

**PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ**  
**FACULTAD DE CIENCIAS E INGENIERÍA**



PONTIFICIA  
**UNIVERSIDAD**  
**CATÓLICA**  
DEL PERÚ

**DISEÑO E IMPLEMENTACIÓN DE UN ALGORITMO PARA EL  
REGISTRO DE IMÁGENES AÉREAS ORTORECTIFICADAS  
APLICADO EN LA SUPERVISIÓN AGRÍCOLA**

Tesis para optar el Título de **INGENIERO ELECTRÓNICO**, que presenta el bachiller:

**Fernando Javier Meza Zurita**

**ASESOR: Andrés Flores Espinoza**

Lima, mayo de 2015

## Resumen

La presente tesis muestra el diseño e implementación de un sistema de registro de imágenes aéreas orientado a la agricultura de precisión. Este sistema cuenta con la particularidad que las imágenes pasarán por un pre procesamiento de corrección geométrica para mejorar la calidad del registro. Este sistema permitirá la generación de panoramas globales de los campos de cultivo con la finalidad de obtener información correctamente localizada del campo analizado.

El documento está dividido en 4 capítulos y su estructura es de la siguiente manera:

En el capítulo 1, se analiza la problemática de la agricultura de precisión y el sensado remoto, así como también se expone la aplicación y la problemática del uso de aeromodelos e imágenes aéreas en la agricultura de precisión.

En el capítulo 2, se exponen los temas de registro y ortorectificación de imágenes como solución cuando se trabaja con imágenes aéreas orientado a la agricultura de precisión

En el capítulo 3, se explican los objetivos de este proyecto y se desarrolla detalladamente el sistema diseñado. Se definen las características particulares del vuelo para la obtención de las imágenes aéreas, a partir de estas características se realiza la selección de algoritmos y se desarrolla el modelo propuesto y una explicación detallada de los pasos y algoritmos usados en este proceso.

Finalmente, en el capítulo 4, se muestra la evaluación del sistema propuesto, el cual estará dividido en dos partes, en una primera instancia se evaluará únicamente la generación de mosaicos a partir de un banco de imágenes que se tendrá de referencia y posteriormente se evaluará todo el sistema completo utilizando imágenes aéreas reales adquiridas en la PUCP.

## Índice

<b>Introducción.....</b>	<b>3</b>
<b>Capítulo 1: Imágenes aéreas en la agricultura de precisión.....</b>	<b>4</b>
1.1 Agricultura de precisión.....	4
1.2 Sensado remoto.....	5
1.3 Problemas del sensado remoto.....	6
1.4 Síntesis de la problemática.....	8
<b>Capítulo 2: Método para el registro y ortorectificación de imágenes aéreas....</b>	<b>9</b>
2.1 Registro de imágenes.....	9
2.2 Ortorectificación de imágenes aéreas.....	11
<b>Capítulo 3: Diseño e implementación del sistema de ortorectificación y registro de imágenes.....</b>	<b>18</b>
3.1 Objetivos.....	18
3.2 Características del vuelo .....	18
3.3 Modelo del sistema.....	22
3.4 Selección de algoritmos.....	24
3.5 Descripción del funcionamiento del sistema.....	25
3.6 Funcionamiento del sistema con imágenes.....	30
<b>Capítulo 4: Ensayos y resultados.....</b>	<b>34</b>
4.1 Alcances del sistema .....	34
4.2 Condiciones de pruebas con imágenes aéreas .....	34
4.3 Evaluación de mosaicos.....	35
4.4 Resultado del banco de imágenes .....	37
4.5 Discusión de resultados de imágenes de prueba.....	42
4.6 Análisis de imágenes aéreas.....	42
<b>Conclusiones.....</b>	<b>51</b>
<b>Recomendaciones.....</b>	<b>52</b>
<b>Bibliografía.....</b>	<b>53</b>

## Introducción

En el Perú, la agricultura siempre ha sido una parte representativa de nuestra cultura esto debido a la gran diversidad de productos agrícolas que genera el país. Sin embargo, esta actividad ha ido perdiendo impacto a lo largo de los años, ya que actualmente las actividades agrícolas aportan sólo el 7.6% del PBI nacional, lo cual es la mitad de lo que generaba hace 40 años [1]. Una de las grandes dificultades de la agricultura nacional se encuentra en el monitoreo y supervisión de los campos de cultivo.

A pesar de los avances tecnológicos, los métodos para el monitoreo de campos de cultivo en la agricultura son muy tradicionales; es decir, se basan en la observación y los conocimientos del agricultor. Esto conlleva a que la interpretación del campo se realice de manera subjetiva, lo cual podría derivar en un tratamiento del campo de cultivo no tan eficiente generando pérdidas. Es bajo este contexto que aparece como solución la agricultura de precisión que consiste en el sensado de ciertas características del suelo mediante el uso de imágenes multiespectrales tomadas por aeronaves, lo que permite una supervisión más sencilla y precisa del campo de cultivo, así como un tratamiento focalizado de este, que no solo generará una mejor cosecha sino también se podrá ahorrar en insumos tales como fertilizante, agua, pesticidas, etc.

