

**Centro de Investigación Científica y de Educación  
Superior de Ensenada, Baja California**



**Programa de Posgrado en Ciencias  
en Ciencias de la Computación**

---

**Sistema de percepción remota y análisis post-vuelo basado  
en fotografía aérea de formato pequeño**

Tesis

para cubrir parcialmente los requisitos necesarios para obtener el grado de  
Maestro en Ciencias

Presenta:

**José Daniel García Rosas**

Ensenada, Baja California, México  
2016

Tesis defendida por

**José Daniel García Rosas**

y aprobada por el siguiente Comité

---

**Dr. Jorge Torres Rodríguez**

**M. en C. Francisco Suárez Vidal †**

**Dr. Hugo Homero Hidalgo Silva**

**Dr. Stephen Holmes Bullock Runquist**



---

**Dra. Ana Isabel Martínez García**

Coordinador del Posgrado en Ciencias de la Computación

---

**Dra. Rufina Hernández Martínez**

Director de Estudios de Posgrado

*José Daniel García Rosas © 2016*

*Queda prohibida la reproducción parcial o total de esta obra sin el permiso formal y explícito del autor*

Resumen de la tesis que presenta **José Daniel García Rosas** como requisito parcial para la obtención del grado de Maestro en Ciencias en Ciencias de la Computación.

## **Sistema de percepción remota y análisis post-vuelo basado en fotografía aérea de formato pequeño**

Resumen aprobado por:

---

Dr. Jorge Torres Rodríguez  
Director de tesis

El uso de la fotografía aérea de formato pequeño (FAFP) en percepción remota ofrece ventajas sobre otras tecnologías convencionales, al contar con mayor resolución espacial y temporal, así como bajos costos de inversión, mantenimiento y logísticos. Al combinar esta tecnología con vehículos aéreos no-tripulados o con aviones ultraligeros el resultado es un amplio potencial de aplicaciones. Para obtener imágenes multiespectrales mediante FAFP se toman dos juegos de fotos, uno en el intervalo del infrarrojo cercano y el otro en el visible. Para su posterior uso en aplicaciones de percepción remota es necesario que las fotografías sean alineadas para formar arreglos multiespectrales y que éstas sean corregidas radiométricamente para poder ser calibradas a los valores de reflectancia reales. Se diseñó e implementó un sistema de software que corrige geométrica y radiométricamente las imágenes adquiridas en una misión aérea. La corrección geométrica se realizó mediante el corregistro de las imágenes, y se obtuvo un error cuadrático medio de alineamiento menor a un pixel. Esto provee suficiente precisión para aplicaciones de mapeo aéreo. El proceso de corrección radiométrica consistió en la corrección del efecto de viñeteo y en la normalización radiométrica aplicada a juegos de imágenes multitemporales. El producto multiespectral generado por el sistema demuestra el potencial que tiene la FAFP para generar información de alta calidad cuando se le aplica el procesado adecuado.

Palabras clave: **Percepción Remota, Fotografía Aérea, Corregistro, Corrección Radiométrica, Corrección de viñeteo**

Abstract of the thesis presented by **José Daniel García Rosas** as a partial requirement to obtain the Master of Science degree in Computer Science

## **Sistema de percepción remota y análisis post-vuelo basado en fotografía aérea**

Abstract approved by:

---

Dr. Jorge Torres Rodríguez  
Thesis director

The use of small format aerial photography (SFAP) in the field of remote sensing offers advantages over traditional technologies, by having greater spatial and temporal resolution, as well as low investment, logistic and maintenance costs. By combining this technology with unmanned aerial vehicles or ultra light aircraft the result is an ample potential for applications. To get multispectral images using SFAP two set of photos are taken, one in the near infrared region and the other on the visible. For using them in remote sensing application is necessary to align the photos to create multispectral arrays and to correct their radiometry to calibrate them to their real reflectance values. A software system that corrects geometrically and radiometrically the images taken from a flight mission was designed. The geometric correction was made using image registration, with a mean squared error under a pixel. That provides enough accuracy for aerial mapping applications. The radiometric correction consisted of normalizing multitemporal image sets and correcting the vignetting effect. The multispectral product generated by the system proves the SFAP potential for generating high quality information when an appropriate process is applied.

Keywords: **Remote Sensing, Aerial Photography, Image Registration, Vignetting Correction**

## **Dedicatoria**

**A mis padres**

## Agradecimientos

Al Doctor Jorge Torres Rodríguez.

A los miembros de mi comité de tesis.

A mis compañeros de la Generación 2013.

Al Centro de Investigación Científica y de Educación Superior de  
Ensenada.

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por su apoyo  
económico para poder realizar mis estudios de Maestría.

A STIG-Geomática por brindarnos acceso a su infraestructura y apoyo en  
la logística de campo.

## Tabla de contenido

	Página
<b>Resumen español</b> .....	ii
<b>Resumen inglés</b> .....	iii
<b>Dedicatorias</b> .....	iv
<b>Agradecimientos</b> .....	v
<b>Lista de figuras</b> .....	ix
<b>Lista de tablas</b> .....	x
<b>Capítulo 1. Introducción</b> .....	1
1.1 Antecedentes .....	2
1.2 Planteamiento del problema.....	3
1.3 Objetivos de la investigación.....	5
1.3.1 Objetivo general.....	5
1.3.2 Objetivos particulares .....	5
1.4 Organización de la tesis.....	5
<b>Capítulo 2. Marco teórico</b> .....	7
2.1 Corrección Radiométrica.....	7
2.1.1 Distorsión agregada durante el proceso de adquisición .....	7
2.1.1.1 Corrección gamma .....	7
2.1.1.2 Matriz de filtros de color.....	8
2.1.2 Corrección del efecto de viñeteo.....	9
2.1.2.1 Polinomio bidimensional.....	12
2.1.2.2 Anillos concéntricos.....	13
2.1.3 Normalización radiométrica relativa .....	14
2.2 Corrección geométrica .....	17
2.2.1 Detección de características .....	18
2.2.1.1 Concepto de Gradiente.....	18
2.2.1.2 Laplaciano de Gaussiana .....	20
2.2.1.3 Diferencia de Gaussianas.....	22
2.2.1.4 Determinante de la Hessiana .....	23

2.2.1.5 Transformada SIFT.....	24
2.2.1.6 Algoritmo SURF.....	26
2.2.2 Empatamiento de características.....	27
2.2.2.1 Relación de distancia entre los vecinos más cercanos.....	28
2.2.3 Estimación del modelo de transformación.....	28
2.2.3.1 Modelos de mapeo globales.....	29
2.2.3.2 Modelos de mapeo locales.....	30
2.2.4 Muestreo y transformación de la imagen.....	31
<b>Capítulo 3. Materiales y Metodología.....</b>	<b>33</b>
3.1 Fotografía aérea.....	33
3.2 Desarrollo del sistema.....	33
<b>Capítulo 4. Sistema de percepción remota.....</b>	<b>35</b>
4.1 Descripción del sistema.....	35
4.1.1 Corrección radiométrica.....	36
4.1.1.1 Normalización radiométrica relativa.....	36
4.1.1.2 Viñeteo.....	37
4.1.2 Corrección geométrica.....	38
4.1.3. Validación post-vuelo.....	38
4.2 Implementación del sistema.....	39
4.2.1 Validación post-vuelo.....	39
4.2.2 Parámetros de cámara.....	42
4.2.3 Normalización relativa.....	43
4.2.4 Corrección de viñeteo.....	45
4.2.5 Corregistro.....	46
<b>Capítulo 5. Resultados.....</b>	<b>48</b>
5.1 Evaluación de la precisión del corregistro de imágenes.....	48
5.1.1 Error de empatamiento.....	49
5.1.2 Error de alineamiento.....	51

5.2 Resultados de la corrección de viñeteo .....	52
5.3 Resultados de la normalización .....	54
<b>Capítulo 6. Conclusiones</b> .....	<b>57</b>
6.1 Trabajo a futuro.....	58
<b>Lista de referencias bibliográficas</b> .....	<b>59</b>

## Lista de figuras

Figura		Página
1	Diagrama general de la arquitectura del sistema.....	36
2	Menú de opciones del sistema.....	39
3	Archivo de texto con las coordenadas de la misión aérea.....	40
4	Mapa de vuelo generado.....	41
5	Comparación entre grupos de imágenes.....	42
6	Configuración de parámetros de cámara.....	43
7	Selección interactiva de puntos estables de alta y baja reflectividad...	44
8	Corrección de viñeteo.....	45
9	Validación antes del correregistro y correregistro individual.....	47
10	Barra de carga durante el correregistro.....	47
11	Imagen visible con sus puntos de interés correspondientes.....	50
12	Imagen infrarroja con sus puntos de interés correspondientes.....	50
13	Perfiles inversos de la función de viñeteo del grupo de imágenes del caso de estudio 'Retorno'. (a) perfil de la banda roja, (b) perfil de la banda verde y (c) perfil de la banda azul.....	53
14	(a) Imagen original. (b) Imagen con el efecto de viñeteo corregido.....	53
15	Imágenes satelitales tomadas en Santa Clara con el sensor Thematic Mapper del satélite Landsat 5. (a) Imagen de referencia tomada en 1979. (b) Imagen objetivo tomada en 1980. (c) Imagen objetivo con radiometría normalizada en base a la imagen de referencia.....	54
16	Histogramas de las 3 bandas del par de imágenes 'Santa Clara' y su normalización. La línea roja coincide con la segunda banda, la línea verde con la tercera y la línea azul con la cuarta. (a) Histograma de la imagen de referencia tomada en 1979. (b) Histograma de la imagen objetivo tomada en el 1980. (c) Histograma de la imagen objetivo con radiometría normalizada en base a la imagen de referencia.....	55

## Lista de tablas

Tabla		Página
1	Error cuadrático medio en grupo de imágenes 'Retorno'.....	52
2	Error cuadrático medio promedio en cada grupo de imágenes...	52
3	Datos estadísticos Santa Clara de 1979.....	55
4	Datos estadísticos Santa Clara de 1980.....	56
5	Datos estadísticos Santa Clara normalizada de 1980.....	56

## Capítulo 1. Introducción

El término *Percepción Remota* se refiere a las tecnologías y procedimientos empleados para adquirir, procesar e interpretar información relacionada con objetos o fenómenos que ocurren en la superficie terrestre, sin establecer contacto físico con éstos. El mecanismo de contacto entre los objetos a distancia y los sensores a bordo de plataformas aéreas o espaciales, es la energía electromagnética radiada por la superficie terrestre (Campbell, 2011; Rees, 2013). De esta manera, los diversos tipos de cubierta del terreno presentan características particulares de emisión y absorción de energía en diferentes regiones del espectro electromagnético. Esto les imprime una *firma espectral* distintiva que permite su identificación mediante el uso de técnicas de procesamiento y análisis de imágenes adquiridas con sensores remotos (Schowengerdt, 2009; Richards y Jia, 2005; Jensen, 2007).

El desarrollo tecnológico ha traído como resultado una amplia variedad de aplicaciones de la percepción remota, las cuales van desde aplicaciones a la ingeniería civil y estudios de asentamientos humanos (Kennie y Matthews, 1985; Ridd y Hipple, 2006); aplicaciones en el manejo de recursos naturales y agricultura (Ustin, 2004; Paalman, 2015); así como en geología, oceanografía y ecología (Gupta, 2003; Ryerson, 2006; Horning et al., 2010).

La tecnología actual de percepción remota tuvo su origen en la década de los años 1970's con el inicio de la serie de satélites Landsat, originalmente denominados ERTS (Earth Resources Technology Satellite) (USGS, 1979; Sabins, 1997). Los sensores que han operado en los satélites Landsat han estado en constante evolución, desde el primer Multi Spectral Scanner (MSS) a bordo de Landsat 1; así como su posterior versión mejorada, el Thematic Mapper (TM, ETM) a bordo de Landsat 4, 5 y 7; y finalmente el reciente sensor Operational Land Imager (OLI) a bordo de Landsat 8. Estos sensores se caracterizan por ser de banda ancha, operando en regiones del espectro visible, infrarrojo cercano, e infrarrojo medio y térmico. Y han servido además como estándares de referencia para las nuevas tecnologías en satélites de alta

resolución espacial, como por ejemplo: Ikonos, QuickBird, GeoEye, WorldView, etc. (Gomarasca, 2009).

En las últimas décadas la *Fotografía Aérea de Formato Pequeño* (FAFP) ha surgido como un campo tecnológico emergente, y como una alternativa a la fotografía aérea tradicional y a las imágenes de satélite de alta resolución espacial. El avance tecnológico y la constante reducción de precios en equipo de cómputo y dispositivos de almacenamiento, así como la versatilidad de las nuevas cámaras digitales tipo DSLR, hacen que la FAFP ofrezca ventajas debido a su accesibilidad, menor costo de mantenimiento y logística de operación. Lo anterior ha traído como consecuencia un auge en la integración de estos sistemas basándose en plataformas aéreas de diferentes tipos, que van desde el uso de globos y papalotes, hasta aviones ultraligeros y drones no tripulados (Aber et al., 2010; Torres, 2012).

En este trabajo se propone el desarrollo de un sistema de percepción remota basado en fotografía aérea de formato pequeño, así como de un sistema de análisis post-vuelo que permita el manejo, corrección y validación de cubrimientos adquiridos en misiones aéreas usando plataformas tripuladas (ultraligeros) y no-tripuladas (drones).

## **1.1 Antecedentes**

Operando en plataformas de vuelo bajo, la FAFP presenta ventajas significativas con resoluciones espaciales de unos cuantos centímetros y con capacidades para programar una misión de vuelo de acuerdo a los tiempos y requerimientos específicos de un determinado proyecto. Debido a su escala de observación, resolución espacial y altura de vuelo, la FAFP ocupa un nicho especial entre las tecnologías de percepción remota empleadas para la observación directa en el campo, y la fotografía aérea convencional e imagen de satélite. La FAFP es particularmente útil en estudios a gran escala de sitios con cubrimiento limitado. Ejemplos de su aplicación en estudios geomorfológicos relacionados con procesos glaciales, erosión y degradación de suelos, así como el impacto del clima y el sobre-pastoreo en la vegetación pueden ser consultados en Aber et al. (2008), Dickenson (2009), Ries y Marzloff (2003), Ries y Hirt

(2008), Marzoff y Ries (2007), y en Seeger et al. (2009). Otros campos de aplicación incluyen el mapeo topográfico, el monitoreo y manejo de recursos naturales, el análisis del cambio de cubierta y uso del suelo, así como estudios de medio ambiente y desarrollo urbano (Hofstee, 1985; Chanond y Leekbhai, 1986; Fryer, 1992; Aldred y Hall, 1992; Aber et al., 2010).

En la actualidad la FAFP se basa en cámaras digitales de peso ligero, incluyendo desde los modelos comerciales tipo “point and shot” orientados al consumidor, hasta los modelos profesionales tipo DSLR (Digital Single Lens Reflex) en sus versiones de formato APS-C y de formato completo. Desde el punto de vista de la logística, el vuelo de baja altura y el mapeo a escalas grandes es viable utilizando varios tipos de plataformas aéreas tripuladas y no-tripuladas. Debido a esto la FAFP se puede adquirir en situaciones donde la operación de aviones grandes puede ser impráctica, riesgosa o hasta imposible. Por otra parte la tecnología de FAFP es de alta portabilidad, de fácil configuración y puesta a punto en el campo, y no requiere personal altamente entrenado. Lo anterior hace de la FAFP una herramienta logísticamente apropiada en un amplio espectro de aplicaciones (Torres, 2012).

Actualmente en el Laboratorio de Geomática del CICESE se lleva a cabo una iniciativa, cuyo objetivo es la implementación de sistemas para el manejo de misiones de fotografía aérea basados en cámaras de formato pequeño y plataformas aéreas tripuladas y no-tripuladas (Torres, 2010; Ramírez y Torres, 2012; Ramírez, 2013).

## **1.2 Planteamiento del problema**

Se identificaron tres problemas a resolver, con el objetivo de implementar un sistema de percepción remota basado en fotografía aérea de formato pequeño. El primer problema consiste en corregir radiométricamente las imágenes. Esta corrección implica la eliminación del efecto de viñeteo originado por la lente, y una normalización radiométrica necesaria para intercalibrar imágenes adquiridas en diferentes fechas o bajo diferentes condiciones de iluminación. El segundo problema consiste en integrar un arreglo multispectral a partir de la alineación geométrica de las bandas

provenientes de fotografías adquiridas por dos cámaras independientes (una de ellas modificada para captar el infrarrojo). Finalmente, el tercer problema consiste en el diseño de una herramienta computacional que permita hacer la validación del cubrimiento de una misión aérea, además de administrar y manipular la base de datos fotográfica.

