Figura 30: Imágenes SAR sintéticas con parámetros fijos  $\alpha_1 = -1,5$ ,  $\alpha_2 = -20$ ,  $\gamma_1 = \gamma_2 = 1$ , y  $L = 8$ , y variando  $k$ . La imagen de la Figura 30(c) es la imagen de referencia.

En la Figura 31 se pueden ver los resultados del tercer experimento descrito para imágenes sintéticas. Por lo tanto, se muestra la cantidad de correspondencias en función del valor de escala

$k$  y donde el intervalo que se muestra en cada barra es la desviación estándar. Como se puede ver, la cantidad de correspondencias disminuye en ambos métodos a medida que  $k$  aumenta. Por otro lado, en todos los casos SAR-SIFT encuentra una mayor cantidad de correspondencias que SIFT, sin embargo, esta diferencia se hace muy pequeña cuando  $k = 16$ . En adición, para cambios de resolución no muy grandes, SAR-SIFT es más resistente al cambio de escala que SIFT ya que, para los valores de  $k$  1, 2 y 4, la cantidad de correspondencias halladas por SAR-SIFT disminuye más lentamente que la de las halladas por SIFT.

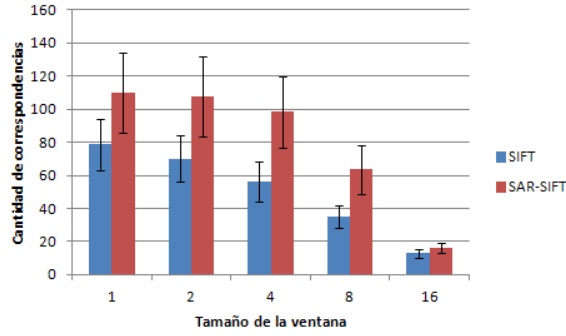


Figura 31: Comparación de cantidad de correspondencias obtenidas con SIFT y con SAR-SIFT entre la imagen de referencia e imágenes sintéticas generadas a partir de la misma imagen binaria con parámetros fijos  $\alpha_1 = -1,5$ ,  $\alpha_2 = -20$ ,  $\gamma_1 = \gamma_2 = 1$ ,  $L = 8$ , y variando el parámetro de escala  $k$ , el cual define el tamaño de la ventana ( $k \times k$ ) que se usa para reducir la resolución de la imagen.

En la Figura 32 se pueden ver algunos ejemplos de las imágenes generadas para el cuarto experimento sobre imágenes SAR sintéticas que consiste en variar el parámetro de rotación  $\theta$  y  $L$ .

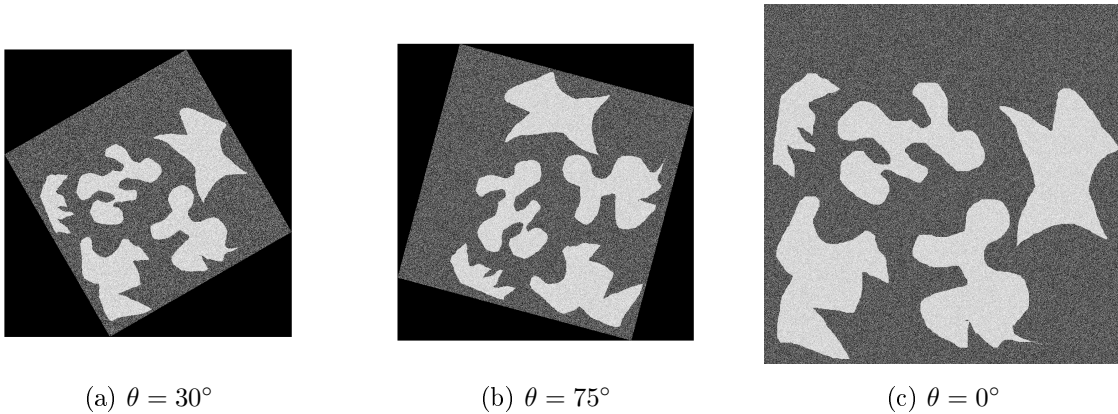


Figura 32: Imágenes SAR sintéticas con parámetros fijos  $\alpha_1 = -1,5$ ,  $\alpha_2 = -20$ ,  $\gamma_1 = \gamma_2 = 1$ , y variando  $\theta$  y  $L$ . La imagen de la Figura 32(c) es la imagen de referencia.