Debido a que las imágenes tomadas por la aeronave son de partes muy reducidas del campo, estas tienen que ser procesadas para poder unir las y generar un panorama global de la zona, este proceso de integrar las imágenes se conoce como registro de imágenes. Cabe resaltar que la aeronave no toma las imágenes perfectamente planas, lo cual genera una perspectiva de una imagen con respecto a otra y si esta diferencia es muy notoria se generará un registro equivocado. Es por esta razón que existe la necesidad de corregir esta perspectiva antes de realizar el registro de las imágenes, esto mejorará la calidad del registro y ayudará a reducir el tiempo de procesamiento del algoritmo, este proceso de corrección es conocido como ortorectificación.

El presente trabajo de tesis tiene como objetivo diseñar e implementar un algoritmo que permita la ortorectificación y posterior registro de imágenes aéreas de campos de cultivos para supervisión agrícola.



## Capítulo 1: Imágenes aéreas en la agricultura de precisión

La producción agrícola requiere de sistemas de monitoreo eficientes con la finalidad de obtener una alta productividad. En la actualidad, la actividad agrícola en el Perú usa técnicas de monitoreo obsoletas, lo cual le impide brindar productos de calidad y mejorar la utilización de sus recursos, es así que aparece la agricultura de precisión como posible solución para mejorar las técnicas de supervisión en base a los últimos avances tecnológicos.

### 1.1 Agricultura de precisión

La agricultura de precisión es un concepto agronómico de gestión de campos de cultivo, basado en el análisis de la variabilidad del cultivo. Para esto se requiere del uso de tecnologías vigentes tales como: Sistema de Posicionamiento Global (GPS), sensores, satélites e imágenes aéreas junto con un Sistema de Información Geográfico para estimar, evaluar y entender dichas variaciones [2].

Lo que se busca con la agricultura de precisión es básicamente el uso de diferentes tecnologías con la finalidad de obtener información de los campos de cultivo para un mejor tratamiento de los mismos y así aumentar la rentabilidad y calidad de la agricultura

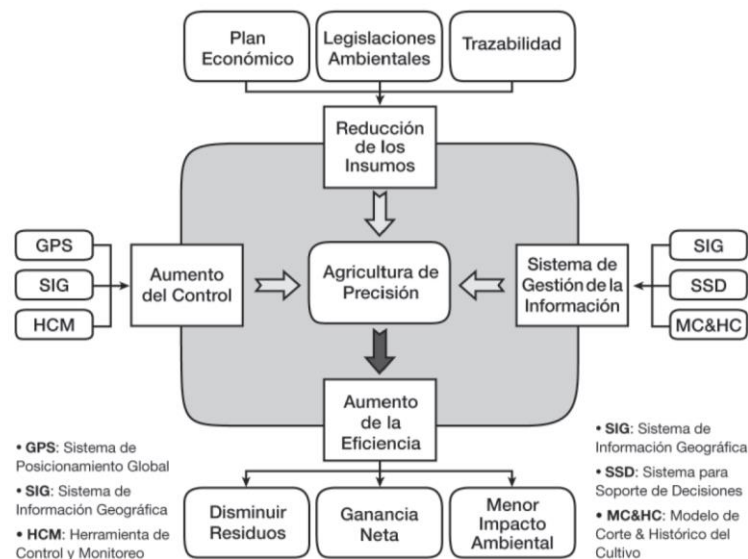


Figura 1: Integración entre los elementos de la agricultura de precisión [2].

Los objetivos específicos de la agricultura de precisión son:

- *Mejorar la producción de la agricultura tanto en calidad como cantidad*
- *Reducción del uso de insumos*
- *Reducción del impacto ambiental (menos pesticidas)*
- *Aumentar la rentabilidad de la agricultura*

## 1.2 Sensado remoto

El sensado remoto es la ciencia y el arte de obtener información sobre un objeto, área o fenómeno a través del análisis de datos obtenidos con un aparato (sensor remoto) que no está en contacto físico con ese objeto, área o fenómeno bajo estudio. El sensor remoto puede estar a pocos centímetros o a varios kilómetros, dependiendo del sistema usado y de la información deseada. Ej.: sensor de nitrógeno, fotografías aéreas, imágenes satelitales, etc. [3].

En la agricultura de precisión, el sensado remoto puede ser usado de diferentes maneras, pero en particular en este documento, se tratará sobre el análisis de concentración de los elementos del suelo mediante el uso de imágenes aéreas multiespectrales. Esto es posible dado que un elemento cualquiera, al ser afectado por una fuente energética, absorbe cierta cantidad de dicha energía en diversas longitudes de onda y esta energía absorbida está relacionado con su nivel de concentración por lo tanto es posible conocer los niveles de concentración de los elementos de los suelos y analizar la concentración de los elementos importantes para la nutrición de los campos de cultivo

Para la adquisición de dichas imágenes aéreas, en el espectro deseado, se cuenta con las siguientes herramientas:

- Aeromodelo

El aeromodelo, que también puede ser diferentes tipos de aeronave no tripulada, es el equipo que volará sobre el perímetro del campo de cultivo a muestrear. Este aeromodelo puede controlarse manualmente mediante un control remoto y también tiene un modo de vuelo automático. Para la toma de datos de esta tesis se utilizó el comando de vuelo automático y se programó una ruta determinada para cubrir la totalidad de un campo específico.

- Cámaras y filtros

El aeromodelo cuenta con cámaras en la parte inferior que le permiten tomar fotos de todo lo que sobrevuela el avión. Para tomar las fotos en el rango de frecuencia de interés (rojo e infrarrojo) se cuentan con 2 filtros especiales que permiten la absorción de energía de dichas ondas.