Para abordar el problema del viñeteo nos basaremos en las técnicas propuestas por Yu (2004) y Goldman (2010), mientras que la implementación de la normalización radiométrica se llevará a cabo tomando como base los algoritmos propuestos por Elvidge et al. (1995) y el trabajo desarrollado por Morán (2014).

Aunque son comunes los trabajos donde se utilizan imágenes multispectrales adquiridas por medio de drones o algún tipo de sistema no-tripulado (e.g. Berni, 2009; Levin, 2011; Saari, 2011), existe poca información sobre técnicas para hacer el correregistro de imágenes. Sin embargo, en el campo de visión por computadora es común encontrar información sobre trabajos relacionados con el emparejamiento de características entre imágenes, algunos de ellos se discuten a continuación.

El algoritmo SIFT (Lowe, 2004) es uno de los más populares para encontrar la correspondencia de características. Debido a que SIFT sólo funciona con imágenes en tonos de gris, al tener que convertir imágenes multispectrales esto puede originar la pérdida de información radiométrica. Esto reduce la utilidad de las características extraídas, además de que se pierde información espectral que puede ser valiosa para diferenciar características espectrales en una imagen, reduciendo además el número de puntos de interés obtenidos.

Para cubrir estas deficiencias, otros autores han propuesto modificaciones a SIFT para casos específicos. Algunos maximizan las coincidencias correctas cambiando los parámetros originales (Sima y Buckley, 2013), mientras que otros cambian la representación de la imagen para mejorar su correspondencia (Li et al, 2014). Estos se apoyan de la relación espacial entre puntos de interés (Hasan et al, 2012), o consideran un enfoque híbrido, como en el caso de apoyar a SIFT mediante procesamiento adaptable por bloques (Lee y Choi, 2013).

En misiones aéreas se obtienen grandes cantidades de imágenes de alta resolución, por lo que es deseable que el algoritmo de empatación de características sea lo más rápido posible. Aunque existen algoritmos más rápidos que SIFT, como es el caso de SURF (Bay et al., 2008), sin embargo son pocos los trabajos que consideran su aplicación para imágenes multiespectrales (Teke y Temizel ,2010).

## **1.3 Objetivos de la investigación**

### **1.3.1 Objetivo general**

Diseñar e implementar un sistema que permita la corrección geométrica y radiométrica de imágenes adquiridas por diferentes sensores para formar una imagen multi-espectral, así como un sistema de análisis post-vuelo que permita validar la cobertura y control de calidad en misiones de fotografía aérea de formato pequeño.

### **1.3.2 Objetivos particulares**

1. Diseñar e implementar un algoritmo para la alineación y corrección de imágenes.
2. Diseñar e implementar un algoritmo de corrección radiométrica por viñeteo.
3. Diseñar e implementar un algoritmo para normalización radiométrica.
4. Desarrollar una herramienta computacional para el análisis post-vuelo de misiones de fotografía aérea de formato pequeño.

## **1.4 Organización de la tesis**

La tesis se organiza de la siguiente manera: en el Capítulo 2 se presentan los conceptos teóricos que fundamentan los algoritmos utilizados para la corrección geométrica y radiométrica de las imágenes. En el Capítulo 3 se hace referencia a los materiales utilizados como parte de la investigación y la metodología que se siguió. En el Capítulo 4 se describe el sistema de cómputo que se desarrolló. Los resultados obtenidos se presentan en el Capítulo 5. Y finalmente, en el Capítulo 6 se presentan las

conclusiones del proyecto de tesis, así como algunas sugerencias para el posible trabajo a futuro.

## Capítulo 2. Marco teórico

---

En este capítulo se presentan los problemas a resolver para que las imágenes obtenidas mediante un vehículo aéreo tripulado (o no tripulado) puedan ser utilizadas en aplicaciones de percepción remota y se describe el fundamento teórico de los algoritmos implementados en el sistema.

### 2.1 Corrección Radiométrica

La información de los valores de intensidad o brillo en una imagen son de utilidad en aplicaciones en donde es necesaria una precisión fotométrica para el análisis cuantitativo de los datos en las imágenes objetivo. Por lo tanto se requiere que la intensidad registrada en la imagen corresponda con la luminosidad original del objeto, o en el caso que esto no sea posible, que las imágenes mantengan consistencia radiométrica entre sí. Para lograr esto es necesario corregir todas aquellas distorsiones radiométricas introducidas durante el proceso de adquisición de imágenes.

#### 2.1.1 Distorsión agregada durante el proceso de adquisición

Al adquirir una imagen mediante un dispositivo digital comercial, como en el caso de las cámaras de formato pequeño, esta pasa por varios procesos que la modifican con fines estéticos o de ahorro de memoria, tales procesos ocasionan una distorsión radiométrica que afecta el uso de imágenes adquiridas con tal dispositivo para fines cuantitativos.

##### 2.1.1.1 Corrección gamma

La corrección gamma es uno de estos procesos, en donde se modifica el valor de intensidad de los píxeles con fines estéticos. Por ejemplo, el sistema visual humano se comporta de una manera no lineal en el cual se detecta un mayor rango de tonos oscuros que de alto brillo. Debido a esto el sistema de procesamiento de imágenes

embebido en el dispositivo trata de emularlo, aplicando un factor gamma (por ejemplo, 1 / 2.2 en el caso del espacio de color sRGB, que es el más usado) a la cantidad de fotones recibidos en un píxel dado. Entonces, si se recibe el doble de la cantidad de fotones, en lugar de generar un valor doble al original (el cual sería un comportamiento lineal) sigue una curva definida por el valor gamma.

La función de ajuste gamma es la siguiente

$$V_{out} = A * V_{in}^{\gamma} \quad (1)$$

en donde  $V_{in}$  es el valor de luminancia real,  $A$  es una constante,  $\gamma$  es el factor gamma y  $V_{out}$  es el valor de luz detectada en el rango de intensidad definido.

Cuando el valor del factor gamma es menor a 1, la función de ajuste se considera una compresión, por ejemplo cuando una cámara digital guarda una foto. Cuando el valor del factor gamma es mayor a 1, la función de ajuste se trata de una expansión, que es lo que sucede cuando se corrige el ajuste gamma de una foto al desplegarla en una computadora. Para realizar tal corrección es necesario conocer el factor gamma con el que fue comprimida, y utilizar el inverso de este valor.

### 2.1.1.2 Matriz de filtros de color

Una matriz de filtros de color (color filter array, CFA por sus siglas en inglés) es un mosaico que se coloca sobre los pozos de fotones en una cámara digital, tales pozos se abren al presionar el botón de captura de la cámara para capturar fotones y la cantidad de fotones en cada pozo indica el nivel de intensidad en esa posición. El filtro de color es necesario para distinguir cuántos fotones de cada color primario (rojo, verde, y azul) se capturaron, permitiendo la entrada de un solo color en cada pozo y bloqueando los otros dos. Como sólo se captura un color por cada pozo, es necesario que la cámara aproxime los otros dos colores primarios para poder tener un color completo en cada píxel.

La matriz de filtros de color más popular es la Bayer, en donde se alternan renglones de filtros rojo-verde y azul-verde. Se cuenta con el doble de sensores verdes para aproximar como el sistema visual humano es más sensible a la luz verde, lo que ayuda a que se produzca una imagen con menos ruido y mayor detalle después de la interpolación para obtener color en cada pixel.

Esta interpolación cromática, en el caso de una matriz Bayer, consiste en obtener los otros dos colores no presentes en el píxel. En el caso de un pixel rojo o azul, para obtener su valor verde se debe de promediar los valores de sus vecinos en un patrón de cruz (vertical y horizontal):

$$\tilde{g}(i,j) = \frac{1}{4} \sum g(i+m, j+n), (m,n) = \{(0,-1), (0,1), (-1,0), (1,0)\} \quad (2)$$

para los colores rojo y azul se aplica la misma ecuación, pero con un patrón diagonal.

En el caso de un píxel verde la ecuación sólo promedia dos valores, los horizontales en el caso del color rojo y los verticales para el color azul.

Esta interpolación es el origen de la distorsión radiométrica introducida a la imagen durante este proceso, además de no poder contar con los valores originales para cada banda espectral que son de utilidad para su análisis cuantitativo.

Posteriormente el sistema de procesamiento de imágenes de la cámara reduce la resolución radiométrica (de 12 a 8 bits), recorta los bordes de la imagen y la comprime (usualmente guardándola en el formato JPEG).

### **2.1.2 Corrección del efecto de viñeteo**

El viñeteo es un efecto de pérdida de intensidad gradual en las regiones periféricas de una imagen debido a condiciones durante su adquisición. Esto ocasiona que sean más

oscuras de lo que deberían ser ciertas regiones de la imagen, por lo que es deseable corregir este efecto para poder trabajar con los valores originales de intensidad.

El efecto de viñeteo durante el proceso de adquisición de una imagen se puede caracterizar en la siguiente función, que muestra la distorsión radiométrica del valor original de radiancia del píxel  $x$  en la imagen  $i$  después de su adquisición:

$$P_{x,i} = R(t_i L_x M(r_{x,i})) \quad (3)$$

en donde  $R$  es la función de respuesta de la cámara,  $M$  es la función de viñeteo,  $t_i$  es el tiempo de exposición del marco  $i$ ,  $L_x$  es la radiancia original del punto en dirección a la cámara y  $r_{x,i}$  es la distancia de la proyección de  $x$  en la imagen  $i$  desde el centro óptico del marco.

Existen distintas fuentes de origen del viñeteo, las cuales clasifican el tipo de viñeteo producido (Goldman, 2010). El viñeteo natural se refiere al declive radial debido a la geometría de la óptica, en donde diferentes regiones del plano de la imagen reciben diferente irradiancia. Por otro lado, en el caso de las cámaras digitales el viñeteo se produce por la sensibilidad angular de la óptica digital, ya que los pozos de fotones en los sensores digitales cuentan con una profundidad finita, lo que causa que la luz incidente sobre un pozo de fotones en un ángulo inclinado sea parcialmente bloqueada por los lados del pozo. El viñeteo óptico, o artificial, ocurre cuando el paso de la luz es bloqueado desde dentro del cuerpo de la lente por el diafragma de la lente. Este efecto es reducido al disminuir la apertura, ya que una apertura menor limita el paso de la luz de igual manera al centro y las orillas del marco. Finalmente, el viñeteo mecánico es aquel en donde algunos caminos de luz son bloqueados por otros elementos físicos de la cámara, generalmente se tratan de filtros o parasoles que se conectan con el cuerpo del lente.

Debido a que existen distintas causas del efecto de viñeteo, es difícil predecir el grado de viñeteo dadas las características de la lente y el sistema de adquisición.

El declive o la pérdida de intensidad gradual se puede definir como

$$falloff = (d/f)^2 \quad (4)$$

en donde  $d$  es la distancia desde el centro de la imagen y  $f$  es la distancia focal de la cámara.

Una de las maneras más simples de corregir el efecto de viñeteo para toda imagen obtenida con el mismo sistema de adquisición y con las misma configuración, es el preparar una fuente de iluminación blanca con un nivel de intensidad conocido sobre un objeto de referencia con un bajo reflejo especular, en donde las variaciones de intensidad se pueden atribuir al efecto de viñeteo. A este método se le llama corrección de campo plano (flat-field correction) (Yu, 2004). Al apuntar una cámara hacia la superficie de referencia se obtiene la respuesta a la intensidad en cada posición de pixel. Y en cada pixel se calcula la siguiente fórmula:

$$I_{LUT}(i,j) = I_{ref,max}/I_{ref}(i,j) \quad (5)$$

en donde  $I_{ref}(i,j)$  es el valor de intensidad en la posición de píxel  $(i,j)$ ;  $I_{ref,max}$  el valor máximo de  $I_{ref}(i,j)$ ; e  $I_{LUT}(i,j)$  es un factor de corrección a ser guardado en una tabla de consulta (LookUp Table) en la posición  $(i,j)$ . A partir de ahí, una imagen capturada con la misma cámara puede ser corregida multiplicando los valores de píxel de la imagen con factores de corrección correspondientes guardados en la tabla de consulta:

$$I'(I,J) = I(i,j) * I_{LUT}(i,j) \quad (6)$$

en donde  $I(i,j)$  e  $I'(I,J)$  son los valores de intensidad en la posición de píxel  $(i,j)$  antes y después de la corrección de viñeteo respectivamente.

El método de tabla de consulta es el más preciso para la corrección de viñeteo pero depende de un alto requerimiento de memoria dependiendo del tamaño de la imagen. Existen otros métodos que calculan el perfil a partir de una sola imagen sin conocer previamente las condiciones de su adquisición sacrificando precisión (Zheng et al, 2009) o a partir de una secuencia de imágenes como en el caso de una imagen panorámica (Goldman, 2010).

Debido a que el método de campo plano requiere obtener una imagen bajo ciertas condiciones para sólo funcionar con imágenes obtenidas bajo estas mismas condiciones, se han investigado otros algoritmos que permitan corregir el viñeteo sin tener conocimiento previo del sistema de adquisición. Uno de ellos se presta a nuestro caso ya que requiere de un gran conjunto de imágenes para aproximar la imagen de referencia mediante el promedio de ellas (Lebourgeois et al, 2008). Esto se debe a que dada una secuencia de imágenes en donde la geometría de la lente se mantiene fija entre capturas, el efecto de viñeteo es el mismo en cada imagen.

Bajo este enfoque se calcula una imagen promedio para cada banda espectral, suponiendo que son directamente proporcionales a la luminancia y se usan para calcular y modelar los perfiles radiométricos. Estos perfiles radiométricos se usan para modelar el efecto de viñeteo, ajustando el perfil a modelos empíricos como funciones polinomiales o hiperbólicas, mediante métodos de mínimos cuadrados. A continuación se presentan dos enfoques distintos para corregir el viñeteo partiendo de la imagen de referencia.

#### **2.1.2.1 Polinomio bidimensional**

Dada una imagen de referencia (en nuestro caso, obtenida mediante el promedio de un conjunto grande de imágenes) se calcula un filtro de corrección de viñeteo para cada banda espectral ajustando una función de distribución polinomial bidimensional sobre la imagen de referencia mediante el método de mínimos cuadrados. Debido a que es bidimensional el polinomio, se expresa el factor de viñeteo para una posición dada en la imagen como una función de coordenadas de columna y de renglón. De esta manera

se obtiene el grado de polinomio óptimo para cada banda espectral de manera empírica. La función polinomial resultante se usa para crear un filtro máscara que se le aplica multiplicativamente a cada banda para eliminar su efecto de viñeteo.

### 2.1.2.2 Anillos concéntricos

Otro método para corregir el efecto de viñeteo, dada una imagen de referencia, es el promediar los valores de intensidad para cada banda espectral a cierta distancia del centro de la imagen  $(x_0, y_0)$ , es decir, seccionando la imagen en una cantidad  $N$  de anillos concéntricos definidos por el radio desde el centro de la imagen hacia el anillo. Como el viñeteo es simétrico radialmente desde el centro de la imagen, el declive de intensidad se puede parametrizar por un radio  $r$ .

Esto permite aproximar el efecto radial que se presenta en el viñeteo, al crear una tabla de consulta para obtener el valor promedio en cada banda espectral para un valor de radio dado.

$$T_{rad} = \frac{1}{n} \sum (x, y) \text{ para todo } (x, y) \text{ que cumpla con } \frac{\sqrt{(x-x_0)^2 + (y-y_0)^2}}{r_{max}} = rad \quad (7)$$

$$r_{max} = \sqrt{x_0^2 + y_0^2}$$

en donde  $rad$  es el radio en relación del punto  $(x, y)$  con el punto central  $(x_0, y_0)$  y  $n$  es el total de píxeles para un valor de radio dado, representado en la ecuación como la variable  $rad$ .

A estos valores de intensidad se les aplica el logaritmo natural para después ajustar los valores de radio de los anillos a polinomios de 7mo grado. Esto se efectúa para cada banda espectral, así como para el promedio de los 3 valores del visible, para la corrección en tonos de grises.

Los coeficientes polinomiales calculados se usan para corregir el viñeteo dividiendo cada valor de píxel en la imagen por el valor del polinomio para el radio de ese píxel en particular.

### 2.1.3 Normalización radiométrica relativa

Para realizar un análisis multitemporal, en donde se comparan grupos de imágenes de la misma zona pero tomadas en diferentes fechas, se requiere que la radiometría de los distintos grupos de imágenes sea lo más cercana posible, ya que si se pretende hacer un análisis cuantitativo de estos datos, las diferencias radiométricas inter-temporales pueden introducir ruido en el modelo de análisis utilizado. Estas diferencias radiométricas se producen ya sea por un cambio en las características del sistema de adquisición de imágenes o por cambios en las condiciones atmosféricas y/o la iluminación por parte del sol en las distintas fechas.