En la Figura 33 se pueden ver los resultados del cuarto experimento descrito para imágenes sintéticas. Por lo tanto, se muestra la cantidad de correspondencias en función del valor de rotación  $\theta$  para cada valor de  $L$  y donde el intervalo que se muestra en cada barra es la desviación

estándar. A partir del mismo se puede notar que SAR-SIFT es más resistente a rotaciones que SIFT para el experimento en cuestión. La cantidad de correspondencias que encuentran ambos métodos decrece para ángulos entre 0 y 45 grados, y luego vuelve a crecer entre 45 y 90. Por otro lado, SAR-SIFT resiste casi totalmente a rotaciones múltiplo de 90 grados. En cambio, la cantidad de correspondencias que encuentra SIFT para cualquier rotación es siempre menor a la mitad de aquellas encontradas sin ninguna rotación, incluso para rotaciones de 180 grados. En adición, a diferencia de los otros experimentos, con valores de  $L = 3$  o  $L = 5$ , ambos métodos encuentran más correspondencias que con  $L = 1$  o  $L = 8$ .

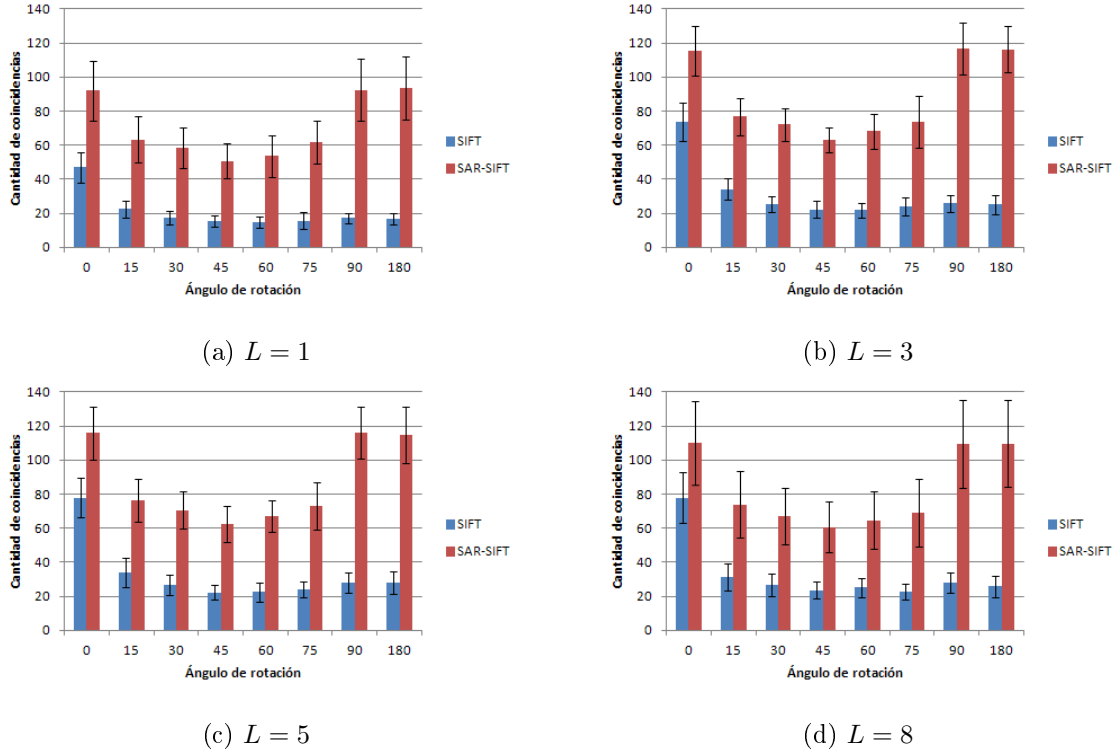


Figura 33: Comparación de cantidad de correspondencias obtenidas con SIFT y con SAR-SIFT entre la imagen de referencia e imágenes sintéticas generadas a partir de la misma imagen binaria con parámetros fijos  $\alpha_1 = -1,5$ ,  $\alpha_2 = -20$ ,  $\gamma_1 = \gamma_2 = 1$  y variando  $\theta$  y  $L$ .