- Circuito electrónico de control y adquisición de datos

El aeromodelo, en su interior, cuenta con una tarjeta electrónica dedicada que se encarga del control de vuelo, el sensado de las diferentes variables del vuelo y un sistema para la adquisición de las imágenes mediante el uso de las cámaras mencionadas en el punto anterior.

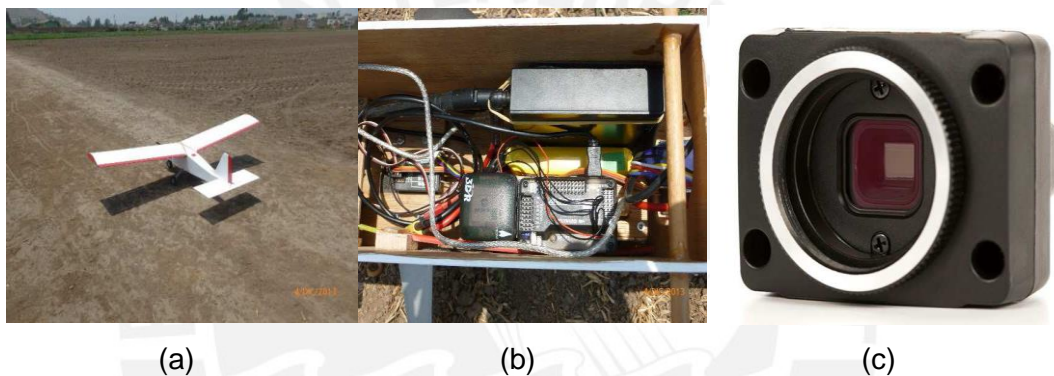


Figura 2: (a) Aeromodelo, (b) Cámaras, (c) Interior del aeromodelo (tarjeta electrónica para el control y adquisición de datos)

### 1.3 Problemas del sensado remoto

El principal problema del sensado remoto, en particular para la toma de imágenes aéreas, se encuentra en la adquisición de datos ya que al ser tomados desde una aeronave en constante movimiento y expuesta a distintos factores climáticos, principalmente viento, genera que esta toma de imágenes no sea paralela al suelo, lo cual es fundamental ya que de lo contrario se verán distorsiones en los tamaños reales del campo que se esté tomando de muestra.

A continuación se detallan los 3 tipos de movimiento que presenta una aeronave a lo largo de sus ejes (longitudinal, vertical y lateral). Estos 3 movimientos en el ámbito de la aeronáutica son conocidos como alabeo (roll), cabeceo (pitch) y guiñada (yaw).

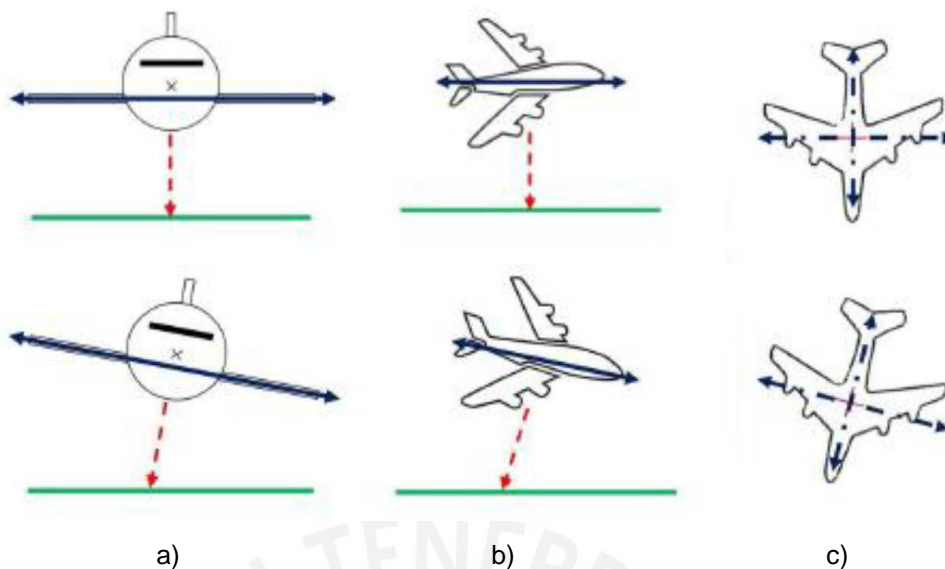


Figura 3: Grados de libertad de una aeronave. a) alabeo. b) cabeceo. d) guiñada.[9]

Estos movimientos de la aeronave generan que la toma de la imagen, que debería ser completamente paralela al suelo, salga distorsionada. Esto puede observarse en la siguiente imagen.

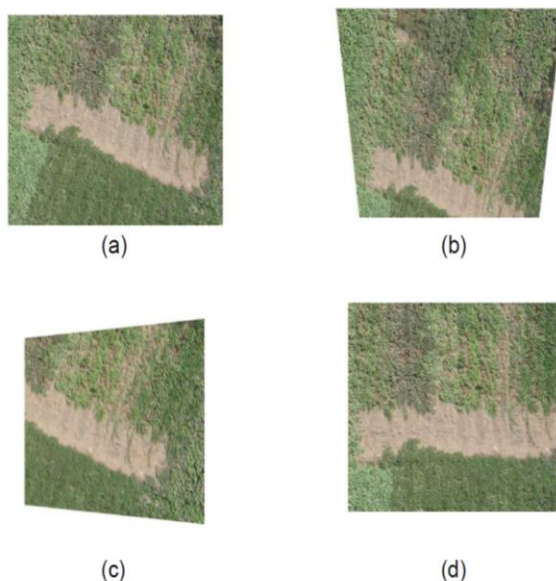


Figura 4: Adquisición de imágenes con los diferentes movimientos de la aeronave. (a) Imagen a partir de la aeronave alineada al suelo. (b), (c), (d) Imágenes a partir del alabeo, cabeceo y guiñada de la aeronave respectivamente. [9]