Las técnicas tradicionales para la normalización radiométrica relativa se basan en la comparación lineal de las características estadísticas de dos imágenes. Estas técnicas consisten en una transformación lineal de la forma

$$\hat{X}_k = a_k X_k + b_k \quad (8)$$

en donde  $X_k$  es la k-ésima banda de la imagen objetivo  $X$ ,  $\hat{X}_k$  es la k-ésima banda de la imagen normalizada resultante, y  $a_k, b_k$  son la ganancia (gain) y compensación (offset) respectivamente, las cuales se usan para lograr la normalización de la banda k en la imagen  $X$ .

Un método propuesto para solucionar la diferencia radiométrica es el de Regresión Automática Controlada por Diagramas de Dispersión (Elvidge et al, 1995) (Automatic Scattergram Controlled Regression - ASCR). En este método se define como imagen de referencia aquella de la fecha que se considera que su radiometría es la más adecuada, y una segunda imagen objetivo adquirida en otra fecha a la cual se le

aplicará el algoritmo para normalizar su radiometría con base a la de referencia. El procedimiento está basado en la condición de que para una banda espectral dada, existe una relación lineal entre los valores de los pixeles para dos imágenes geográficamente corregistradas en la misma área.

El procedimiento ASCR consiste en las cuatro siguientes etapas:

1. Se calculan los diagramas de dispersión de las imágenes X y Y para sus bandas en el infrarrojo cercano.
2. Se realiza una búsqueda para localizar los centros de agrupamiento correspondientes a las superficies de baja reflectividad, típicamente agua,  $(i_{lmax}, j_{lmax})$  y las superficies de alta reflectividad, tierra  $(i_{umax}, j_{umax})$ , basándose en máximos locales en los diagramas de dispersión. Una vez que se identificaron los centros se procede a encontrar los coeficientes a,b (ganancia y compensación respectivamente) para calcular una región inicial de no cambio (NC) para las bandas  $k = 3,4$  (bandas infrarrojas).

$$a_{k0} = \frac{j_{umax\ k} - j_{lmax\ k}}{i_{umax\ k} - i_{lmax\ k}} \quad (9)$$

$$b_{k0} = j_{lmax\ k} - a_{k0} * i_{lmax\ k} \quad (10)$$

La región de no cambio (NC) depende de la mitad de ancho vertical (Half Vertical Width - HVW) que ayuda a obtener una cota de la región, esta se obtiene con la siguiente ecuación:

$$HVW_k = HPW \sqrt{1 + a_{k0}^2} \quad (11)$$

en donde HPW es la mitad de ancho perpendicular (Half Perpendicular Width) cuyo valor depende del ancho de los agrupamientos de agua y tierra. Elvidge, et al. (1995) proponen un valor inicial de 10.

El conjunto de píxeles incluidos en la región de no cambio se selecciona como

$$NC = (x, y): |y_3 - b_{30} - a_{30} x_3| \leq HVW_3 \cap: |y_4 - b_{40} - a_{40} x_4| \leq HVW_4 \quad (12)$$

3. Se calculan líneas de regresión para todas las bandas de la imagen multiespectral. Sólo se usan aquellos píxeles que cayeron dentro de la región de no cambio. Los coeficientes de regresión para la banda k se calculan a partir de una ecuación de mínimos cuadrados:

$$Q_k = \sum_{nc} (Y_k - b_k - a_k X_k)^2 \quad (13)$$

donde nc es el número de píxeles en la región de no cambio.

La solución de la ecuación anterior está dada por:

$$a_k = \frac{S_{xk yk}}{S_{xk xk}} \quad (15)$$

$$b_k = \bar{y}_k - a_k \bar{x}_k \quad (16)$$

donde  $S_{xk yk}$  es la covarianza entre ambas imágenes,  $S_{xk xk}$  es la varianza de la imagen X,  $\bar{y}_k$  es la media de la imagen Y, y  $\bar{x}_k$  es la media de la imagen X. Se resuelve la ecuación de mínimos cuadrados para cada banda k.

4. Se normalizar la imagen X usando los coeficientes calculados anteriormente. Es decir, para  $k = 1, 2, 3, 4$ , se calcula

$$\hat{x}_k = a_k x_k + b_k \quad (17)$$

## 2.2 Corrección geométrica

El registro, o correregistro, de imágenes es el proceso mediante el cual una imagen objetivo es transformada geoméricamente al modelo de otra imagen de referencia. Ambas son tomadas de la misma escena pero con características diferentes, ya sea un ángulo distinto, con diferente sensor, o en diferente fecha. El propósito del correregistro es empatar los puntos en común entre la imagen objetivo y los de la imagen de referencia.

La mayoría de los métodos de correregistro consisten en las siguientes cuatro etapas (Zitová, Flusser, 2004):

**Detección de características.** Se detectan manual o automáticamente objetos distintivos o sobresalientes en las imágenes. Esas características se pueden representar en la forma de un descriptor o como un vector de características mediante sus atributos, como su orientación, magnitud de gradiente, escala, etc.

**Empatamiento de características.** Se establece correspondencia entre las características detectadas en la imagen objetivo y las características detectadas en la imagen de referencia. Usualmente se usan medidas de similitud entre sus relaciones espaciales.

**Estimación del modelo de transformación.** Se estiman el tipo y los parámetros de las funciones de mapeo que alinean la imagen objetivo con la de referencia. Los parámetros se obtienen de la correspondencia de características establecida anteriormente.

**Muestreo y transformación de la imagen.** La imagen objetivo es transformada mediante las funciones de mapeo. Valores de imagen en coordenadas no enteras se computan mediante la técnica de interpolación más apropiada.

## 2.2.1 Detección de características

De acuerdo a los requerimientos del sistema propuesto, se requiere obtener correspondencias entre grandes conjuntos de imágenes adquiridas durante misiones de vuelo utilizando plataformas aéreas (tripuladas o no tripuladas, lo que hace impráctico el detectar manualmente objetos distintivos en las imágenes. Los algoritmos de detección automática se clasifican con base a dos enfoques principales: aquellos basados en área e intensidad, en donde mediante distintas estadísticas se pueden establecer correspondencias en regiones de las imágenes; y los algoritmos basados en características, en donde se detectan rasgos sobresalientes como puntos, bordes, esquinas, etc.

Debido a que se desea llevar a cabo el correregistro entre imágenes visibles e infrarrojas, se descartan los algoritmos basados en área, debido a que existirán diferencias en la intensidad de brillo de los píxeles en las diferentes bandas espectrales, como es el caso de la reflectividad de la vegetación en la banda del infrarrojo cercano (Aguilera et al, 2012), por consiguiente en este trabajo se consideran los métodos basados en características. En las siguientes secciones se describirán conceptos y técnicas comúnmente empleadas en la detección de características.

### 2.2.1.1 Concepto de Gradiente

El Gradiente es un operador vector para cualquier función escalar de N dimensiones  $f(x_1, \dots, x_N) = f(x)$ , en donde  $x = [x_1, \dots, x_N]^T$  es un vector variable de N dimensiones (N-D). El gradiente de una función N-D es un vector compuesto de las N derivadas parciales de  $f(x)$ :

$$g(x) = \nabla f(x) = \frac{d}{dx} f(x) = \left[ \frac{\partial f(x)}{\partial x_1}, \dots, \frac{\partial f(x)}{\partial x_N} \right]^T \quad (18)$$

en donde la derivada parcial de una función de varias variables se define como la derivada de la función con respecto a una variable y las otras permanecen constantes.

$$\frac{\partial f}{\partial x_1}(a_1, \dots, a_n) = \lim_{h \rightarrow 0} \frac{f(a_1, \dots, a_i+h, \dots, a_n) - f(a_1, \dots, a_i, \dots, a_n)}{h} \quad (19)$$

Todas las variables  $a_1$  hasta  $a_n$  permanecen fijas, a excepción de  $x_i$ .

Por otra parte, la dirección  $\angle g$  de un vector gradiente  $g$  es la dirección en el espacio N-D en la cual la función  $f(x)$  aumenta más rápido; es el ángulo que guarda con respecto a una línea horizontal. Mientras que la magnitud  $|g|$  del gradiente  $g$  es la tasa de incremento.

En el campo del procesado de imágenes sólo se considera un espacio 2-D:

$$\nabla = \begin{bmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{bmatrix} \quad (20)$$

Cuando el operador de gradiente es aplicado a una función 2-D  $f(x,y)$ , éste produce una función vector:

$$\nabla f(x) = [f_x, f_y]^T \quad (21)$$

en donde  $f_x = \partial f / \partial x$  y  $f_y = \partial f / \partial y$ .

Por lo tanto, la dirección y magnitud de  $g$  son respectivamente:

$$\angle g = \tan^{-1} \left( \frac{f_y}{f_x} \right), \quad \|g\| = \sqrt{f_x^2 + f_y^2} \quad (22)$$

Para imágenes digitales discretas, la derivada en la operación gradiente

$$D_x[f(x)] = \frac{d}{dx} f(x) = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} \quad (23)$$

se convierte en la diferencia

$$D_n[f[n]] = f[n + 1] - f[n], \text{ o } \frac{f[n+1]-f[n-1]}{2} \quad (24)$$

El operador gradiente es un detector efectivo para bordes abruptos, en donde los niveles de brillo de los píxeles cambian bastante rápido en una dirección determinada.

### 2.2.1.2 Laplaciano de Gaussiana

El operador de Laplace es un operador escalar definido como el producto interno entre dos operadores vector gradiente:

$$\Delta = \nabla \cdot \nabla = \nabla^2 = \left[ \frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_N} \right] \begin{bmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_N} \end{bmatrix} = \sum_{n=1}^N \frac{\partial^2}{\partial x_n^2} \quad (25)$$

Cuando N=2, como es el caso de imágenes:

$$\Delta = \nabla \cdot \nabla = \nabla^2 = \begin{bmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \end{bmatrix} \begin{bmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{bmatrix} = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \quad (26)$$

y la derivada de segundo orden se convierte en una diferencia de segundo orden:

$$(f[n + 1] - f[n]) - (f[n] - f[n - 1]) = f[n + 1] - 2f[n] + f[n - 1] \quad (27)$$

La operación provoca el mismo efecto que un filtro pasa altas, pero produce un mejor resultado al detectar bordes abruptos.

Si en el vecindario de un píxel dado existen valores mayores y menores que cero, entonces el píxel se considera un cruce por cero, y si la diferencia entre el máximo y el mínimo es mayor que un umbral dado, se considera al píxel como la ubicación de un borde. Esto se realiza considerando la posibilidad de existir ruido aleatorio.

Aun así, el operador de Laplace es sensible al ruido, por lo que es deseable suavizar la imagen para tratar de eliminarlo. Esto se puede realizar mediante la convolución de la imagen con un filtro Gaussiano de anchura sigma:

$$G_{\sigma}(x, y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \quad (28)$$

en donde sigma es la desviación estándar de la Gaussiana y representa el ancho de esta.

Aplicando este filtro para suprimir el ruido, antes de aplicar el operador de Laplace para detectar bordes, se tiene:

$$\Delta[G_{\sigma}(x, y) * f(x, y)] = [\Delta G_{\sigma}(x, y)] * f(x, y) = LoG * f(x, y) \quad (29)$$

el cual, gracias a la propiedad asociativa de la convolución, se puede implementar como un operador o núcleo de filtro independiente, al cual se le conoce como Laplaciano de Gaussiana:

$$LoG \triangleq \Delta G_{\sigma}(x, y) = \frac{\partial^2}{\partial x^2} G_{\sigma}(x, y) + \frac{\partial^2}{\partial y^2} G_{\sigma}(x, y) = \frac{x^2+y^2-2\sigma^2}{\sigma^4} e^{-(x^2+y^2)/2\sigma^2} \quad (30)$$

El operador LoG para dos dimensiones se puede aproximar como un núcleo de filtro de convolución de 5x5.

Se pueden obtener núcleos de distintos tamaños aproximando la expresión continua del operador LoG. El proceso de detección de bordes es igual al del operador de Laplace,

se aplica el operador LoG, se obtienen los cruces por cero, y se descartan aquellos cruces cuya diferencia entre su valor maximal y minimal sean menores que un umbral dado.

### 2.2.1.3 Diferencia de Gaussianas

Este procedimiento consta de dos etapas. La imagen se suaviza con un filtro Gaussiano de amplitud 1:

$$g_1(x, y) = G_{\sigma_1}(x, y) * f(x, y) \quad (31)$$

Posteriormente, con otro filtro sigma2 de amplitud distinta se obtiene otra imagen suavizada:

$$g_2(x, y) = G_{\sigma_2}(x, y) * f(x, y) \quad (32)$$

De igual manera que en el Laplaciano de Gaussiana, se puede separar el operador aplicado a la imagen en un núcleo de filtro independiente. A este filtro se le denomina Diferencia de Gaussianas:

$$DoG \triangleq G_{\sigma_1} - G_{\sigma_2} = \frac{1}{\sqrt{2\pi}} \left( \frac{1}{\sigma_1} e^{-\frac{x^2+y^2}{2\sigma_1^2}} - \frac{1}{\sigma_2} e^{-\frac{x^2+y^2}{2\sigma_2^2}} \right) \quad (33)$$

En realidad el operador DoG es un filtro pasa-bandas, debido a que implica la diferencia entre dos imágenes a las que se les ha aplicado un filtro pasa-bajos diferente. Por lo tanto este filtro remueve los componentes de alta frecuencia relacionados con el ruido, así como algunos componentes de baja frecuencia que representan las áreas homogéneas en la imagen. Aquellos componentes de frecuencia intermedia que pasa por la banda, se supone están asociados con los bordes en la imagen.

La curva obtenida por el operador DoG es bastante similar a la de LoG, por lo que se puede decir que es una aproximación a éste, cuyo beneficio es en términos computacionales, ya que se realizan menos operaciones. El procedimiento para detectar bordes es idéntico al de LoG descrito anteriormente.

#### 2.2.1.4 Determinante de la Hessiana

El determinante de la matriz Hessiana puede ser una herramienta útil para detectar puntos de interés. La matriz Hessiana es una matriz cuadrada que incluye las derivadas parciales de segundo orden de una función.

Dado un punto  $X = (x,y)$  en una imagen  $I$ , la matriz Hessiana  $H(X,)$  en el punto  $X$  a una escala se define como sigue:

$$\mathcal{H}(x, \sigma) = \begin{bmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma) & L_{yy}(x, \sigma) \end{bmatrix} \quad (34)$$

donde  $L_{xx}(x, \sigma)$  es la convolución de la derivada de segundo orden Gaussiana  $d^2/dx^2 G()$  con la imagen  $I$  en el punto  $X$ . Siendo de igual manera para  $L_{xy}(x, \sigma)$  y  $L_{yy}(x, \sigma)$ .

Mediante el determinante de la matriz Hessiana, el cual se define como

$$\det H L (x, y; t) = t^2(L_{xx} L_{yy} - L_{xy}^2) \quad (35)$$

en donde  $t$  es la escala, se pueden detectar regiones en un espacio de escalas comparando el valor del determinante obtenido en el punto  $X$  con los determinantes de los puntos vecinos en espacio y escala

$$(\hat{x}, \hat{y}; \hat{t}) = \operatorname{argmax}_{(x,y;t)} (\det HL (x, y; t) ) \quad (36)$$

Si el determinante es mayor que un umbral dado y que todos sus vecinos, entonces se considera al punto  $X$  como un punto de interés.

Debido a que el determinante de la Hessiana es maximizado, en el proceso se descartan las estructuras alargadas que tienen derivadas de segundo orden pequeñas, o sea, cambios en la señal en una sola dirección.

### 2.2.1.5 Transformada SIFT

La Transformada SIFT (Scale-Invariant Feature Transform) es un algoritmo que combina detección, descripción y empatación de características. Es un algoritmo que es invariante a rotación, escala y parcialmente a iluminación, por lo que se considera que funciona de manera robusta (Lowe,2004). Durante varios años fue considerado el estado del arte, al obtener los mejores resultados en casos generales.

Durante el proceso de detección de SIFT, en donde se obtienen aquellos puntos que se consideran de interés, se utiliza la Diferencia de Gaussianas, el cual es un filtro central que se aplica sobre una pirámide a diferentes escalas. Esto para detectar respuestas extremas (maximales o minimales) en el vecindario en que se le aplica el filtro y a distintas escalas.