### 6.2.2. Imágenes reales

En la Figura 34 se pueden ver algunos ejemplos de las imágenes generadas para el primer experimento sobre el primer grupo de imágenes SAR reales que consiste en variar el parámetro de escala  $k$ .

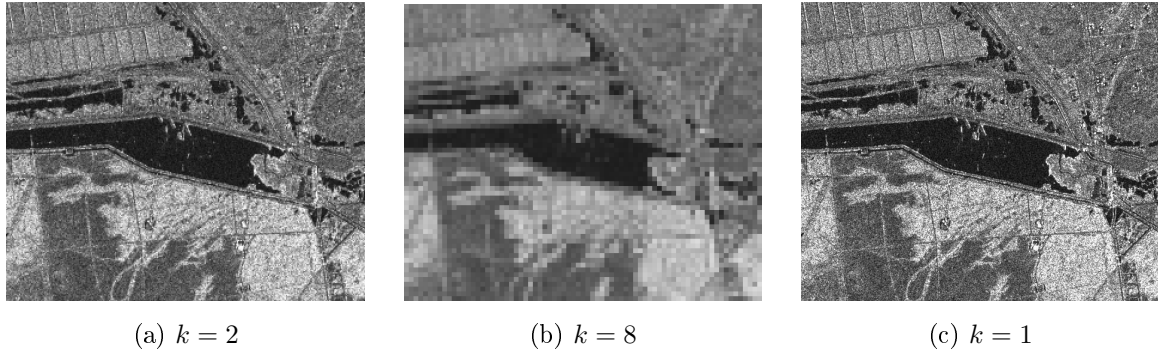
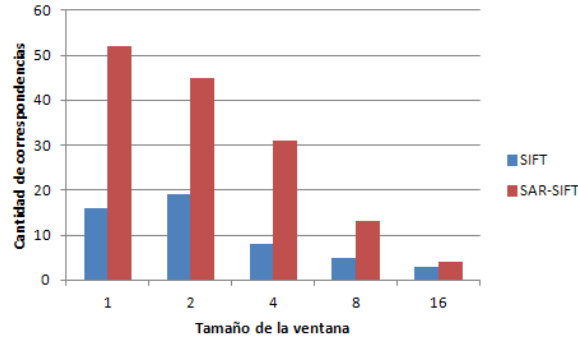
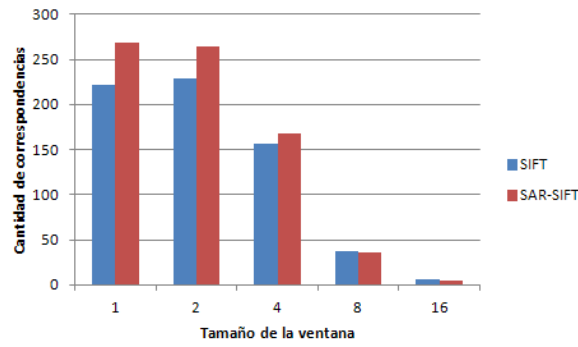


Figura 34: Imagen SAR real variando el parámetro de la resolución  $k$ . La imagen de la Figura 34(c) es la imagen de referencia.