Como se puede observar en la figura 4, esta distorsión es un grave problema ya que da una idea errónea de las dimensiones y posición de la imagen que se está muestreando con las cámaras. Para esto hay diversas soluciones como utilizar un estabilizador mecánico para mantener la cámara perfectamente alineada con el suelo y así asegurar una adquisición de datos correctos. Esto implica la adquisición de

dicho sistema mecánico, el cual es costoso, complicado y pesado, lo cual es una gran desventaja para aeromodelos que tienen muy poca capacidad de carga y tiempo de vuelo limitado. Otra alternativa de solución a este problema, es la corrección de dichas imágenes a partir de la información obtenida por los sensores que dan los ángulos de inclinación de la aeronave (yaw, roll y pitch) y una vez obtenida esta información, realizar la corrección de dicha imagen haciendo uso de la técnica de ortorectificación, esta técnica será explicada en el capítulo 2. La ventaja de utilizar este tipo de enfoque para resolver el problema es que es rápido, sencillo y eventualmente podría ser implementado en tiempo real dentro del aeromodelo de manera que todas las imágenes se corrijan automáticamente y así obtener siempre una toma de datos correcta.

Otro problema de usar imágenes aéreas es que las imágenes obtenidas son de partes reducidas del campo. Para solucionar este problema podrían utilizarse imágenes satelitales que cubran la totalidad de la zona a estudiar; sin embargo, estas son costosas y dependientes del clima ya que una nube podría tapar la vista de la cámara. Otra posible solución es tomar imágenes de toda la zona de cultivo; es decir, a una mayor altura para cubrir toda la zona; sin embargo, se perdería resolución de la imagen adquirida. Otra solución, y la cual fue desarrollada en este documento, es utilizar técnicas de procesamiento digital de imágenes para unir estas imágenes que cubren pequeñas porciones del campo de cultivo y de esta manera generar el panorama global de la zona de cultivo a evaluar. La técnica de unir estas imágenes se llama registro de imágenes y la ventaja de utilizar este tipo de enfoque es que permite resolver el problema de manera sencilla, permite obtener una buena resolución de la zona a evaluar y es independiente del clima. La única desventaja de la técnica es que requiere mucha capacidad de procesamiento lo cual lo vuelve lento cuando se arman panoramas de gran tamaño.

#### **1.4 Síntesis de la problemática**

Como ya se mencionó las técnicas actuales para el monitoreo de campos en la agricultura nacional son muy artesanales y producen resultados ineficientes; es así que existen diversas ventajas del uso de la agricultura de precisión, específicamente el monitoreo con imágenes aéreas; sin embargo el trabajar con imágenes aéreas también presenta varios problemas que serán resueltos mediante el uso de técnicas de procesamiento digital de imágenes con la finalidad de compensar dichas desventajas y así obtener una manera confiable y segura de trabajar con imágenes aéreas para el monitoreo de los campos de cultivo.



## Capítulo 2: Método para el registro y ortorectificación de imágenes aéreas

### 2.1 Registro de imágenes

Como ya se mencionó antes los campos de cultivo son de grandes extensiones por lo cual para obtener una imagen con buena definición es necesario tomar dichas imágenes a una distancia no tan alta, lo cual provoca que las imágenes sean de reducidas partes del campo. Es así que se requiere del registro de imágenes para poder procesar dichas imágenes y generar un panorama global del campo.

El *Registro de Imágenes* es el proceso de superponer dos imágenes de la misma escena tomada en diferentes tiempos, diferentes perspectivas o diferentes sensores. Esta técnica consiste en alinear geométricamente dos imágenes de las cuales una será considerada la imagen referencia y la otra será la imagen de muestra [4].

A continuación se describe el proceso básico para el registro de imágenes [5]. El cual puede ser dividido en las siguientes 5 etapas:

- I. **Pre procesamiento:** Este primer paso previo consiste en preparar las imágenes para la detección de puntos de control. Para ello se pueden utilizar métodos de ajuste de escala, segmentación y otros. Este paso previo se realiza con información previa; por ejemplo, si se sabe que la imagen de referencia es dos veces más grande que la imagen de muestra, se puede escalar dicha imagen al tamaño de la imagen referencia y así facilitar la correspondencia entre los puntos de control.
- II. **Detección de puntos de control (*features*):** Esta etapa consiste en la detección de puntos característicos de las imágenes y sus respectivos descriptores que son, básicamente, información sobre la orientación y magnitud de dicho punto de control. Para esto y dependiendo de la aplicación existen distintos métodos tales como el detector Harris-Laplace, SURF, SIFT entre otros.
- III. **Correspondencia entre los puntos de control:** En este paso, se realiza la comparación entre los puntos característicos (*features*) hallados en el paso II tanto para la imagen de referencia como la imagen de muestra. Este paso comparación es aplicado generalmente sobre los descriptores y el método más común para esto consiste en distancias euclidianas entre los descriptores.