Para asegurar la invariancia a rotación se le asigna una orientación a cada punto de interés. Para lograr esto se selecciona la imagen suavizada más cercana a la escala del punto, para que todo el cálculo sea invariante a escala también. Para cada muestra de imagen  $L(x,y)$ , a una escala dada, la magnitud  $m(x,y)$  y la orientación  $\theta(x,y)$  del Gradiente se calculan usando las diferencias de píxeles:

$$m(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2} \quad (37)$$

$$\theta(x,y) = \tan^{-1}((L(x,y+1) - L(x,y-1))/(L(x+1,y) - L(x-1,y))) \quad (38)$$

Posteriormente se crean histogramas de orientación con base en la orientación de los puntos de muestra en la región alrededor de un punto de interés. El histograma de orientación cuenta con 36 cubetas (bins) para cubrir cada 10 grados de los 360 grados

del rango de orientación. Cada muestra agregada al histograma es pesada por su magnitud y por una ventana circular Gaussiana con un valor  $\sigma = 1.5$  que es la escala del punto de interés a la que se concluyó empíricamente que era la que obtenía mejores resultados. Los picos en los histogramas de orientación representarán la orientación dominante del gradiente local.

Después de detectar los puntos de interés, estos se describen mediante los histogramas de gradiente obtenidos de los puntos alrededor del punto de interés y a distintas escalas. Esto hace que sea más descriptivo y más robusto a los cambios de iluminación que a utilizar valores crudos de los píxeles, o histogramas de color, etc. A este proceso se le denomina “descriptor de punto de interés”, y es la principal aplicación de este algoritmo.

El descriptor es creado a partir de un vecindario de  $16 \times 16$  alrededor del punto de interés, el cual se divide en 16 sub-bloques de  $4 \times 4$ . Por cada sub-bloque se crea un histograma de orientación de 8 cubetas. Posteriormente se genera un vector compuesto de cada uno de los sub-bloques, por lo que la dimensión del descriptor será de 128 (16 sub-bloques \* 8 cubetas de orientación).

Para empatar los descriptores con los obtenidos de otra imagen, mediante una métrica de distancia basada en los descriptores, se compara la relación entre los dos más cercanos y si tal relación es baja significa que el primero es mucho más cercano que el segundo, por lo que se considera que es un buen candidato para el empatamiento. En caso contrario ( $>.8$ , según los autores), esto significa que sus distancias son parecidas, por lo que se rechaza el empatamiento, debido a que puede ocurrir una falsa correspondencia por causa del ruido.

### 2.2.1.6 Algoritmo SURF

El algoritmo SURF (Speeded-Up Robust Features) es similar a SIFT en cuanto a que propone un conjunto de técnicas para detección, descripción y empataamiento, el cual funciona para casos en general (Bay,2008).

Para el proceso de detección, el algoritmo SURF utiliza el determinante de una matriz Hessiana, el cual no produce características tan estables como en el caso del operador LoG (Tuytelaars y Mikolajczyk, 2008). Sin embargo esto se compensa con la velocidad del algoritmo, mediante el uso de un filtro de caja (box filter) para aproximar la matriz Hessiana. Esto logra una aceleración en el tiempo de ejecución, como su nombre lo indica.

El valor en cualquier punto  $(x,y)$  en la imagen entera  $I$  es la suma de todos los píxeles por arriba y hacia la izquierda de  $(x,y)$ :

$$I(x, y) = \sum_{\substack{x' \leq x \\ y' \leq y}} i(x', y') \quad (39)$$

Una imagen entera en  $(x,y)$  se puede calcular eficientemente, en términos de tiempo de ejecución, en una sola pasada sobre la imagen usando la siguiente propiedad:

$$I(x, y) = i(x, y) + I(x - 1, y) + I(x, y - 1) - I(x - 1, y - 1) \quad (40)$$

Una vez calculada la imagen, la tarea de evaluar cualquier rectángulo se puede realizar en tiempo constante con sólo cuatro accesos a referencias de arreglo. Esto ayuda en la detección Hessiana al no tener que calcular el valor en cada punto del paso de muestra para obtener la traza y el determinante de cada capa, lo que agiliza el proceso.

A diferencia de SIFT, no se crea un espacio de escalas o una pirámide en sí, sino que se calcula cada escala cambiando el tamaño del filtro, lo cual hace posible su implementación paralela.

Después de obtener los determinantes se compara su vecindario 3x3x3 (en espacio y escala) para tratar de encontrar los máximos, los cuales serán los puntos de interés.

Para describir los puntos de interés obtenidos en el primer paso, se asigna una orientación a la región circular alrededor del punto de interés. Después se construye una región cuadrada de 20x20 alineada a la orientación asignada y se divide en subregiones de 4x4, a las cuales se les aplican filtros Wavelet de Haar en dirección vertical y horizontal, cuyas respuestas se toman para formar el siguiente vector, el cual es el descriptor para una subregión dada

$$v = (\sum d_x, \sum d_y, \sum |d_x|, \sum |d_y| ) \quad (41)$$

donde  $dx$  y  $dy$  son las respuestas horizontales y verticales respectivamente del Wavelet de Haar. Para obtener el descriptor del punto de interés se agregan los vectores descriptores de las 16 subregiones en un sólo vector de dimensión 64, a diferencia del descriptor SIFT que cuenta con dimensión 128, lo que hace más lento el proceso de empataamiento.

Habiéndose obtenido este descriptor se pueden obtener correspondencias.

### 2.2.2 Empataamiento de características

Una vez que se cuenta con suficientes puntos de interés o características para ambas imágenes (la de referencia y la objetivo) es necesario establecer correspondencias entre los puntos de una imagen y los de la otra. A partir de estos puntos se obtendrán los puntos de control necesarios para el correregistro de las imágenes.

### **2.2.2.1 Relación de distancia entre los vecinos más cercanos**

La búsqueda del vecino más cercano se define de la siguiente manera: dado un conjunto  $S$  de puntos en un espacio  $M$  y un punto de consulta  $q \in M$ , encontrar el punto más cercano a  $q$  en  $S$ .

Para el caso de descriptores invariantes como SIFT y SURF, el vecino más cercano se considera como aquel punto de interés con la distancia Euclidiana mínima.

Durante la búsqueda se consideran los dos vecinos más cercanos y se rechazan todas las correspondencias en donde la relación de distancia entre los vecinos más cercanos es mayor a 0.8, lo cual elimina el 90% de correspondencias falsas, mientras que sólo se descarta menos del 5% de correspondencias correctas.

La implementación de este algoritmo se puede acelerar mediante una aproximación con el algoritmo de Best-Bin-First, el cual hace una búsqueda de orden modificada, tal que las cubetas en el espacio de características sean buscadas en el orden de la distancia más cercana a la ubicación de la búsqueda.

### **2.2.3 Estimación del modelo de transformación**

Una vez que se cuenta con suficientes puntos de control, estos son utilizados para estimar modelos de transformación que permiten superponer la imagen objetivo sobre la imagen de referencia, de manera que los puntos de control correspondan. El tipo de función de mapeo seleccionada debe coincidir con el tipo de deformación geométrica que se supone tiene la imagen objetivo, así como con el método de adquisición de la imagen y con la precisión requerida para el correregistro (Gostasby, 2011).

Existen dos categorías de funciones de mapeo: los modelos globales que usan todos los puntos de control para estimar una función de mapeo para toda la imagen; y los modelos locales que consideran la imagen como una composición de parches donde los parámetros de la función dependen de su ubicación en la imagen.

### 2.2.3.1 Modelos de mapeo globales

La *transformación de similitud* (rígida) es el modelo más simple, en donde sólo se consideran las distorsiones por rotación, traslación y escalamiento.

$$u = s(x \cos \alpha - y \sin \alpha) + t_x \quad (42)$$

$$v = s(x \sin \alpha + y \cos \alpha) + t_y \quad (43)$$

donde  $s$  es un factor de escalamiento,  $\alpha$  es el ángulo de rotación y  $t_x$  y  $t_y$  son la traslación en  $x$  y  $y$  respectivamente. Este modelo mantiene los ángulos y curvaturas.

La *transformación afín*, llamada distorsión de cizalla (*shear*), además de las distorsiones mencionadas en la transformada de similitud, es un modelo más general que puede mapear un paralelogramo a un cuadrado. Esta transformación mantiene las líneas rectas y el paralelismo entre estas.

$$u = a_0 + a_1x + a_2y \quad (44)$$

$$v = b_0 + b_1x + b_2y \quad (45)$$

donde  $a_0$  y  $b_0$  son factores de traslación horizontal y vertical respectivamente, y los coeficientes  $a_1, a_2, b_1, b_2$  definen el tipo de transformación, ya sea rotación, reflexión o escalamiento.

Se puede emplear para el correregistro de imágenes adquiridas desde distintos puntos de vista, cuando la distancia de la cámara a la escena es grande en comparación con el tamaño del área fotografiada, o cuando la escena es plana y la distorsión geométrica presente no cuenta con factores locales.

Si la condición de la distancia entre la cámara y la escena no se satisface, entonces se debe usar el *modelo de proyección perspectiva*

$$u = \frac{a_0 + a_1x + a_2y}{1 + c_1x + c_2y} \quad (46)$$

$$v = \frac{b_0 + b_1x + b_2y}{1 + c_1x + c_2y} \quad (47)$$

en donde el punto extra (o los dos coeficientes)  $c_1$  y  $c_2$  permiten especificar la deformación por las distancias relativas entre los pares de puntos. Este modelo permite mapear un cuadrángulo general a un cuadrado preservando las líneas rectas.

Cuando las imágenes cuentan con distorsiones no lineales es necesario utilizar polinomiales de segundo o tercer orden, esto es, transformaciones no lineales. Por ejemplo, un modelo polinomial de segundo orden será:

$$x' = k_{x1} + k_{x2}x + k_{x3}y + k_{x4}x^2 + k_{x5}xy + k_{x6}y^2 \quad (48)$$

$$y' = k_{y1} + k_{y2}x + k_{y3}y + k_{y4}x^2 + k_{y5}xy + k_{y6}y^2 \quad (49)$$

donde  $k_n$  son los coeficientes correspondientes a los 6 puntos de control, mínimo necesario para el modelo.

Debido a que el número de puntos de control es usualmente mayor que el mínimo requerido para determinar la función de mapeo, entonces se deben seleccionar aquellos puntos que minimizan la suma de los errores cuadrados en los puntos de control, mediante un ajuste por Mínimos Cuadrados.

### 2.2.3.2 Modelos de mapeo locales

Un mapeo polinomial global no puede manejar de manera adecuada las imágenes distorsionadas localmente. Para el correregistro de áreas locales se debe considerar la información disponible acerca de la distorsión geométrica local. Algunos métodos locales son: el mapeo lineal por segmentos, el mapeado cúbico por segmentos y la transformada de promedio local pesado.

Esta última se define dadas dos matrices  $m \times 2$ , en donde  $m$  es la cantidad de puntos de control empataados en la imagen de referencia y en la objetivo. La transformada crea un mapeo, al inferir un polinomio en cada punto de control usando sus puntos de control vecinos. El mapeo en cualquier ubicación dependerá de un promedio pesado de esos polinomiales. Los  $n$  puntos más cercanos son usados para inferir una transformación polinomial de segundo grado para cada uno de los puntos de control. El valor mínimo de  $n$  es 6 para un polinomio de segundo grado normal, pero son recomendados más puntos para obtener buenos resultados.

#### **2.2.4 Muestreo y transformación de la imagen**

Después de obtener un modelo de transformación para el correregistro, es necesario ahora aplicarlo. Existen dos maneras de hacerlo; en la primera se considera cada punto de control en la imagen objetivo y se le aplica el modelo de transformación, a esto se le llama una transformación directa (forward transform). El problema con este procedimiento es que puede ocurrir que varios puntos queden fuera del tamaño de la imagen, lo que ocasiona huecos que necesitan ser interpolados en la imagen, o bien, que dos puntos se traslapen al caer en la misma ubicación.

Por otro lado, el segundo enfoque se denomina transformada inversa (backward transform) y ofrece mejores resultados. En este método se considera cada punto de la imagen de referencia y se le aplica el inverso del modelo de transformación, para obtener la ubicación correspondiente en la imagen objetivo. Como la interpolación se realiza sobre la imagen objetivo, no ocurren traslapos o huecos.

La interpolación consiste en cualquier técnica utilizada para construir puntos de datos dentro del rango de un conjunto discreto de puntos conocidos. En imágenes, la interpolación se realiza mediante la convolución de una imagen con un núcleo de interpolación.

Los métodos de interpolación más comúnmente usados son los siguientes: vecino más cercano (nearest neighbor), interpolación bilineal y bicúbica, splines cuadráticos, B-splines de mayor orden, Gaussianas, y funciones sinc.

En el muestreo de imágenes, cuando la velocidad de cálculo no es un factor importante, la interpolación bicúbica se elige sobre vecino más cercano e interpolación bilineal, ya que produce una superficie más suave (imagen más nítida).

En el caso de la interpolación bilineal, para un punto (x,y) su valor interpolado se obtiene al promediar el valor pesado de sus 4 puntos más cercanos en un vecindario de 2x2. La función de peso para cada uno de los 4 puntos dependerá de su distancia al punto (x,y).

La interpolación bicúbica considera a los 16 vecinos más cercanos al punto (x,y), basándose en la fórmula para una interpolación cúbica

$$f(p_0, p_1, p_2, p_3, x) = \left(-\frac{1}{2}p_0 + \frac{3}{2}p_1 - \frac{3}{2}p_2 + \frac{1}{2}p_3\right)x^3 + \left(p_0 - \frac{5}{2}p_1 + 2p_2 - \frac{1}{2}p_3\right)x^2 + \left(-\frac{1}{2}p_0 + \frac{1}{2}p_2\right)x + p_1 \quad (50)$$

en este caso se utiliza un polinomio de tercer grado para interpolar en un intervalo [0,1], considerando 4 puntos para resolverlo.

En el caso de una interpolación bicúbica primero se interpola cúbicamente en una dirección, por ejemplo verticalmente, para el primer renglón de 4 elementos del vecindario de 4x4, y después para los siguientes tres. Cada valor obtenido se utiliza para interpolar en la otra dirección, horizontalmente, como se muestra en la siguiente fórmula:

$$g(x, y) = f(f(p_{00}, p_{01}, p_{02}, p_{03}, y), f(p_{10}, p_{11}, p_{12}, p_{13}, y), f(p_{20}, p_{21}, p_{22}, p_{23}, y), f(p_{30}, p_{31}, p_{32}, y), x) \quad \dots(51)$$

Aunque la interpolación bicúbica produce mejores resultados, generalmente tiene un costo mayor de procesamiento, por lo que se debe considerar si se justifica ese compromiso entre calidad y tiempo.

## Capítulo 3. Materiales y Metodología

---

### 3.1 Fotografía aérea

En el Laboratorio de Geomática del CICESE se cuenta con acceso a un avión ultraligero tipo ala-delta, equipado con un arnés sujetador y sistema para el disparo automático de cámaras, así como con un dron multi-rotor tipo hexacóptero equipado con sistema de vuelo autónomo. Se cuenta además con una cámara Canon EOS 5D modelo Mark II y una cámara Canon EOS 500D modificada para captar imágenes en el infrarrojo cercano (830 – 1200 nm). Ambas cámaras funcionan de manera simultánea para formar imágenes multispectrales RGB-NIR con un ancho de banda espectral que va desde el visible hasta el infrarrojo cercano (400 – 1200 nm).

Se eligieron dos sitios de prueba con el objeto de hacer vuelos con el hexacóptero para adquirir fotografía aérea en el visible e infrarrojo cercano. El primer sitio corresponde al nodo de retorno vial ubicado frente al CICESE, a este sitio se le denominó “*Retorno*”. El segundo sitio corresponde a la parte superior del campus de CICESE, en la zona del edificio de Telemática, a este sitio se le denominó “*Telemática*”.

Se efectuaron un par de vuelos sobre cada sitio de prueba, un vuelo para adquirir la fotografía en el visible y el otro para tomar la fotografía infrarroja. Ambos vuelos se efectuaron a una altura de 100 m sobre el sitio de despegue del multi-rotor y a una velocidad de crucero de 5 m/s. Se tomaron conjuntos de fotografías de 27 y 32 tomas sobre los sitios Retorno y Telemática respectivamente. Cada fotografía tuvo aproximadamente un cubrimiento de 86 X 65 m en el terreno y una resolución de 2 cm por pixel. En los capítulos 4 y 5 se muestra el uso de estos conjuntos de fotografías para probar el funcionamiento del sistema que se implementó en este trabajo.

### 3.2 Desarrollo del sistema

Se eligió Python como el lenguaje de programación a utilizar en el desarrollo del sistema debido a su compatibilidad multiplataforma, su ágil proceso de desarrollo y a su

creciente aceptación en el ámbito del cómputo científico, lo cual le proporciona bibliotecas de acceso libre que son útiles para el manejo de imágenes y datos en grandes cantidades.

Se utilizó Numpy, que es una librería de código abierto, para el manejo de arreglos de N dimensiones, álgebra lineal y cómputo científico en general, para facilitar el manejo de imágenes en el desarrollo de los algoritmos de corrección.