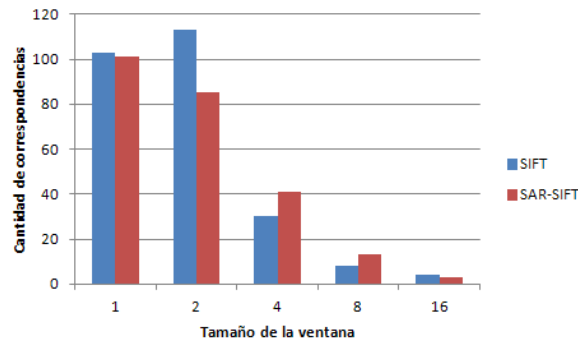
En la Figura 35 se pueden ver los resultados del primer experimento descrito para el primer grupo de imágenes SAR reales. Por lo tanto, se muestra la cantidad de correspondencias en función del valor de escala  $k$ . Como se puede ver, la cantidad de correspondencias aumenta únicamente para el método SIFT cuando se cambia el valor de  $k$  a 2, mientras que en ambos métodos en el resto de los casos disminuye a medida que  $k$  aumenta. Por otro lado, en cuanto a la comparación entre SIFT y SAR-SIFT no se encuentran resultados consistentes, ya que la cantidad de correspondencias halladas no muestra a uno de los dos métodos como superior al otro para todos los pares de imágenes.



(a) Resultados correspondientes a la comparación entre las Figuras 23(a) y 23(b) variando  $k$ .



(b) Resultados correspondientes a la comparación entre las Figuras 23(c) y 23(d) variando  $k$ .



(c) Resultados correspondientes a la comparación entre las Figuras 23(e) y 23(f) variando  $k$ .

Figura 35: Comparación de cantidad de correspondencias obtenidas con SIFT y con SAR-SIFT entre la imagen de referencia entre pares de imágenes reales distintas entre sí correspondientes a una misma región variando el parámetro de escala  $k$ .

En la Figura 36 se pueden ver algunos ejemplos de las imágenes generadas para el segundo experimento sobre el primer grupo de imágenes SAR reales que consiste en variar el parámetro de rotación  $\theta$ .



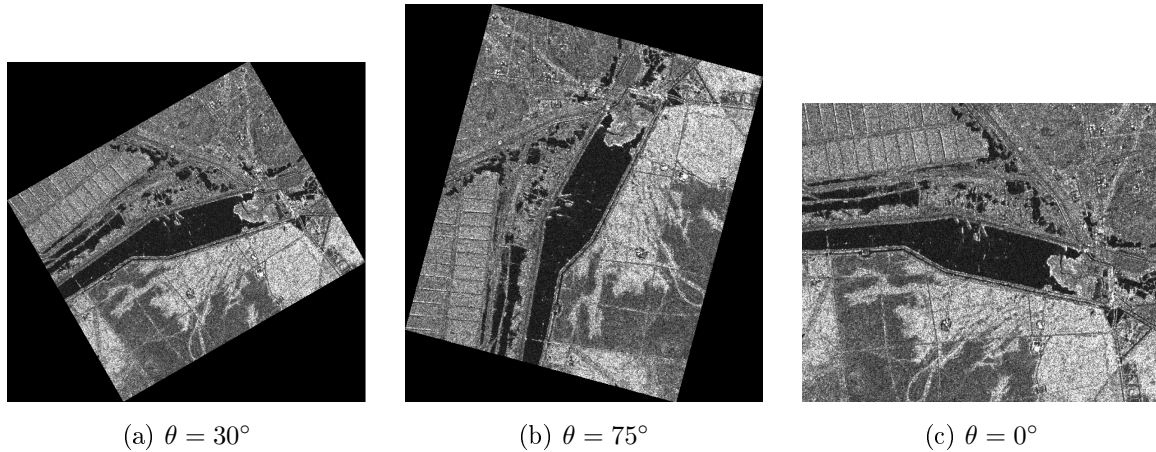
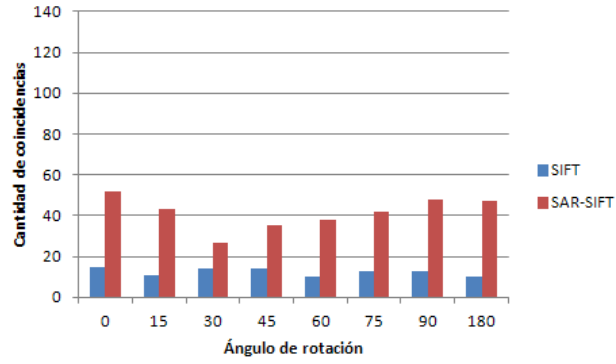
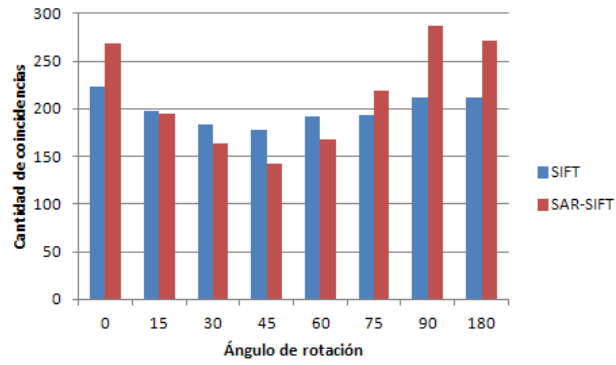