- IV. Estimación del modelo de transformación:** En esta etapa, se estiman los parámetros para establecer la función que pueda alinear la imagen de muestra con la imagen referencia. Estos parámetros son estimados a partir de los puntos de correspondencia hallados en III.
- V. Transformación de la imagen:** Se realiza la transformación de la imagen de muestra a partir del modelo de transformación obtenido en IV. Esta transformación se realiza a nivel de los sistemas coordinados de la imagen de muestra y lo que se hace es transformar el sistema de coordenadas de la imagen de muestra al sistema de coordenadas de la imagen referencia. Una vez hecho esto se ubican los píxeles en el nuevo sistema coordinado y para ello se emplean métodos de interpolación para realizar de manera exacta la ubicación de dichos píxeles.

A continuación se presentan algunas de las aplicaciones del registro de imágenes:

- *Fusión de bandas*

Integra datos de diferentes bandas de frecuencia captados por diferentes sensores y en diferentes instantes. Esto permite desarrollar una sola vista con mayor información espacial y espectral.

- *Generación de mosaicos o panoramas:*

Integración de 1 o más imágenes de diferentes escenas pero con áreas en común con la finalidad de poder superponerlas y generar una sola imagen y que este sea panorama global de todas las imágenes.

- *Sensado remoto:*

Es la aplicación concerniente a este tema de tesis y consiste en la generación de mosaicos a partir de imágenes aéreas las cuales contarán con información multispectrales de los suelos para monitoreo o sensado de los campos de cultivo.

- *Imágenes médicas:*

Es aplicado en varios campos de la medicina como por ejemplo: el monitoreo de la evolución de un tumor ya que el registro de imágenes permite diferenciar cambios pequeños de una misma imagen pero tomada en diferentes instantes de tiempo. Además es útil en estudios como resonancias magnéticas, ultrasonidos entre otros ya que permite relacionar información obtenida desde diferentes sensores.

## 2.2 Ortorectificación de imágenes aéreas

El proceso de ortorectificación remueve la distorsión geométrica presente en las imágenes y la cual es ocasionada por la orientación de la cámara o el sensor, el desplazamiento debido al relieve y los errores sistemáticos asociados con la imagen. En este caso, se trabaja con campos de cultivos planos por lo cual la distorsión generada por el relieve será mínima y en lo que se pondrá especial énfasis es en la distorsión geométrica creada por la orientación de la cámara. Dicha distorsión es un problema ya que; por ejemplo, si se toma la imagen de un cuadrado como referencia y esta imagen es tomada con la cámara un poco inclinada, el cuadrado ya no se verá como cuadrado en la cámara sino como un trapecio (ver figura 5).



Figura 5: a) imagen con la cámara alineada, b) imagen con la cámara rotada o inclinada.

En el caso de la presente tesis, y como ya se mencionó en el capítulo 1, esta distorsión viene dada por el movimiento de la aeronave con la que se toman las fotos. A continuación se presenta una imagen que describe los 3 grados de libertad que puede presentar un avión y sus respectivos ejes que ayudan a dar una idea sobre cómo se medirán los ángulos de distorsión cuando el avión no se encuentre perfectamente alineado con el suelo.

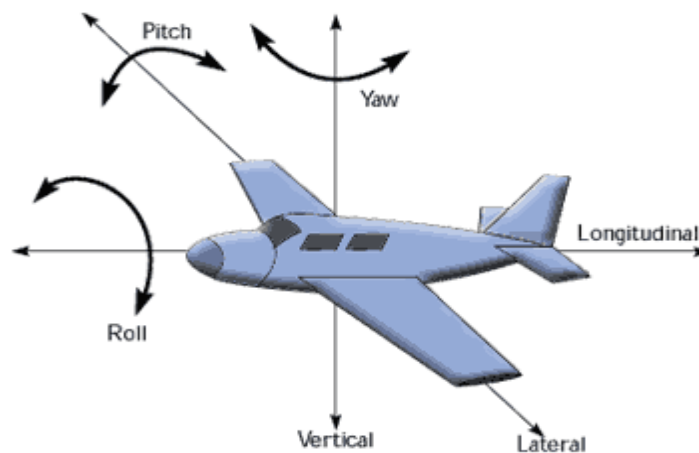


Figura 6: Ejes de rotación de una nave

Para entender cómo se realiza esta corrección de imágenes es importante dar unos alcances previos sobre la adquisición de las imágenes en la cámara usada así como también conceptos básicos de transformaciones de imágenes y en particular la transformación proyectiva que es utilizada para resolver este tipo de distorsiones generadas por las diferentes perspectivas generadas en el vuelo.

### 2.2.1 Firefly Mv

Cámara utilizada en la adquisición de imágenes, es la Firefly Mv [6] cuyas principales características se muestran a continuación:

- Cámara USB 2.0 CMOS monocromática
- Resolución de 640x480
- Captura imágenes con una frecuencia de 60 fotogramas por segundo (60FPS) [6].

Adicionalmente la cámara cuenta con un objetivo, que es el conjunto de lentes que sirven para controlar la entrada de luz y el foco de la imagen. El objetivo usado en la presente tesis es el modelo M0814-MP2 COMPUTAR [7] y cuya distancia focal es de 8 mm. Cabe resaltar que la distancia focal es el único parámetro utilizado para realizar la ortorectificación, los demás parámetros como apertura, ángulo de visión entre otros no serán tomados en cuenta.



Figura 7: Cámara con el objeto adherido.