Se aprovecharon los algoritmos de visión por computadora proporcionados por la librería OpenCV para el análisis de correspondencia entre imágenes, necesario en el proceso de corrección geométrica.

Se decidió usar PyQt como entorno gráfico, al ser multiplataforma y tratarse de un binding sobre el código de Qt, que es popular como interfaz gráfica en distintos lenguajes de programación, y proporcionar una arquitectura modelo-vista, que permite mantener separadas las distintas competencias del código.

Se usó el paquete de desarrollo de software científico Python(x,y) que proporciona las librerías anteriormente mencionadas y cuenta con un entorno de desarrollo integrado llamado Spyder.

El desarrollo del software se basó en una arquitectura Modelo-Vista, ya que es la que utiliza Qt. Basándose en la famosa arquitectura Modelo-Vista-Controlador se optó por integrar la vista con el controlador, debido a que el sistema de manejo de eventos de Qt, denominado señales y ranuras (signals and slots), es el más apropiado para el desarrollo de aplicaciones de escritorio.

## Capítulo 4. Sistema de percepción remota

---

En este capítulo se describen los requerimientos del sistema y las decisiones de diseño que se siguieron a partir del flujo de procesamiento post-vuelo necesario para el uso de las imágenes en aplicaciones de percepción remota, además se describe el funcionamiento de la implementación del sistema.

### 4.1 Descripción del sistema

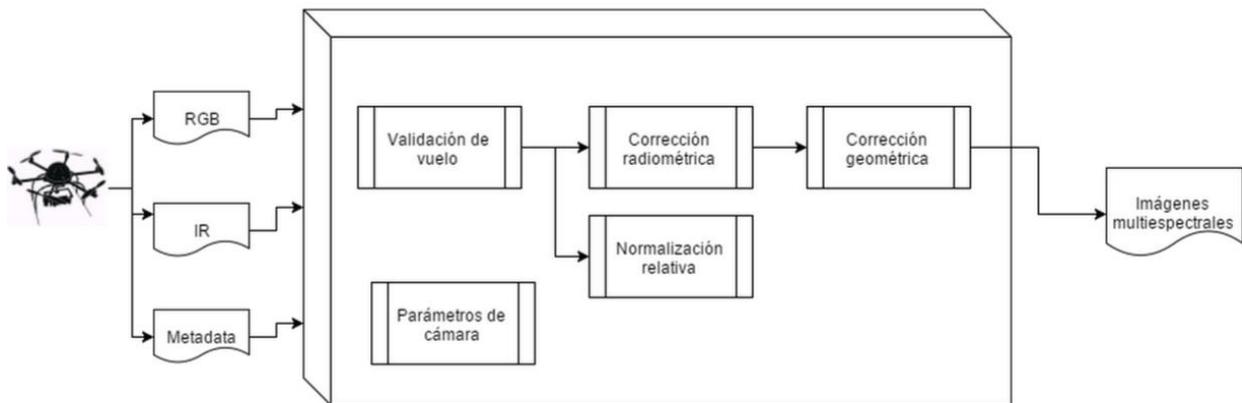
Se supone que se cuenta con dos juegos de imágenes, uno tomado en el visible y el otro en el infrarrojo. La logística necesaria para la adquisición de las imágenes y el diseño y manejo de la misión aérea se consideran trabajo previo ya realizado (Ramírez, 2013).

Como estos juegos de imágenes se graban en memorias SD, el sistema debe de ser capaz de seleccionar en el sistema de archivos de cualquier sistema operativo la carpeta que contenga las imágenes objetivo y la carpeta de las imágenes de referencia.

En el módulo de análisis post-vuelo (previo a las correcciones) el sistema permite validar que la misión de vuelo se haya realizado exitosamente, permitiendo observar el mapa índice de vuelo de las imágenes tomadas, lo que permite identificar si hubo huecos a causa de errores en la adquisición algunas fotos o mediante la comparación de ambos grupos de imágenes entre sí. El sistema tiene la opción de renombrar y reenumerar cada una de las imágenes en cada carpeta, creando copias de ellas sobre la cual se va a trabajar, lo cual facilita su identificación y su posterior procesamiento. Además de ofrecer la capacidad de visualizar, agregar y editar los campos Exif (Exchangeable image file format) el cual contiene meta-información relacionada con la imagen; ya sea información sobre el proceso de captura (distancia focal, tiempo de exposición, valor de apertura, ubicación, etc.) o sobre la misma imagen (dimensiones, formato, etc.). Tales acciones se pueden realizar individualmente por imagen (como en el caso de la ubicación) o sobre un grupo de imágenes (como las características del dispositivo de adquisición). En el caso de este sistema, sólo ciertos datos del proceso

de captura son útiles, en particular la información de ubicación, la cual se almacena en Exif como información de GPS, es decir, la latitud, longitud y altitud registradas en el momento de la captura de la imagen.

En la figura 1 se muestran los distintos componentes o módulos que componen el sistema.



**Figura 1. Diagrama general de la arquitectura del sistema**

#### 4.1.1 Corrección radiométrica

Se implementaron dos distintos tipos de corrección radiométrica, los cuales son procesos independientes el uno del otro, y se pueden aplicar en cualquier orden.

##### 4.1.1.1 Normalización radiométrica relativa

Se implementó el algoritmo de manera semi-interactiva con el objeto de poder trabajar con imágenes con distintos tipos de cubrimiento de terreno, esto en lugar de depender de que las imágenes incluyan superficie con cubierta de agua y suelo, como en el caso de la regresión automática con diagramas de dispersión (Elvidge, 1995). Por lo tanto es necesario contar con una interfaz gráfica interactiva, para que el usuario seleccione los puntos de control que se considerarán como el extremo superior y el extremo inferior de

reflectividad. Para apoyar la selección de los puntos extremos, debajo de cada imagen que se despliega el diagrama de dispersión del tiempo correspondiente, en donde al seleccionar un punto sobre una de las imágenes, mostrará la ubicación en el diagrama del valor radiométrico de tal punto, con el objetivo de encontrar los puntos extremos que se encuentren aproximadamente en el centro de sus cúmulos correspondientes.

El valor de HPW (Half Perpendicular Width) podrá ser modificado por el usuario y contará con un valor por omisión de 10, correspondiendo al reportado como el que daba mejores resultados en las imágenes de prueba en Elvidge, et al. (1995).

#### **4.1.1.2 Viñeteo**

Al acceder al módulo de viñeteo se requerirá que se seleccione la carpeta con el grupo de imágenes a procesar, dando la opción al usuario de sólo generar el filtro y guardarlo como un archivo de texto (.txt) con el cual se podrá corregir cada imagen individualmente, o bien, generar el filtro en memoria y aplicarlo a cada una de las imágenes en la carpeta. Esta operación puede resultar costosa en tiempo, por lo que se deberá ejecutar la corrección en otro hilo diferente al de la interfaz gráfica y se desplegará una barra de progreso en donde se podrá cancelar el proceso. Además se creará otra carpeta con las imágenes corregidas, para preservar las imágenes originales.

Durante el flujo de trabajo semiautomático se corregirá el efecto de viñeteo en ambas carpetas, la de imagen de referencia y la objetivo, creando un filtro para cada una de ellas, sin necesidad de la intervención del usuario. Esto deberá realizarse antes del corregistro de las imágenes.

#### **4.1.2 Corrección geométrica**

Para corregir geoméricamente las imágenes es necesario corregir las imágenes objetivo con las de referencia una a una. Para esto es necesario que exista una correspondencia para cada imagen objetivo, sin embargo debido a factores externos es posible que durante la misión de vuelo no se logren capturar imágenes en algunos focos, o bien, que la imagen objetivo esté demasiado desfasada con respecto a la imagen de referencia, por lo que no sea de utilidad. En esos casos es necesario que el sistema permita al usuario revisar ambos grupos de imágenes (referencia y objetivo), descartar aquellas que no tengan imagen correspondiente y así seleccionar aquellos pares de imágenes que pasarán posteriormente al proceso de corrección. Este proceso será automático y tendrá como resultado una carpeta nueva que a su vez contendrá una carpeta por cada par de imágenes, en donde se guardará un archivo en formato Envi con las 6 bandas y con entrelazado BSQ, además de otro archivo de texto (.txt) con los valores de la homografía aplicada a la imagen objetivo, así como el resultado de tal operación en una imagen JPEG, es decir, la imagen objetivo con la geometría de la imagen de referencia.

#### **4.1.3. Validación post-vuelo**

El proceso de validación post-vuelo es necesario debido a diversos factores involucrados en la adquisición de imágenes durante el flujo de una misión aérea (e.g. la precisión del sistema de posicionamiento, el tiempo de captura de la cámara, las condiciones de viento, etc.). Debido a lo anterior puede ocurrir que una misión no se haya realizado exitosamente, como es el caso en que falten una o más fotografías en alguno de los conjuntos de imágenes, por lo que es necesario el poder validar los datos obtenidos antes de ingresarlos al sistema de corrección y corrección. Una manera de lograr esto es mediante la generación de un mapa índice de vuelo que permita identificar la existencia de huecos en los dos conjuntos de fotografías mediante la comparación de ambos grupos de imágenes, permitiendo descartar aquellas que no tengan una imagen correspondiente.

## 4.2 Implementación del sistema

Como se mencionó en el Capítulo 3, el sistema fue implementado utilizando el lenguaje de programación Python bajo una arquitectura Modelo-Vista. El sistema cuenta con distintos módulos independientes a los cuales se puede acceder desde el menú inicial (Figura 2). A continuación se describe el funcionamiento para cada uno de los módulos.

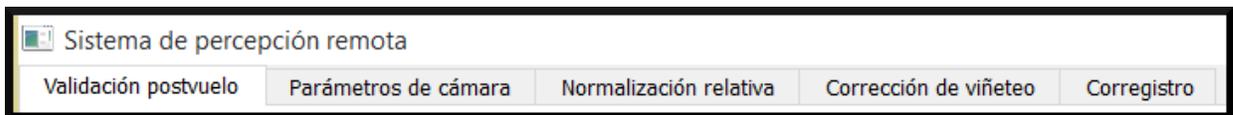
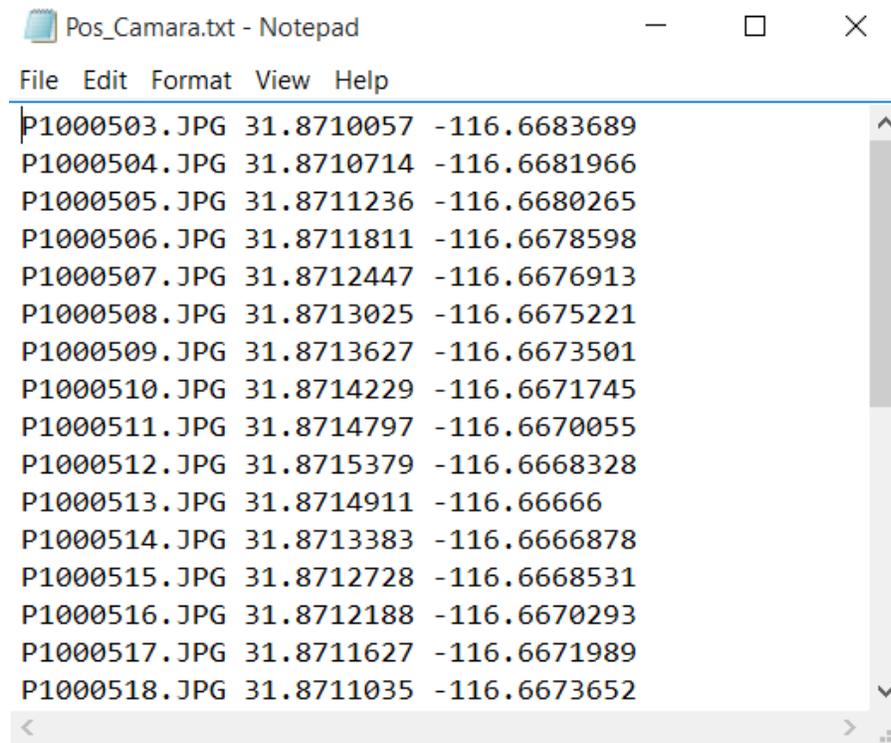


Figura 2. Menú de opciones del sistema

### 4.2.1 Validación post-vuelo

La validación post-vuelo consta de dos etapas, en la primera se hace una evaluación visual mediante un mapa índice de vuelo para identificar huecos en el cubrimiento, y en la segunda etapa se comparan y seleccionan pares entre ambos conjuntos de imágenes para confirmar su correspondencia.

Para generar un mapa índice de vuelo es necesario seleccionar el folder en donde se encuentre el conjunto de imágenes obtenido en una misión aérea, así como contar con un archivo de texto (.txt) en donde se haga una relación entre cada imagen con las coordenadas del fotocentro donde fue tomada en un formato separado por espacios y con un fin de línea entre cada imagen (Fig. 3)



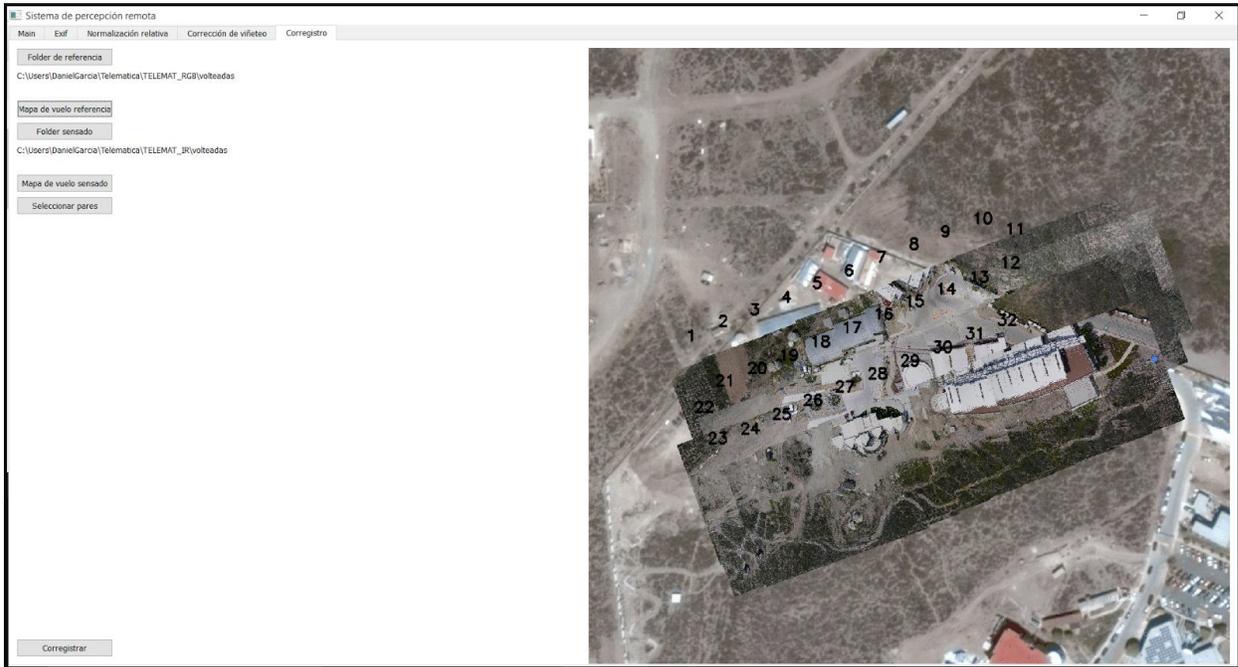
Pos\_Camara.txt - Notepad

File Edit Format View Help

```
P1000503.JPG 31.8710057 -116.6683689
P1000504.JPG 31.8710714 -116.6681966
P1000505.JPG 31.8711236 -116.6680265
P1000506.JPG 31.8711811 -116.6678598
P1000507.JPG 31.8712447 -116.6676913
P1000508.JPG 31.8713025 -116.6675221
P1000509.JPG 31.8713627 -116.6673501
P1000510.JPG 31.8714229 -116.6671745
P1000511.JPG 31.8714797 -116.6670055
P1000512.JPG 31.8715379 -116.6668328
P1000513.JPG 31.8714911 -116.66666
P1000514.JPG 31.8713383 -116.6666878
P1000515.JPG 31.8712728 -116.6668531
P1000516.JPG 31.8712188 -116.6670293
P1000517.JPG 31.8711627 -116.6671989
P1000518.JPG 31.8711035 -116.6673652
```

**Figura 3. Archivo de texto con las coordenadas de la misión aérea**

Después de seleccionar el archivo de texto con las coordenadas de la misión, el sistema descarga un mapa estático con la cobertura correspondiente y despliega las imágenes del folder seleccionado sobre la coordenada que indica el archivo de texto, creando así una visualización general de la misión, lo que ayuda en el proceso de validación a identificar huecos en caso de imágenes no tomadas (Figura 4).