Figura 36: Imagen SAR real variando el parámetro de la rotación  $\theta$ . La imagen de la Figura 36(c) es la imagen de referencia.

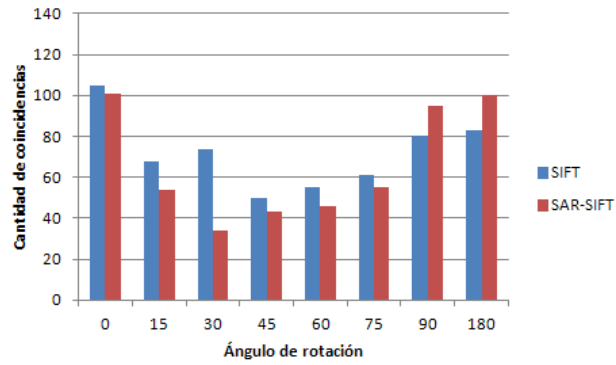
En la Figura 37 se pueden ver los resultados del segundo experimento descrito para el primer grupo de imágenes SAR reales. Por lo tanto, se muestra la cantidad de correspondencias en función del valor de rotación  $\theta$ . La cantidad de correspondencias que encuentran ambos métodos decrece para ángulos entre 0 y 45 grados, y luego vuelve a crecer entre 45 y 90. Por otro lado, SAR-SIFT resiste casi totalmente a rotaciones múltiplo de 90 grados. En adición, en cuanto a la comparación entre SIFT y SAR-SIFT no se encuentran resultados consistentes, ya que la cantidad de correspondencias halladas no muestra a uno de los dos métodos como superior al otro para todos los pares de imágenes. Por último, se puede observar que la cantidad de correspondencias varía más aplicando el método SAR-SIFT que al utilizar el método SIFT.



(a) Resultados correspondientes a la comparación entre las Figuras 23(a) y 23(b) variando  $\theta$ .



(b) Resultados correspondientes a la comparación entre las Figuras 23(c) y 23(d) variando  $\theta$ .



(c) Resultados correspondientes a la comparación entre las Figuras 23(e) y 23(f) variando  $\theta$ .

Figura 37: Comparación de cantidad de correspondencias obtenidas con SIFT y con SAR-SIFT entre la imagen de referencia entre pares de imágenes reales distintas entre sí correspondientes a una misma región variando el parámetro de rotación  $\theta$ .

En la Figura 38 se pueden ver algunos ejemplos de las imágenes generadas para el primer experimento sobre el segundo grupo de imágenes SAR reales que consiste en variar el parámetro de escala  $k$ .