### 2.2.2 Distancia focal

La distancia focal es la distancia que existe entre el centro óptico del lente y el foco, que es el punto por donde todos los rayos de luz confluyen. Esta distancia focal es empleada en las cámaras para enfocar imagen que se desea adquirir.

### 2.2.3 Cámara estenopeica

Es una cámara que no dispone de lentes, tiene un pequeño orificio por el cual entra la luz y dispone de un material fotosensible, en el cual se proyecta la imagen tomada. La distancia focal de esta cámara corresponde a la distancia que existe entre el agujero y el material fotosensible. En la figura 8 se observa este modelo de cámara.

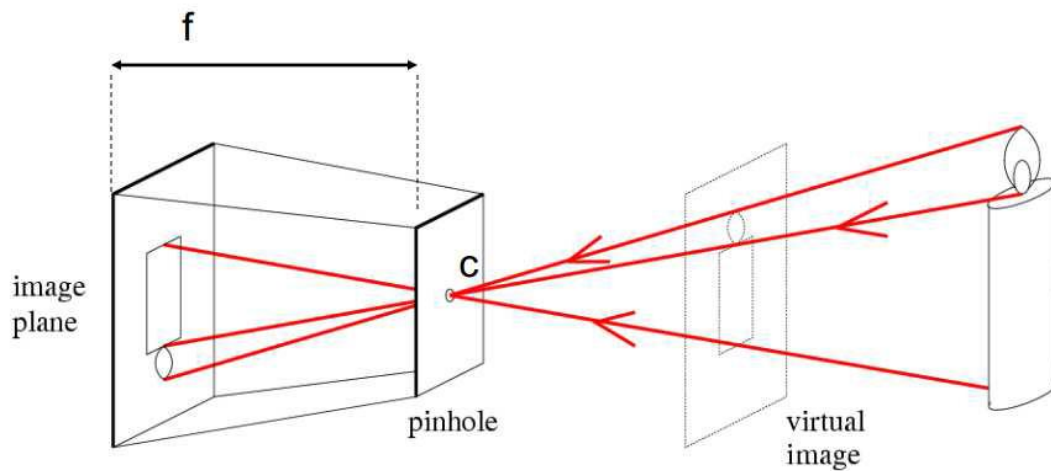


Figura 8: Diagrama de una cámara estenopeica [8]

### 2.2.4 Funciones de transformación

La transformación de imágenes consiste básicamente en modificar el valor de la posición de los píxeles de una imagen hacia otra locación dependiendo del tipo de transformación que se realice. Se pueden distinguir 2 tipos de transformación: las transformaciones lineales y las no lineales.

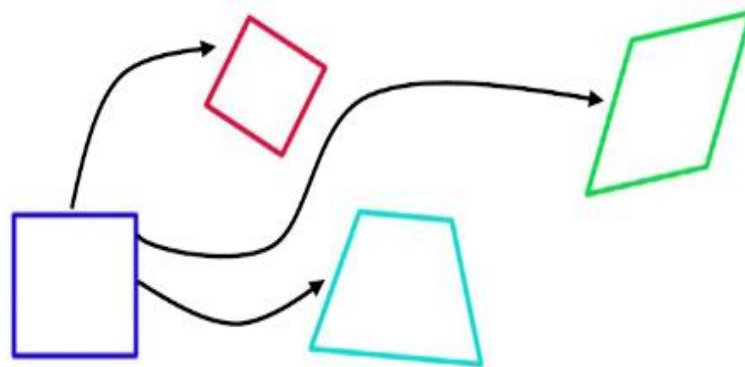


Figura 9. Transformaciones lineales. (Azul) imagen referencia. (Rojo) Similitud. (Verde) Afinidad. (Celeste) Proyectiva [9]

De la figura x se puede observar un comportamiento básico de cómo funcionan estas transformaciones lineales; a continuación se detallaran un poco más las transformación descritas en la imagen y en particular la transformación proyectiva que será clave tanto para la corrección geométrica de las imágenes como para la transformación de la imagen en la etapa del registro de imágenes.

- **Transformación afín [5]**

Este tipo de transformación, es una transformación lineal que se caracteriza por tener 6 grados de libertad ya que puede provocar los efectos de rotación, traslación, escala y transvección (shear) en dos dimensiones. La matriz de transformación afín cuenta con 6 variables las cuales están distribuidas de la siguiente forma (2.2.1)

$$H = \begin{bmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{bmatrix} \quad (2.2.1)$$

El tipo de efecto que genera esta matriz dependerá de los valores que tomen dichas variables; a continuación se detalla los posibles efectos que podrían obtenerse con esta función.

### Rotación

Para obtener un efecto de rotación los valores de la matriz de transformación H tendrían que ser los mostrados en (2.2.2) donde  $\theta$  es el ángulo a rotar la imagen.

$$H = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2.2.2)$$

### Traslación

Para obtener un efecto de traslación los valores de la matriz de transformación H tendrían que ser los mostrados en (2.2.3) donde dx y dy representan el traslado de la imagen tanto en el eje x como el eje y.

$$H = \begin{bmatrix} 1 & 0 & dx \\ 0 & 1 & dy \\ 0 & 0 & 1 \end{bmatrix} \quad (2.2.3)$$

### Shear horizontal

Para obtener un efecto de shear horizontal los valores de la matriz de transformación  $H$  tendrían que ser los mostrados en (2.2.4) donde  $\alpha$  representa el desplazamiento de los pixeles en la dirección horizontal

$$H = \begin{bmatrix} 1 & 0 & 0 \\ \alpha & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2.2.4)$$