**Figura 4. Mapa de vuelo generado**

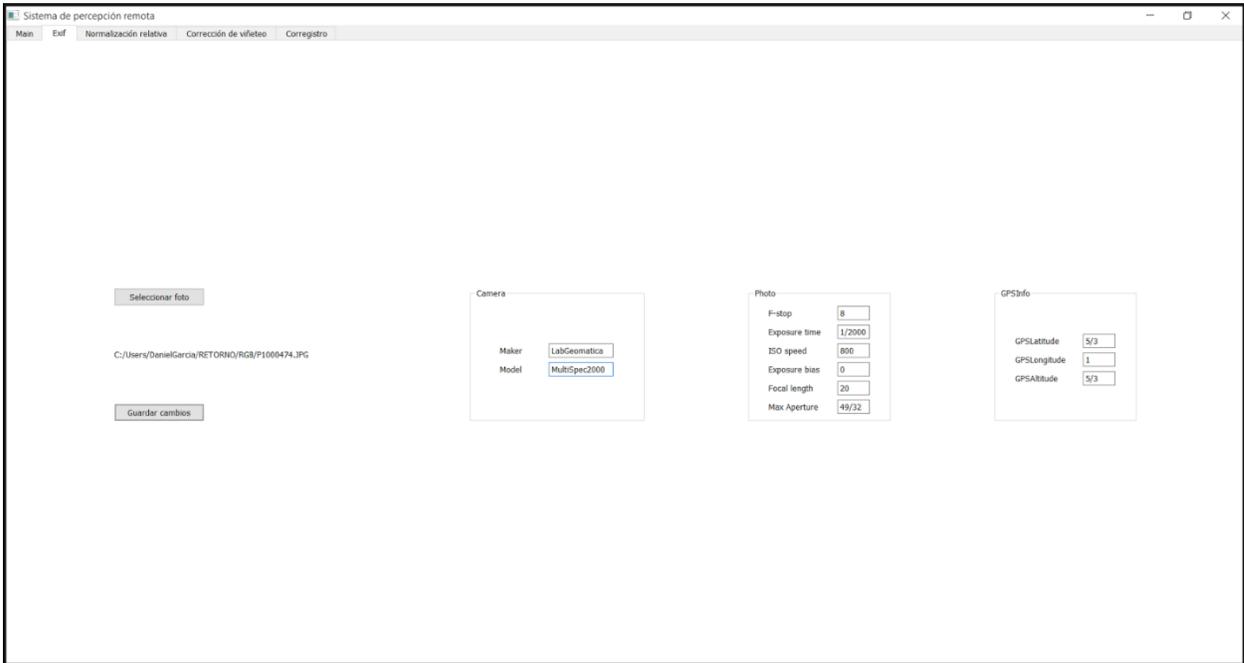
Para poder comparar y seleccionar pares en dos grupos de imágenes de la misma misión es necesario seleccionar la carpeta en donde se encuentren las imágenes de referencia y la carpeta en donde se encuentren las imágenes objetivo, y hacer click en el botón “Seleccionar pares”. Esto muestra las primeras imágenes en cada carpeta (Fig. 5), permitiendo navegar entre ellas presionando los botones de siguiente o anterior, una etiqueta en la parte superior izquierda de cada imagen muestra cuál número de fotografía se está mostrando actualmente y de cuantas es el total del conjunto de imágenes. Para facilitar la comparación entre los pares, se resalta el centro de cada imagen mediante una X que se sobrepone a ella, en donde la intersección de las líneas representa el centro de la imagen. Después de hacer la comparación visual el usuario puede descartar aquellas imágenes que no tengan correspondencia, con el objetivo de dejar ambos grupos con la misma cantidad de imágenes. Para descartar una imagen, ésta se debe estar mostrando y el usuario debe de hacer click en el botón ‘Descartar imagen’ y confirmar que lo desea hacer.



Figura 5. Comparación entre grupos de imágenes

#### 4.2.2 Parámetros de cámara

En el módulo de parámetros de cámara se pueden cambiar los valores ya sea de una imagen o de un grupo de imágenes seleccionando la carpeta que las contenga. Los parámetros que se pueden modificar o agregar se separan en tres categorías: *Camera*, *Photo* y *GPSInfo*. La categoría *Camera* cuenta con los campos de Fabricante y Modelo de la cámara, los que se pueden modificar escribiendo cualquier cadena de texto. En la categoría *Photo* se muestran los parámetros de configuración internos de la cámara solamente para su consulta y son los siguientes: F-stop, Exposure time, ISO speed, Focal length y Max Aperture. En la categoría *GPSInfo* se le agrega o modifica a una imagen individual sus datos geográficos, es decir su latitud, longitud y altitud, en el momento de su captura. Para que se guarden los cambios hechos se debe de hacer click sobre el botón 'Guardar cambios'.



**Figura 6. Configuración de parámetros de cámara**

### 4.2.3 Normalización relativa

La corrección mediante normalización relativa se aplica cuando se cuenta con imágenes multitemporales, y se desea empatar la radiometría de una imagen con la de otra, por lo que para iniciar el proceso se deben de seleccionar dos imágenes, ya sea en formatos convencionales (.jpg, .png) o multiespectrales (.hdr), una servirá como la de referencia y la otra será la imagen objetivo que se va a normalizar.

El algoritmo utilizado en el sistema (Elvidge et al, 1995) requiere conocer la ubicación de aquellos puntos de alta y baja reflectividad cuyo valor en ambas imágenes se mantenga lo más estable posible al paso del tiempo. El sistema permite al usuario introducir los valores de la ubicación del pixel para alta y baja reflectividad. En el caso que se desconozcan, el sistema cuenta con una opción de 'Selección interactiva' en donde se muestran ambas imágenes y los diagrama de dispersión correspondientes a las bandas espectrales utilizadas por el algoritmo. Para seleccionar el punto que el usuario considere sea de mayor reflectividad estable, debe de hacer click con el botón izquierdo del mouse sobre cualquier punto de ambas imágenes, lo que hace que aparezca un cuadrado rojo sobre el punto seleccionado en ambas imágenes además

de mostrar su ubicación en los diagramas de dispersión, con esto el usuario se puede guiar para encontrar la ubicación en el diagrama de dispersión que se encuentre alrededor del centro del cluster de valores de alta reflectividad (Fig. 7). Al seleccionar un punto, su ubicación en coordenadas x,y se escribe en los cuadros de texto correspondientes. El proceso es el mismo para encontrar el punto de baja reflectividad estable, sólo que se tiene que seleccionar haciendo click con el botón derecho del mouse y aparece un cuadrado azul. Después de seleccionar ambos puntos se debe de introducir el valor para HPW (Anchura perpendicular media) y hacer click sobre el botón 'Normalizar' para finalizar el proceso, teniendo como resultado la imagen objetivo normalizada de acuerdo a la imagen de referencia.

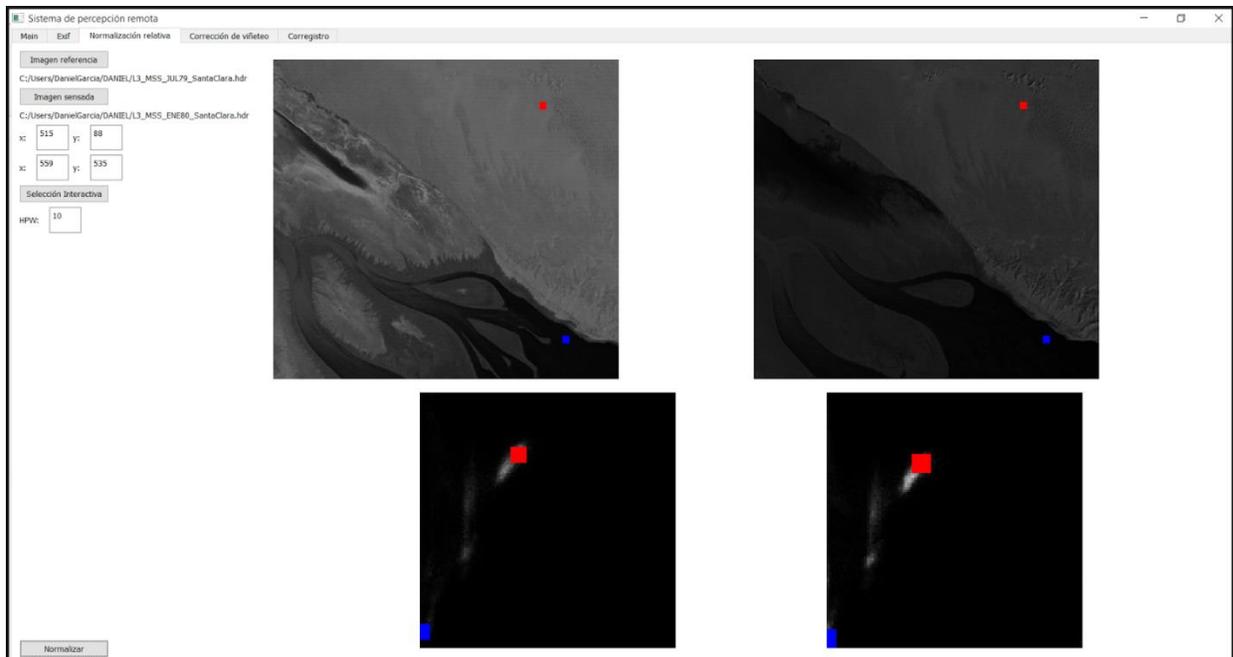


Figura 7. Selección interactiva de puntos estables de alta y baja reflectividad.

#### 4.2.4 Corrección de viñeteo

Como se mencionó en el Capítulo 2, el algoritmo de corrección de viñeteo utilizado en este sistema requiere de un grupo grande de imágenes para poder generar una imagen promedio de la cual obtener el filtro a aplicar sobre ellas, por lo tanto lo único necesario en el módulo de Corrección de viñeteo es seleccionar la carpeta en donde se encuentren las imágenes a corregir y hacer click sobre el botón 'Corregir' como se muestra en la figura 8.

El resultado de este proceso es una carpeta en donde se encuentran todas las imágenes anteriores pero ahora corregidas por el filtro calculado, por lo que se considera que el efecto de viñeteo ha sido eliminado.

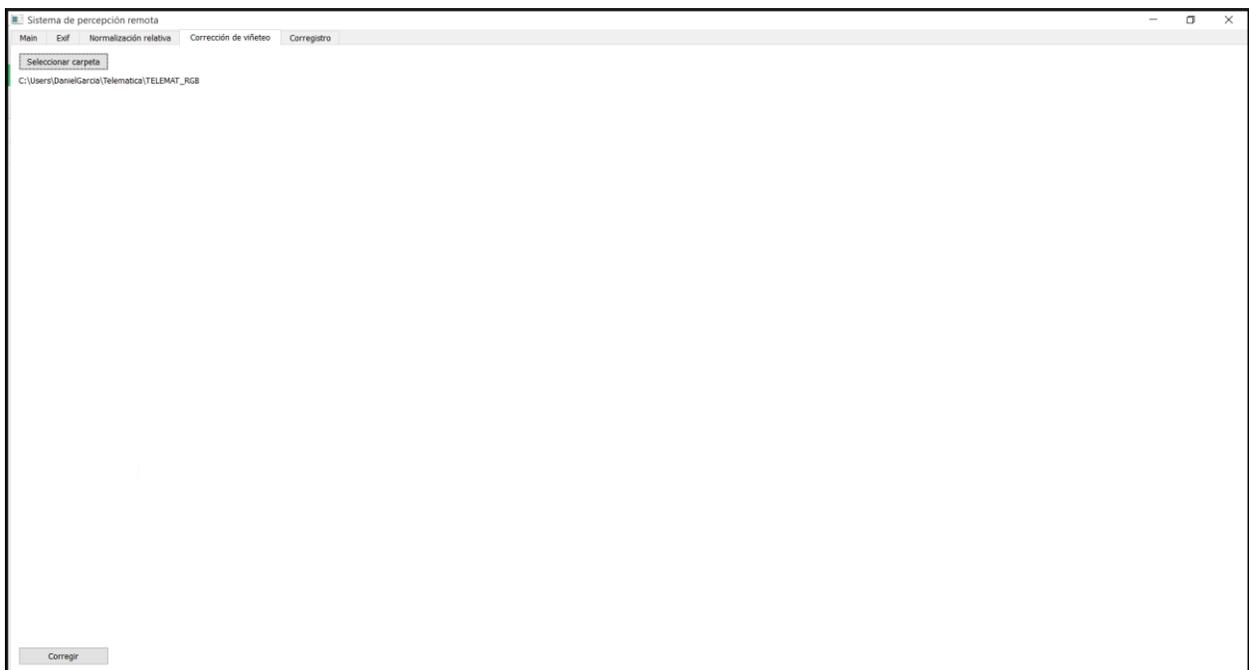


Figura 8. Corrección de viñeteo

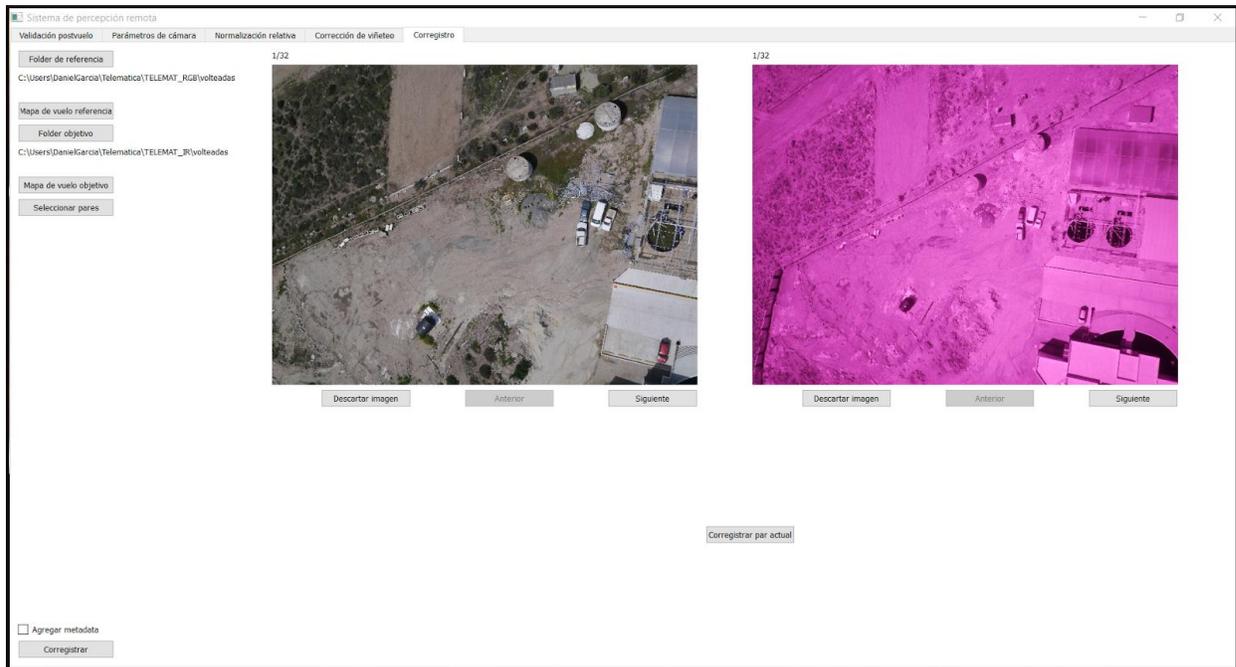
#### 4.2.5 Corregistro

Aunque el corregistro se puede realizar a todo un grupo de imágenes o de manera individual, es necesario seleccionar las carpetas en donde se encuentran la/las imágenes de referencia y la/las imágenes objetivo. Después de haberlas seleccionado, se puede hacer la selección de pares de forma similar que en el módulo de Validación post-vuelo, con la diferencia que aquí además de la comparación y descarte se puede decidir corregistrar cualquier par seleccionado haciendo click sobre el botón 'Corregistrar par actual' (Figura 9).