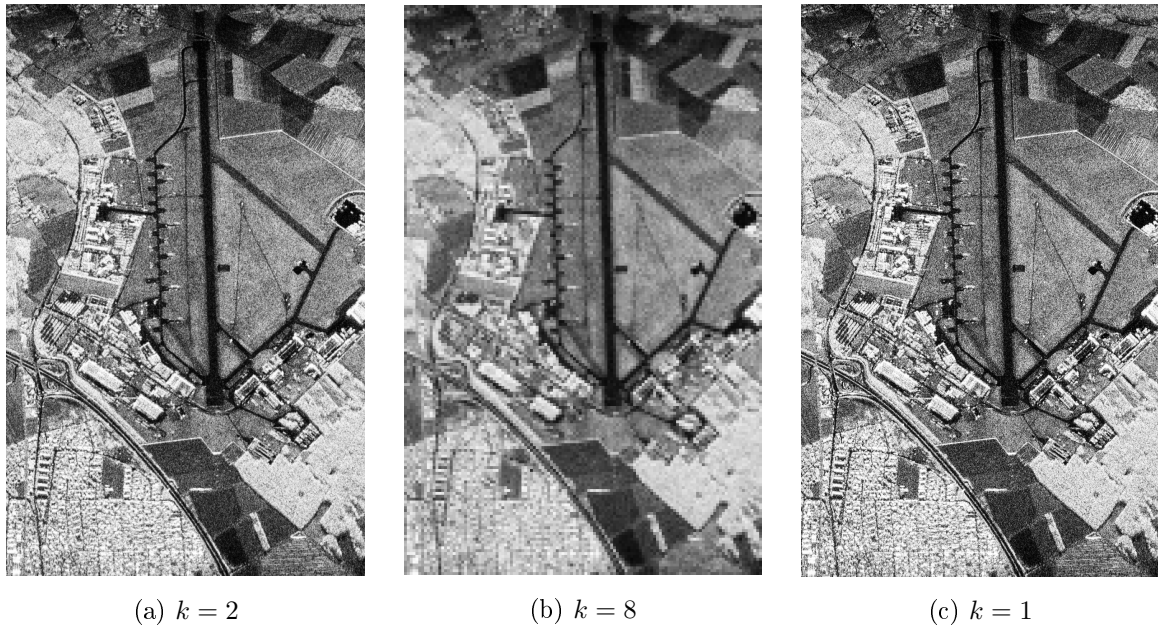


Figura 38: Imagen SAR real variando el parámetro de la resolución  $k$ . La imagen de la Figura 38(c) es la imagen de referencia.

En la Figura 39 se pueden ver los resultados del primer experimento descripto para el segundo grupo de imágenes SAR reales. Por lo tanto, se muestra el promedio de la cantidad de correspondencias en función del valor de escala  $k$ . Como se puede ver, la cantidad de correspondencias es mayor al utilizar el método SIFT, sin embargo, SAR-SIFT resulta ser más resistente a cambios de escala ya que la cantidad de correspondencias halladas disminuye más lentamente que al aplicar el primer método al variar el parámetro  $k$ .

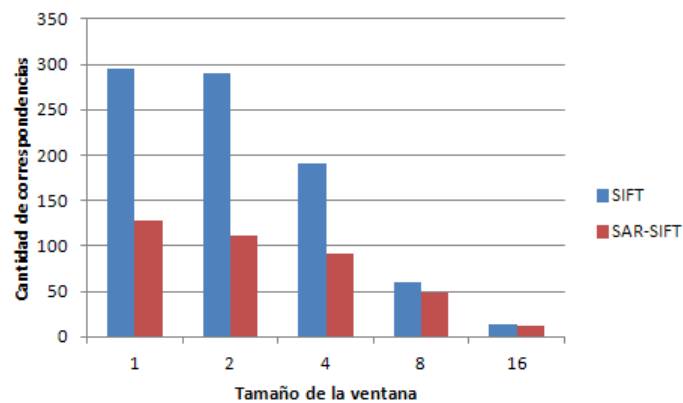


Figura 39: Comparación de cantidad de correspondencias obtenidas con SIFT y con SAR-SIFT entre las bandas HV, VH y VV y la banda HH tomada como imagen de referencia, variando el parámetro de escala  $k$ , el cual define el tamaño de la ventana ( $k \times k$ ) que se usa para reducir la resolución de la imagen.

En la Figura 40 se pueden ver algunos ejemplos de las imágenes generadas para el segundo experimento sobre el segundo grupo de imágenes SAR reales que consiste en variar el parámetro de rotación  $\theta$ .