### Shear vertical

Para obtener un efecto de shear vertical los valores de la matriz de transformación  $H$  tendrían que ser los mostrados en (2.2.5) donde  $\beta$  representa el desplazamiento de los pixeles en la dirección vertical.

$$H = \begin{bmatrix} 1 & \beta & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2.2.5)$$

En la siguiente figura se pueden observar algunos de los efectos obtenidos a partir de este tipo de matrices de transformación

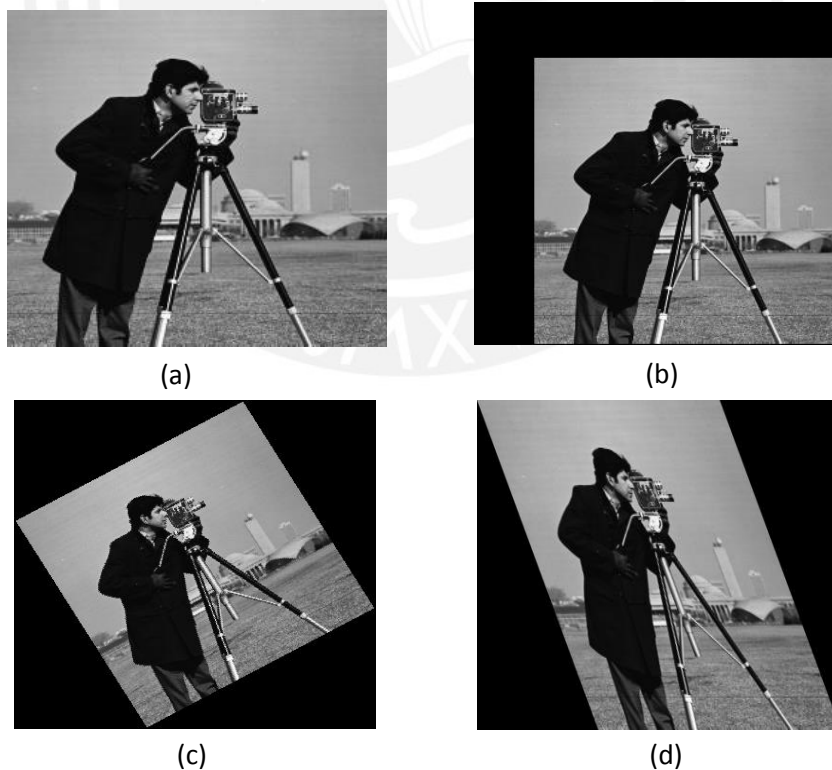


Figura 10: Imágenes con tipos de transformación afín. (a) Imagen original, (b) Imagen con translación, (c) Imagen con rotación y (d) Imagen con shear horizontal



- **Transformación de similitud**

Es un tipo de transformación lineal, similar a la afín, pero la cual solo cuenta con cuatro grados de libertad ya que puede generar los efectos de rotación, escalamiento y traslación en 2 dimensiones. La matriz que describe esta transformación cuenta con 6 elementos los cuales tienen la distribución mostrada en (2.2.6), donde  $s$  es el factor de escalamiento,  $\theta$  define la rotación de la imagen,  $t_x$  y  $t_y$  representan la traslación en sus respectivos ejes. Además esta matriz cuenta con la particularidad que su determinante es 1, esto permite conservar la forma de la imagen durante la transformación.

$$H = \begin{bmatrix} s \cos\theta & -s \sin\theta & t_x \\ s \sin\theta & s \cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix} \quad (2.2.6)$$

De la figura 11 se puede observar los efectos de usar la transformación de similitud, también se puede observar que a pesar de que la imagen haya sido, rotada trasladada y escalada, la forma de la imagen se mantiene ya que las flechas que son el contorno de la imagen se mantienen.

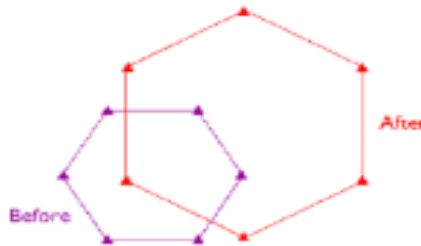


Figura 11: Transformación de similitud. Cambio de rotación, traslación y tamaño. [Referencia Roberto]

- **Transformación Projectiva**

Esta función es del tipo lineal y permite relacionar vistas de una región plana tomada desde diferentes ángulos y puntos de observación. La matriz de transformación para este caso cuenta con 8 variables y permite relacionar dos sistemas coordenados. LA ecuación que satisface esta función de transformación es la siguiente:

$$\begin{bmatrix} \mu(x,y) \\ v'(x,y) \\ w'(x,y) \end{bmatrix} = H \times \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}, \quad H = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{bmatrix}, \quad \mu(x,y) = \frac{\mu'(x,y)}{w'(x,y)}, v(x,y) = \frac{v'(x,y)}{w'(x,y)}$$

Ecuación del modelo de la transformación proyectiva (2.2.7)

Donde  $(x, y)$  representa el sistema coordenado de la imagen original y  $(u,v)$  representa el sistema coordenado de la imagen a proyectar. Del modelo de transformación descrito arriba (2.2.7), se pueden obtener relaciones útiles entre el sistema coordenado original y el sistema coordenado proyectado con la finalidad de poder obtener los valores de la matriz de transformación.

$$u = x.a + x.b + c + 0.d + 0.e + 0.f - u.x.g - u.v.h \quad (2.2.8)$$

$$v = 0.a + 0.b + 0.c + x.d + y.e + f - v.x.g - u.y.h \quad (2.2.9)$$

De las ecuaciones (2.2.8) y (2.2.9) se puede observar que para resolver el sistema de ecuaciones es necesario por lo menos 4 pares de puntos no colineares, los cuales nos ayudaran a obtener los parámetros de la matriz de transformación. En la presente tesis la transformación proyectiva se realizara en la etapa de ortorectificación y en el registro de imágenes, dichas etapas serán explicadas con mayor detalle en el capítulo 3.