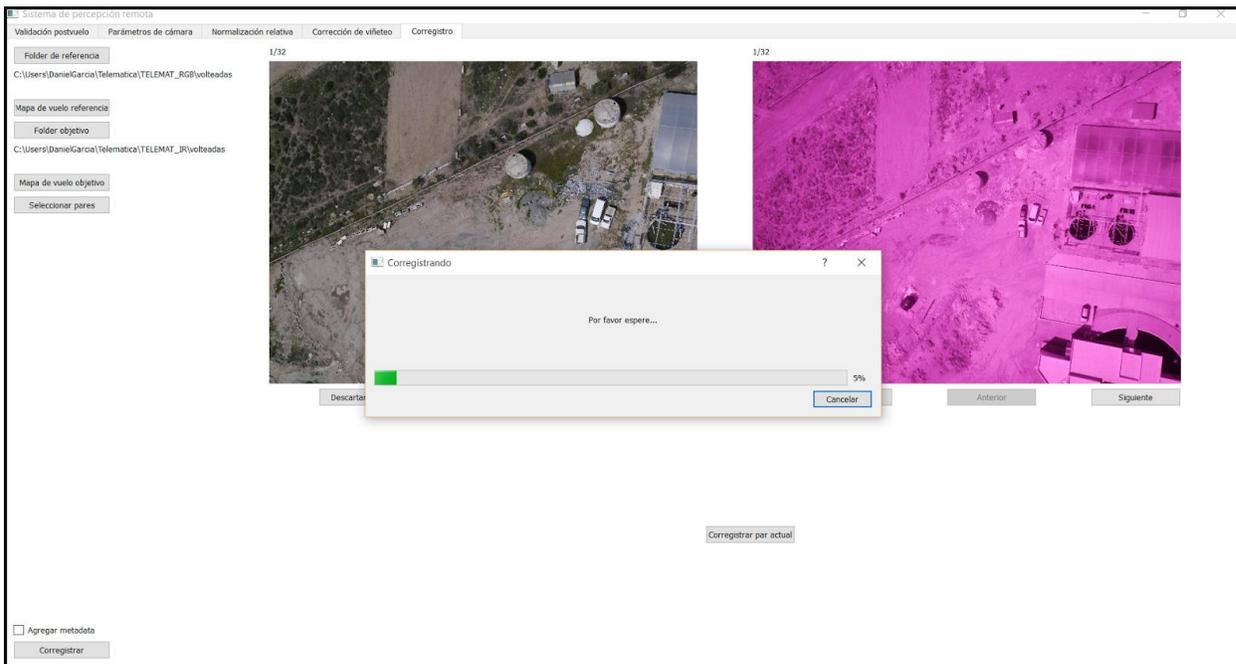
Antes de empezar el corregistro el usuario puede decir si quiere que se generen metadatos acerca del proceso al seleccionar la casilla de confirmación 'Agregar metadata'. Los metadatos en este proceso es la información de los puntos correspondientes, la matriz de homografía usada en la transformación, e imágenes en falso color infrarrojo en formato .jpg.

Al hacer click sobre el botón 'Corregistrar' se inicia el proceso de corrección para ambos grupos de imágenes, lo que puede ser muy tardado, por lo que se informa al usuario sobre el progreso del proceso mediante una barra de carga (Figura 10).

El producto generado por el proceso son imágenes multiespectrales (6 bandas) en formato .bsq y su encabezado .hdr correspondiente.



**Figura 9. Validación antes del corregistro y corregistro individual**



**Figura 10. Barra de carga durante el corregistro**

## Capítulo 5. Resultados

---

El objetivo del sistema es generar imágenes multiespectrales para su uso en percepción remota y en distintas áreas de aplicación, por lo que es necesario conocer la precisión del producto generado, ya que ello indicará el tipo de aplicaciones en donde son de utilidad. En este capítulo se presenta una evaluación del proceso de correregistro y se presentan los resultados de la corrección de viñeteo.

### 5.1 Evaluación de la precisión del correregistro de imágenes

Para la corrección geométrica, el proceso de correregistro debe dejar alineadas las imágenes lo más preciso posible, de esto depende su utilidad en el análisis cuantitativo en donde se selecciona un píxel y se desea obtener su información radiométrica para la mayor cantidad de bandas posibles.

Las clases de errores en el correregistro de imágenes se clasifican de acuerdo al proceso en donde son producidos.

*Error de localización.* Es el desplazamiento de las coordenadas de los puntos de control debido a su detección no precisa. Debido a que se trata de un error intrínseco al método de detección, el error de localización no se puede medir directamente en una imagen dada.

*Error de empatamiento.* Se mide por el número de falsos empatamientos cuando se establece la correspondencia entre los puntos de control. Ya que se trata de un error importante al momento del correregistro, es necesario usar algoritmos de empatamiento robustos, como fue nuestro caso usando FLANN y RANSAC.

*Error de alineamiento.* Es la diferencia entre el modelo de mapeo usado para el correregistro y la distorsión geométrica real entre las imágenes.

Debido a que en este estudio no se trabaja con imágenes georreferenciadas, se tomará en consideración la evaluación de los dos últimos tipos de errores.

### **5.1.1 Error de empataamiento**

Una manera de identificar errores en el empataamiento es mediante un método llamado comprobación de consistencia, que consiste en aplicar distintos métodos de empataamiento a los mismos grupos de puntos de control. Para realizar las pruebas se utilizaron FLANN (implementado en el sistema) y BFSURF(un método de fuerza bruta orientado a SURF). Se realizó una prueba con los datos de la misión aérea “Retorno” y como resultado se obtuvo que de aquellos puntos que coinciden en ambos métodos, sólo el 0.01073% de ellos eran falsos o tenían un nivel de error bajo.

Para complementar esta comprobación, también se realizó un análisis visual sobre las primeras imágenes durante el desarrollo del algoritmo de correregistro, con el objetivo de asegurar el funcionamiento del empataamiento, lo cual permitió hacer una evaluación subjetiva. Por lo tanto se consideran sólo los puntos de interés que el algoritmo RANSAC considera que no son atípicos y que generan los mejores parámetros para el modelo de la transformación. Estos puntos fueron desplegados y enumerados sobre la imagen a la que corresponden y al realizar la evaluación visual se puede corroborar que aquellos puntos con el mismo valor se encuentren en la misma ubicación en el terreno. En las figuras 11 y 12 se muestra un ejemplo para un par de imágenes.

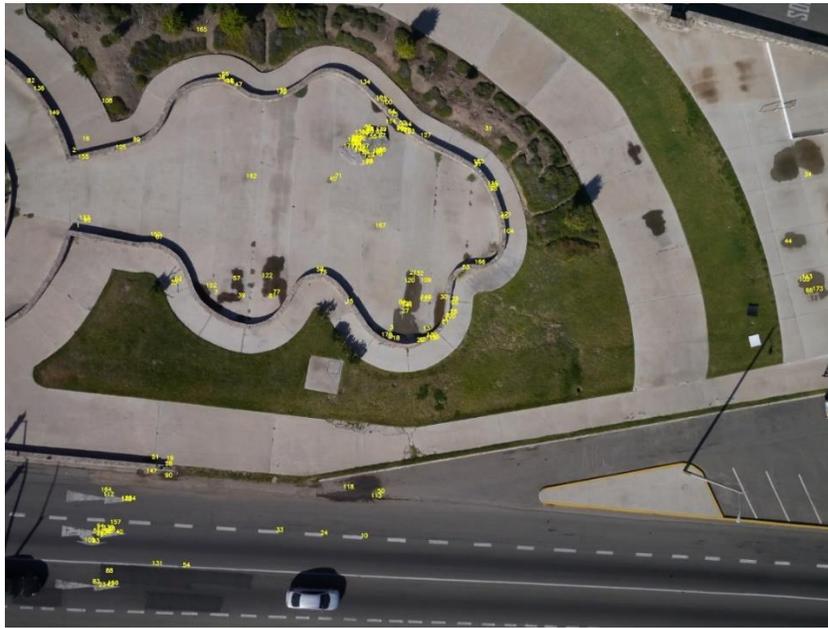


Figura 11. Imagen visible con sus puntos de interés correspondientes.

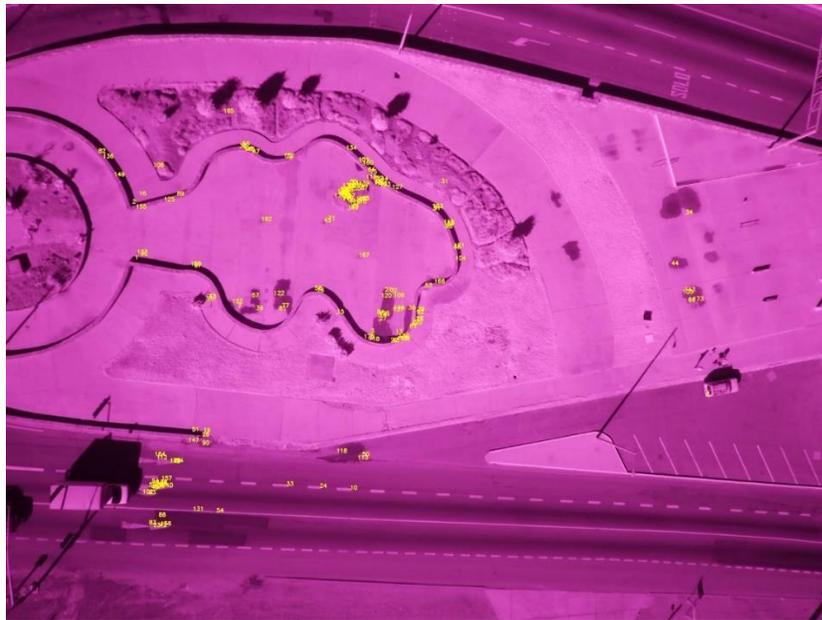


Figura 12. Imagen infrarroja con sus puntos de interés correspondientes.

### 5.1.2 Error de alineamiento

Se utiliza el Error Cuadrático Medio (ECM o MSE por sus siglas en inglés) para la evaluación del error de alineamiento.

El ECM se define como:

$$ECM = \frac{1}{n} \sum_{j=1}^n (\hat{Y}_i - Y_i)^2$$

En donde  $\hat{Y}_i$  es un vector de n predicciones, y  $Y_i$  es el vector de los valores observados correspondientes a las entradas de la función que generó las predicciones.

Para la evaluación se decidió obtener el ECM para el residual en X, para el residual en Y y para la distancia euclidiana.

El residual se obtiene a partir de la homografía que se calculó para transformar la imagen, y se define como

$$resX = |refX - objX * M|, resY = |refY - objY * M|$$

en donde  $refX$  y  $refY$  son los valores X y Y respectivamente de la imagen de referencia,  $objX$  y  $objY$  son los valores X y Y respectivamente de la imagen objetivo. Estos valores  $resY$  y  $resX$  indican el error en el corregistro en pixeles para cada punto en particular. En la ecuación del ECM  $refX$  reemplaza a  $\hat{Y}$  y  $objX * M$  reemplaza a  $Y$ .

También se midió el error en distancia euclidiana entre el punto de la imagen de referencia y el obtenido después de la transformación como

$$dist = \sqrt{(objX - resX)^2 + (objY - resY)^2}$$

y en la ecuación del ECM  $dist$  reemplaza  $\hat{Y}_i - Y_i$ .

El ECM mide el error de alineamiento en un par de imágenes dadas (Tabla 1) pero también se puede medir el promedio del ECM para todas las imágenes en una misión aérea dada (Tabla 2).

**Tabla 1. Error cuadrático medio en grupo de imágenes 'Retorno'**

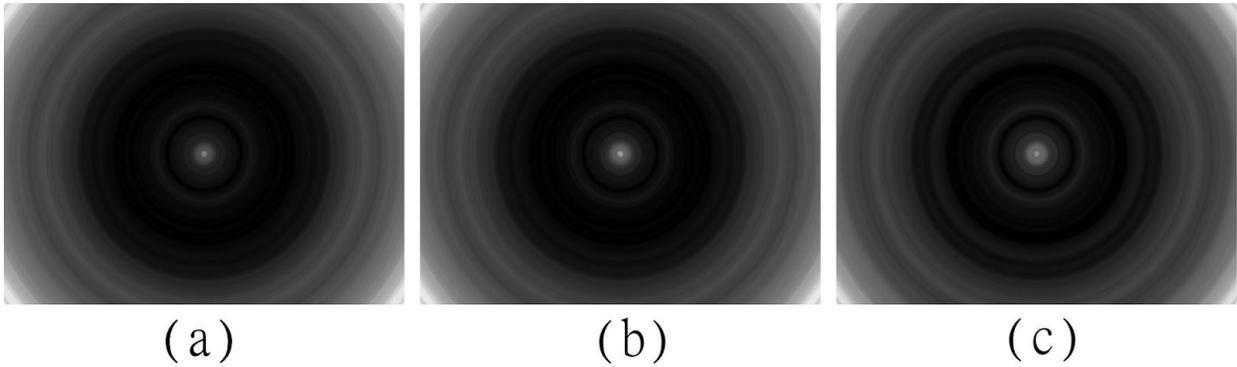
<b>Imágenes</b>	<b>Residual X (en px)</b>	<b>Residual Y (en px)</b>	<b>Distancia (en px)</b>
1.jpg	0.83911	0.60065	1.13968
2.jpg	0.88686	0.78070	1.30725
3.jpg	0.80789	0.66258	1.15537
4.jpg	0.81891	0.69086	1.17791
5.jpg	0.75785	0.64739	1.11234

**Tabla 2. Error cuadrático medio promedio en cada grupo de imágenes**

<b>Misión</b>	<b>Residual X (en px)</b>	<b>Residual Y (en px)</b>	<b>Distancia (en px)</b>
Telemática	0.77469	0.70803	1.16120
Retorno	0.83222	0.67850	1.18807

## **5.2 Resultados de la corrección de viñeteo**

El método utilizado para la corrección de viñeteo en el sistema implementado depende de contar con un grupo de imágenes obtenidas con el mismo sensor. A partir de esas imágenes se obtiene un perfil del efecto de viñeteo del sistema de adquisición, en la figura 13 se ve un ejemplo del inverso del perfil de viñeteo, el cual se aplica como un filtro para la corrección.



**Figura 13.** Perfiles inversos de la función de viñeteo del grupo de imágenes del caso de estudio 'Retorno'. (a) perfil de la banda roja, (b) perfil de la banda verde y (c) perfil de la banda azul.

Con base en el perfil generado y un análisis visual (Figura 14), se concluye que la corrección de viñeteo produce imágenes con una iluminación más uniforme y más cercana a la reflectividad real de la escena, lo cual será de utilidad en un posterior análisis cuantitativo de la radiometría de la escena en aplicaciones de percepción remota.



**Figura 14.** (a) Imagen original. (b) Imagen con el efecto de viñeteo corregido.

### 5.3 Resultados de la normalización

Se realizaron pruebas de la implementación del algoritmo con 3 pares de imágenes satelitales tomadas en 1979 y en 1980 con el sensor Thematic Mapper del satélite Landsat 5. Esas imágenes se encuentran alineadas geométricamente pero cuentan con diferente radiometría, lo que las hace adecuadas para este proceso.

En la figura 15 se muestra el producto de la normalización para el par de imágenes tomado en Santa Clara, en donde para su visualización cada imagen se trata de un falso color infrarrojo compuesto por dos bandas infrarrojas y la del rojo, que corresponden a las bandas 4, 3 y 2 en Thematic Mapper. Se puede ver como la radiometría de la imagen normalizada es más parecida a la de la imagen de referencia que a la imagen objetivo de la cual se obtuvo.

Tal normalización se puede apreciar de una manera más puntual mediante el uso de los histogramas de la radiometría de cada imagen (Fig. 16). En donde se puede observar cómo se diferenciaban sus histogramas originales, en el caso de la imagen de 1980 sus valores se sesgaban hacia la izquierda, que coincide con su baja reflectancia visualmente. En cambio, en la imagen de 1979 su distribución es un poco más uniforme de 0 a 140.