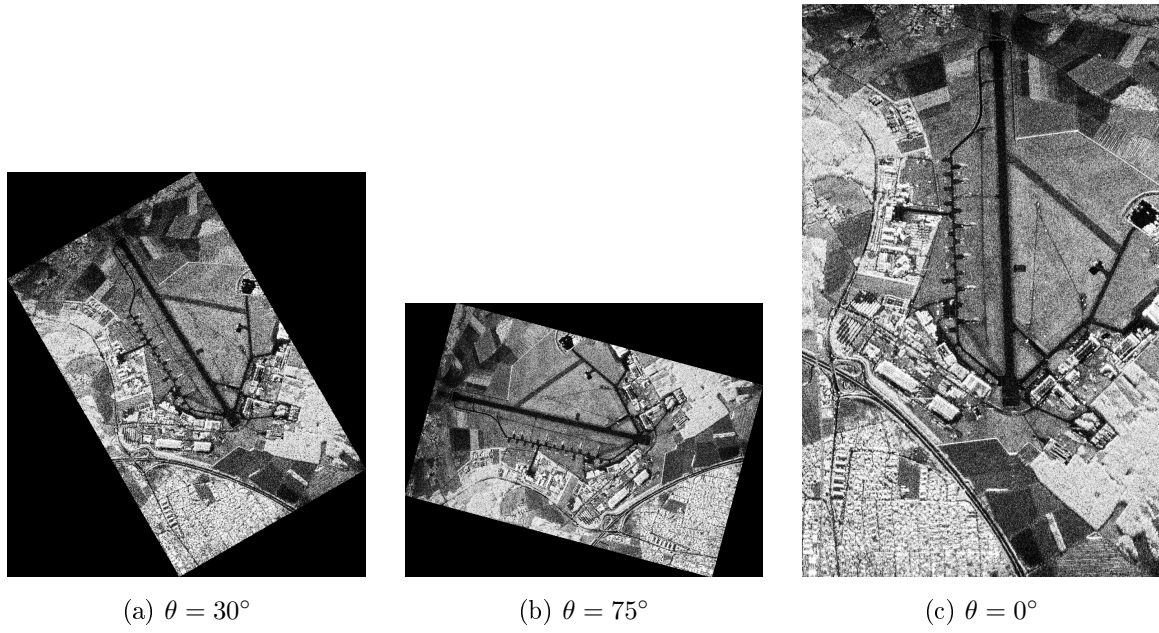


Figura 40: Imagen SAR real variando el parámetro de la rotación  $\theta$ . La imagen de la Figura 40(c) es la imagen de referencia.

En la Figura 41 se pueden ver los resultados del segundo experimento descrito para el segundo grupo de imágenes SAR reales. Por lo tanto, se muestra la cantidad de correspondencias en función del valor de rotación  $\theta$ . La cantidad de correspondencias que encuentran ambos métodos decrece para ángulos entre 0 y 45 grados, y luego vuelve a crecer entre 45 y 90. Por otro lado, SAR-SIFT resiste casi totalmente a rotaciones múltiplo de 90 grados. Por último, se puede observar que la cantidad de correspondencias varía más aplicando el método SIFT que al utilizar el método SAR-SIFT.

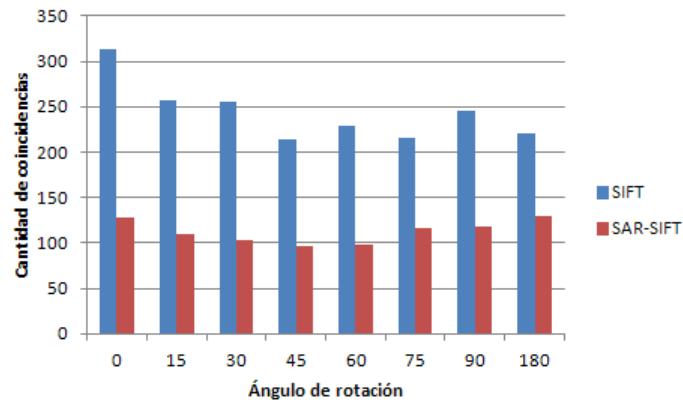


Figura 41: Comparación de cantidad de correspondencias obtenidas con SIFT y con SAR-SIFT entre las bandas HV, VH y VV y la banda HH tomada como imagen de referencia, variando el parámetro de rotación  $\theta$ .