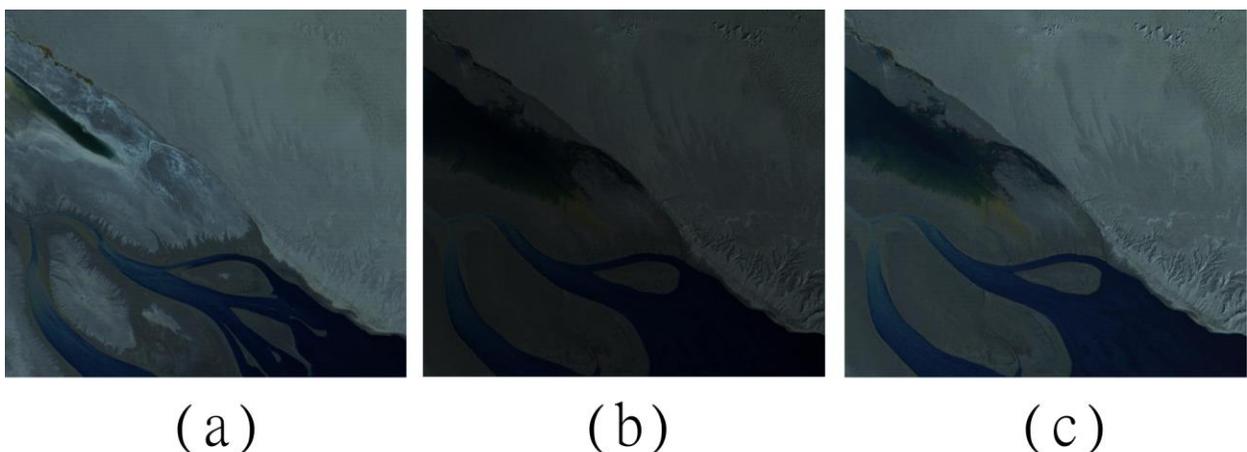
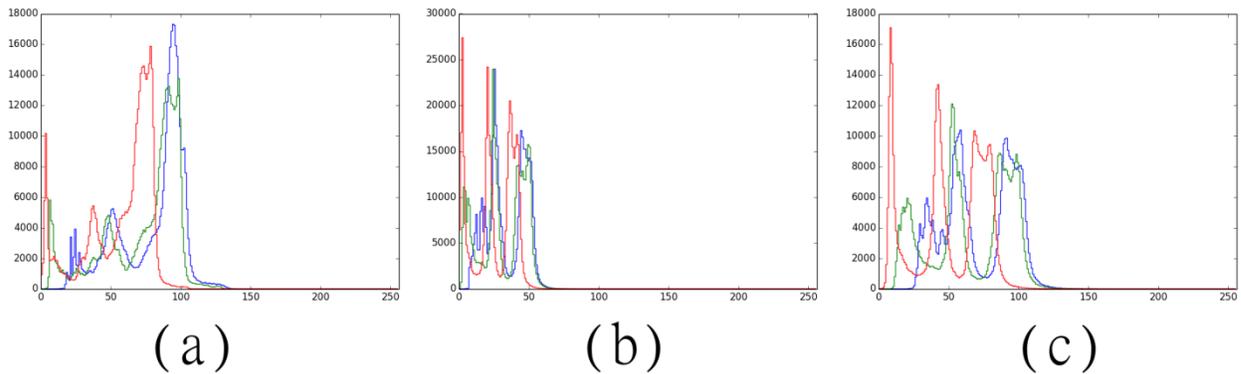


Figura 15. Imágenes satelitales tomadas en Santa Clara con el sensor Thematic Mapper del satélite Landsat 5. (a) Imagen de referencia tomada en 1979. (b) Imagen objetivo tomada en 1980. (c) Imagen objetivo con radiometría normalizada en base a la imagen de referencia.

En el histograma de la imagen normalizada tal sesgo hacia la izquierda se elimina, al extender el rango dinámico con una distribución más uniforme, parecida a la de la imagen de referencia. Las diferencias entre los picos se deben en parte a los cambios del suelo después de un año.



**Figura 16. Histogramas de las 3 bandas del par de imágenes ‘Santa Clara’ y su normalización. La línea roja coincide con la segunda banda, la línea verde con la tercera y la línea azul con la cuarta. (a) Histograma de la imagen de referencia tomada en 1979. (b) Histograma de la imagen objetivo tomada en el 1980. (c) Histograma de la imagen objetivo con radiometría normalizada en base a la imagen de referencia.**

Las tablas 3-5 muestran los datos estadísticos de las imágenes de la figura 15a, 15b y 15c. Aunque los valores de mínimo, máximo y media de la imagen normalizada suben, debido al estiramiento de su rango dinámico, en su desviación estándar se puede ver con más claridad como su distribución se aproxima a la de la imagen de referencia.

**Tabla 3. Datos estadísticos Santa Clara de 1979**

	<b>Min</b>	<b>Max</b>	<b>Media</b>	<b>Desviación estándar</b>
<b>Banda 1</b>	14	127	54.6630	13.2801
<b>Banda 2</b>	13	127	78.8223	24.0336
<b>Banda 3</b>	2	127	73.1866	27.2805
<b>Banda 4</b>	0	111	56.9355	23.6787

Tabla 4. Datos estadísticos Santa Clara de 1980

	<b>Min</b>	<b>Max</b>	<b>Media</b>	<b>Desviación estándar</b>
<b>Banda 1</b>	9	54	24.2204	6.5372
<b>Banda 2</b>	5	89	33.5165	13.8809
<b>Banda 3</b>	0	91	31.0044	16.1640
<b>Banda 4</b>	0	75	24.5465	14.5182

Tabla 5. Datos estadísticos Santa Clara normalizada de 1980

	<b>Min</b>	<b>Max</b>	<b>Media</b>	<b>Desviación estándar</b>
<b>Banda 1</b>	27	92	49.34.16	9.4798
<b>Banda 2</b>	21	168	70.6375	24.4072
<b>Banda 3</b>	9	172	64.7932	28.9660
<b>Banda 4</b>	6	140	49.6911	25.8798

## Capítulo 6. Conclusiones

---

Se diseñó e implementó un sistema de software para el procesado y análisis post-vuelo de fotografía aérea de formato pequeño. El procesado consiste en la corrección geométrica y radiométrica de imágenes con el objetivo de generar imágenes multiespectrales a partir de fotografía aérea de formato pequeño para su uso en aplicaciones de percepción remota. El procesamiento consiste en la aplicación de herramientas de visualización y procesado de imágenes adquiridas en misiones aéreas, con el objetivo de validar su calidad radiométrica, así como la existencia de posibles huecos en su cubrimiento.

Para la corrección geométrica se implementó un algoritmo de corrección de imágenes, el cual alinea las imágenes tomadas con una cámara modificada para captar el espectro infrarrojo, con imágenes convencionales en el visible. El producto del corrección se evaluó cuantitativamente y se encontró que el error promedio en el alineamiento es menor que un pixel, lo cual indica la obtención de imágenes de alta calidad visual, así como su utilidad en aplicaciones de mapeo y de percepción remota donde pudiera ser necesario trabajar con una alta precisión.

En el tema de la corrección radiométrica se corrigió el efecto de viñeteo en las imágenes y se implementó un algoritmo de normalización para juegos de imágenes multitemporales. Como resultado de esto se obtuvieron imágenes con una radiometría mejor balanceada en todo el plano focal, así como una mayor estabilidad temporal en su radiometría, lo cual es de utilidad en aplicaciones cuantitativas de percepción remota.

Las imágenes generadas son similares espectralmente a otros productos comúnmente empleados en percepción remota, como por ejemplo las imágenes de alta resolución espacial de los satélites QuickBird y GeoEye, y otros productos históricos de los sensores MultSpectral Scanner que orbitaron a bordo de los satélites Landsat 1, 2, 3, 4 y 5.

## 6.1 Trabajo a futuro

A continuación se mencionan algunas sugerencias sobre líneas de trabajo que se pueden seguir relacionadas a la temática de esta tesis.

1. Modularizar el sistema de tal manera que el componente de análisis post-vuelo se pueda implementar como una aplicación móvil independiente, para facilitar su uso en el campo durante la misión aérea.
2. Implementar técnicas de calibración espectral de las imágenes generadas, con el objeto de hacerlas compatibles con otros sistemas multiespectrales de uso común.
3. Realizar la calibración métrica de las cámaras para obtener imágenes rectificadas geométricamente para aplicaciones de mapeo fotogramétrico.
4. Diseñar un módulo que permita la adquisición y procesamiento de los datos generados por el sistema de posicionamiento y orientación (GPS, IMU) para poder realizar el georreferenciado directo de las imágenes.
5. Integrar esta información multiespectral con otras fuentes de información de percepción remota, como imágenes térmicas y cubrimientos con LiDAR aéreo.

## Lista de referencias bibliográficas

- Aguilera, C., Barrera, F., Sappa, A. D., & Toledo, R. (2012). A Novel SIFT-Like-Based Approach for FIR-VS Images Registration. *Proc. Quantitative InfraRed Thermography*.
- Aber, J. S., I. Marzloff and J. B. Ries (2010). Small-Format Aerial Photography : principles, techniques and geoscience applications. Elsevier, 266 pp. Aber, J. S., S. W.
- Aber, J. Janoeko, R. Zabielski and M. Gorska-Zabielska (2008). High-altitude kite aerial photography. Kansas Academy of Science, *Transactions* 111, p. 4960.
- Aldred, A. H. and R. J. Hall (1992). Forest regeneration appraisal with large scale photographs. *The Forestry Chronicle*, V. 68, 142-150.
- Bay, H., Tuytelaars, T., and Gool, L. Van. (2008). SURF : Speeded Up Robust Features. *Computer Vision and Image Understanding*, 110(3), 346-359.
- Berni, J. A. J., Member, S., Zarco-tejada, P. J., Suárez, L., and Fereres, E. (2009). Thermal and Narrowband Multispectral Remote Sensing for Vegetation Monitoring From an Unmanned Aerial Vehicle. *Geoscience and Remote Sensing, IEEE Transactions On*, 47(3), 722–738. doi:[10.1109/TGRS.2008.2010457](https://doi.org/10.1109/TGRS.2008.2010457)
- Campbell, J. B. (2011). *Introduction to Remote Sensing*. The Guilford Press
- Chanond, C., and Leekbhai, C., (1986). Small format aerial photography for analysing urban housing problems : a case study in Bangkok. *ITC Journal*, V.3, 197-205.
- Dickenson, J., (2009). Pacific atoll living : How long already and until when ?. *Geology*, 19(3), 4-10.
- Elvidge, C. D., Yuan, D., Weerackoon, R. D., and Lunetta, R. S. (1995). Relative radiometric normalization of Landsat Multispectral Scanner (MSS) data using an automatic scattergram-controlled regression. *Photogrammetric Engineering & Remote Sensing*, 61, 1255–1260.
- Fryers, J. G., (1992). Photogrammetric monitoring of cliffs. *The Australian Surveyor*, V.37, 270-274.
- Goldman, D. B. (2005). Vignette and exposure calibration and compensation. *Tenth IEEE International Conference on Computer Vision (ICCV'05) 1*, 899–906. doi:[10.1109/ICCV.2005.249](https://doi.org/10.1109/ICCV.2005.249)
- Gomasasca, M. A. (2009). *Basics of geomatics*. Springer Science and Business Media.

- Goshtasby, A. (2011). Transformation functions for image registration.II. *CVPR tutorial on tools and methods for image registration*, 1-53.
- Gupta. R.P. (2003). *Remote sensing geology*. Springer.
- Goshtasby, A. (2011). Transformation functions for image registration.II. *CVPR tutorial on tools and methods for image registration*, 1-53.
- Hofstee, P. (1985). An introduction to small format aerial photography for human settlement analysis. *ITC Journal*, V. 2, 121-127
- Horning, N. et al. (2010). *Remote sensing for ecology and conservation: A handbook of techniques*. Oxford University Press.
- Jensen, I. R. (2007). *Remote sensing of the environment*. Pearson Prentice Hall.
- Johnson, L., and Herwitz, S. (2003). Collection of ultra high spatial and spectral resolution image data over California vineyards with a small UAV. *Proc. 30th International Symposium on Remote Sensing of Environment (ISRSE)*,3–5.
- Kennie, T. J. and Matthews, M.C. (1985). *Remote sensing in civil engineering*. Surrey University Press.
- Laliberte, A. S., Goforth, M. a., Steele, C. M., and Rango, A. (2011). Multispectral remote sensing from unmanned aircraft: image processing workflows and applications for rangeland environments. *Remote Sensing*, 3(12), 2529–2551. doi:[10.3390/rs3112529](https://doi.org/10.3390/rs3112529)
- Lebourgeois, V., Bégué, A., Labbé, S., Mallavan, B., Prévot, L., and Roux, B. (2008). Can commercial digital cameras be used as multispectral sensors? A crop monitoring test. *Sensors*, 8(11), 7300–7322. doi:[10.3390/s8117300](https://doi.org/10.3390/s8117300)
- Levin, E., and Curelli, S. (2011). Photogrammetric Small UAV in geospatial research and education at Michigan Tech University. In *American Society for Photogrammetry and Remote Sensing Annual Conference 2011* (pp. 271–280).
- Li, Q., Wang, G., Liu, J., and Chen, S. (2009). Robust scale-invariant feature matching for remote sensing image registration. *IEEE Geoscience and Remote Sensing Letters*, 6(2), 287–291. doi:[10.1109/LGRS.2008.2011751](https://doi.org/10.1109/LGRS.2008.2011751)
- Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2), 91–110. doi:[10.1023/B:VISI.0000029664.99615.94](https://doi.org/10.1023/B:VISI.0000029664.99615.94)
- Marzloff, I. and Ries, J. B. (2007). Gully monitoring in semi-arid landscapes. *Zeitschrift für Geomorphologie*, 51(4), 405-425.

- Morán Mendoza, Antonio Ignacio. (2014) *Algoritmos para la corrección radiométrica de imágenes de satélite*. Tesis de maestría en ciencias. Centro de Investigación Científica y de Educación Superior de Ensenada, Baja California.
- Ramírez Canales, Héctor Iván. (2013) *Restitución fotogramétrica a partir de fotografía aérea de pequeño formato*. Tesis de maestría en ciencias. Centro de Investigación Científica y de Educación Superior de Ensenada, Baja California.
- Ramirez, H.I. y Torres, J. (2012). Restitución fotogramétrica a partir de fotografías aéreas de pequeño y gran formato. *2do Congreso Nacional y 1er Latinoamericano de Ciencia y tecnología Aeroespacial*. Sociedad Mexicana de Ciencia y Tecnología Aeroespacial.
- Rees, W. G. (2013). *Physical principles of remote sensing*. Cambridge University Press.
- Ridd, M.K. and Hipple, J.H. (2006). Remote sensing of human settlements. *Manual of Remote Sensing, Vol. 5*. ASPRS.
- Ries, J.B. and Marzloff, I. (2003). Monitoring of gully erosion in the Central Ebro Basin by large scale aerial photography taken from a remotely controlled blimp. *Catena*, 50(2), 309-328.
- Ries, J. B. and Hirt, U. (2008). Permanence of soil surface crusts on abandoned farmland in the Central Ebro Basin/Spain. *Catena*, 72(2), 282-296.
- Richards, J.A. and Jia, X. (2005). *Remote sensing digital image analysis: An introduction*. Springer.
- Ryerson, R.A. (2006). Remote sensing of the marine environment. *Manual of remote sensing*. ASPRS.
- Sabins, F.F. (1996). *Remote sensing: Principles and interpretation*. W.H. Freeman & Company. New York.
- Schowengerdt, R. A. (2009). *Remote sensing: Models and methods for image processing*. Elsevier Science and Technology.
- Seeger, M., Marzloff, I. and Ries, J.B. (2009). Identification of gully development processes in semi arid NE Spain. *Zeitschrift für Geomorphologie*, 53(4), 417-131.
- Torres, J. (2012). Mapeo aéreo usando fotografía de formato pequeño. E. Sánchez y R. Díaz (Eds.) *Dinámicas locales del cambio ambiental global: Aplicaciones de percepción remota y análisis espacial en la evaluación del territorio*. UACJ, pp. 347-353.
- Torres, J. (2010). Utilización de fotografía aérea de formato pequeño en tareas de reconocimiento y mapeo topográfico. *Memorias del XIV Simposio Internacional SELPER* (CD-ROM).

- Tuytelaars, T., and Mikolajczyk, K. (2007). Local invariant feature detectors: A survey. *Foundations and Trends® in Computer Graphics and Vision*, 3(3), 177–280. doi:[10.1561/06000000017](https://doi.org/10.1561/06000000017)
- USGS (1979). *Landsat data users handbook (revised)*. U.S. Department of the Interior.
- Ustin, S. (2004). Remote sensing for natural resource management and environmental monitoring. *Manual of remote sensing*, Vol. 4. Wiley.
- Watts, A. C., Ambrosia, V. G., and Hinkley, E. a. (2012). Unmanned aircraft systems in remote sensing and scientific research: Classification and considerations of use. *Remote Sensing*, 4(12), 1671–1692. doi:[10.3390/rs4061671](https://doi.org/10.3390/rs4061671)
- Yu, W. (2004). Practical anti-vignetting methods for digital cameras. *IEEE Transactions on Consumer Electronics*, 50, 975–983. doi:[10.1109/TCE.2004.1362487](https://doi.org/10.1109/TCE.2004.1362487)
- Zheng, Y., Lin, S., Kambhamettu, C., Yu, J., and Kang, S. B. (2009). Single-image vignetting correction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31, 2243–2256. doi:[10.1109/TPAMI.2008.263](https://doi.org/10.1109/TPAMI.2008.263)
- Zitová, B., and Flusser, J. (2003). Image registration methods: a survey. *Image and Vision Computing*, 21(11), 977–1000. doi:[10.1016/S0262-8856\(03\)00137-9](https://doi.org/10.1016/S0262-8856(03)00137-9)