## 7. Conclusiones Finales

En este trabajo se analiza la aplicación del método SIFT y una variante desarrollada para ser aplicada en imágenes SAR llamada SAR-SIFT. En primer lugar se menciona cuál es la importancia de las imágenes SAR, y que la aplicación de métodos de procesamiento de imágenes convencionales, como lo es SIFT, se ven dificultados por el ruido *speckle* que es inherente a este tipo de imágenes. Luego se desarrolla cómo se puede desplegar una imagen SAR a partir de un archivo y cómo mediante el uso de la distribución  $\mathcal{G}_I^0$  se pueden generar imágenes sintéticas. Esto último es posible ya que la distribución mencionada permite modelar los datos provenientes de imágenes SAR. Los parámetros  $\alpha$  relacionado con la textura de la zona,  $\gamma$  correspondiente a la luminosidad y  $L$  que describe la cantidad de *looks* caracterizan a dicha distribución.

En primer lugar se puede observar que el cálculo del mapa de estimaciones del parámetro  $\alpha$  no resuelve el problema de la aplicación del método SIFT tal como se ve en la Figura 25 . Por otro lado, se puede ver que el método SAR-SIFT supera a SIFT al ser aplicado a imágenes sintéticas, frente a distintos valores de los parámetros  $\alpha$ ,  $\gamma$  y  $L$ , además de ser más resistente a cambios de rotación y de resolución. Sin embargo, no se obtienen resultados contundentes que afirmen que SAR-SIFT es mejor que el método SIFT al aplicar a ambos a imágenes reales, puesto que no siempre el primero halla más correspondencias que el segundo.

## Referencias

- [1] Faraj Alhwarin, Danijela Ristic-Durrant, and Axel Gräser. Speeded up image matching using split and extended SIFT features. In *VISAPP 2010 - Proceedings of the Fifth International Conference on Computer Vision Theory and Applications, Angers, France, May 17-21, 2010 - Volume 1*, pages 287–295, 2010.
- [2] Matthew Brown and David G. Lowe. Automatic panoramic image stitching using invariant features. *International Journal of Computer Vision*, 74(1):59–73, dec 2006.
- [3] John Canny. A computational approach to edge detection. In *Pattern Analysis and Machine Intelligence*, page 679–698. IEEE Trans, 1986.
- [4] Julia Cassetti, Juliana Gambini, and Alejandro Frery. Estimación de parámetros utilizando distancias estocásticas para datos con ruido speckle. In *14th Argentine Symposium on Technology*, pages 37–48, 2013.
- [5] Flora Dellinger, Julie Delon, Yann Gousseau, Julien Michel, and Florence Tupin. SAR-SIFT: A SIFT-like algorithm for applications on SAR images. In *2012 IEEE International Geoscience and Remote Sensing Symposium*. IEEE, jul 2012.
- [6] Chris Harris and Mike Stephens. A combined corner and edge detector. In *In Proc. of Fourth Alvey Vision Conference*, pages 147–151, 1988.
- [7] D.G. Lowe. Object recognition from local scale-invariant features. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, volume 2, pages 1150–1157 vol.2, 1999.
- [8] Sanhai Ren, Wenge Chang, and Xiangjun Liu. SAR image matching method based on improved SIFT for navigation system. *Progress In Electromagnetics Research M*, 18:259–269, 2011.
- [9] P. Schwind, S. Suri, P. Reinartz, and A. Siebert. Applicability of the SIFT operator to geometric SAR image registration. *International Journal of Remote Sensing*, 31(8):1959–1980, apr 2010.
- [10] Tao Tang, Deliang Xiang, and Yi Su. A new local feature descriptor for SAR image matching. In *Progress In Electromagnetics Research Symposium Proceedings, Guangzhou, China, Aug. 25-28, 2014*, pages 1823–1827, 2014.