Figura 12: Imagen con trasformación proyectiva

### **Capítulo 3: Diseño e implementación del sistema de ortorectificación y registro de imágenes**

#### **3.1 Objetivos**

##### **3.1.1 Objetivo principal**

Desarrollo e implementación de un algoritmo que permita la ortorectificación y registro de imágenes aéreas de campos de cultivo para el monitoreo y supervisión agrícola.

##### **3.1.2 Objetivos específicos**

1. Estudio y comprensión de los diferentes algoritmos utilizados en el registro y ortorectificación de imágenes.
2. Implementación del sistema de ortorectificación y registro de imágenes aéreas en MATLAB y lenguaje C.
3. Implementación de una medida global y específica para el registro de imágenes.

#### **3.2 Características del vuelo**

El vuelo realizado para obtener las imágenes del campo de cultivo es hecho de una manera automática con la finalidad de obtener imágenes de toda la zona sin dejar espacios vacíos. Dicho aeronave cuenta con una tarjeta electrónica (ArduPilot) y un software especializado (Mission Planner) el cual permite programar una ruta predeterminada (ver figura 13), lo cual asegurará la completa cobertura de la zona de cultivo, además se puede programar la velocidad de vuelo y la frecuencia a la cual se toman las imágenes para asegurar un buen traslape y también se cuenta con un sistema de adquisición de datos que registra los ángulos en los que son tomados las imágenes, estos datos son importantes ya que con la información de dichos ángulos se puede realizar la corrección geométrica de las imágenes y usarlo como paso previo del registro de imágenes para obtener un mejor registro.

Sea la figura 14, la zona que se planea volar, y sobre la cual se realizará una descripción de vuelo como en la figura 13, la cual consiste en una ida y vuelta de la nave de manera que abarque la zona designada. Del plan de vuelo y de las características mencionadas se puede observar que las fotos serán tomadas con un buen porcentaje de aérea común y una distribución ordenada tal como se muestra en la figura 14.



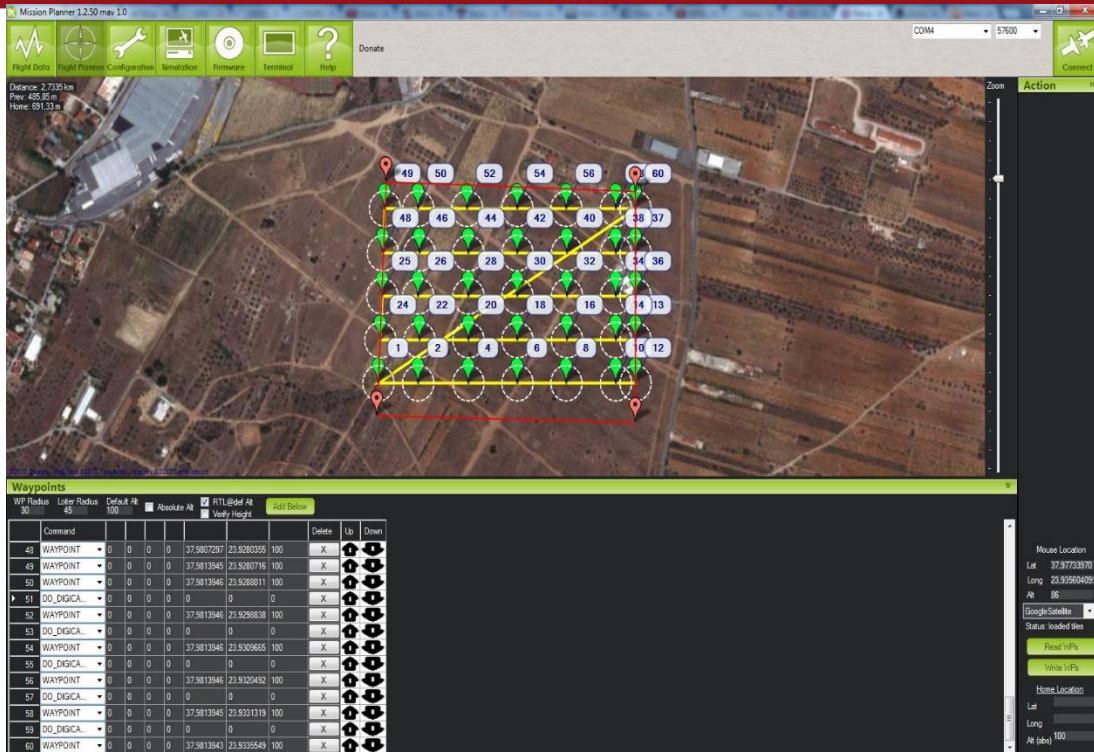


Figura 13: Plan de vuelo usando el Mission Planner. Las líneas amarillas representan la ruta que sigue el aeromodelo. [22]

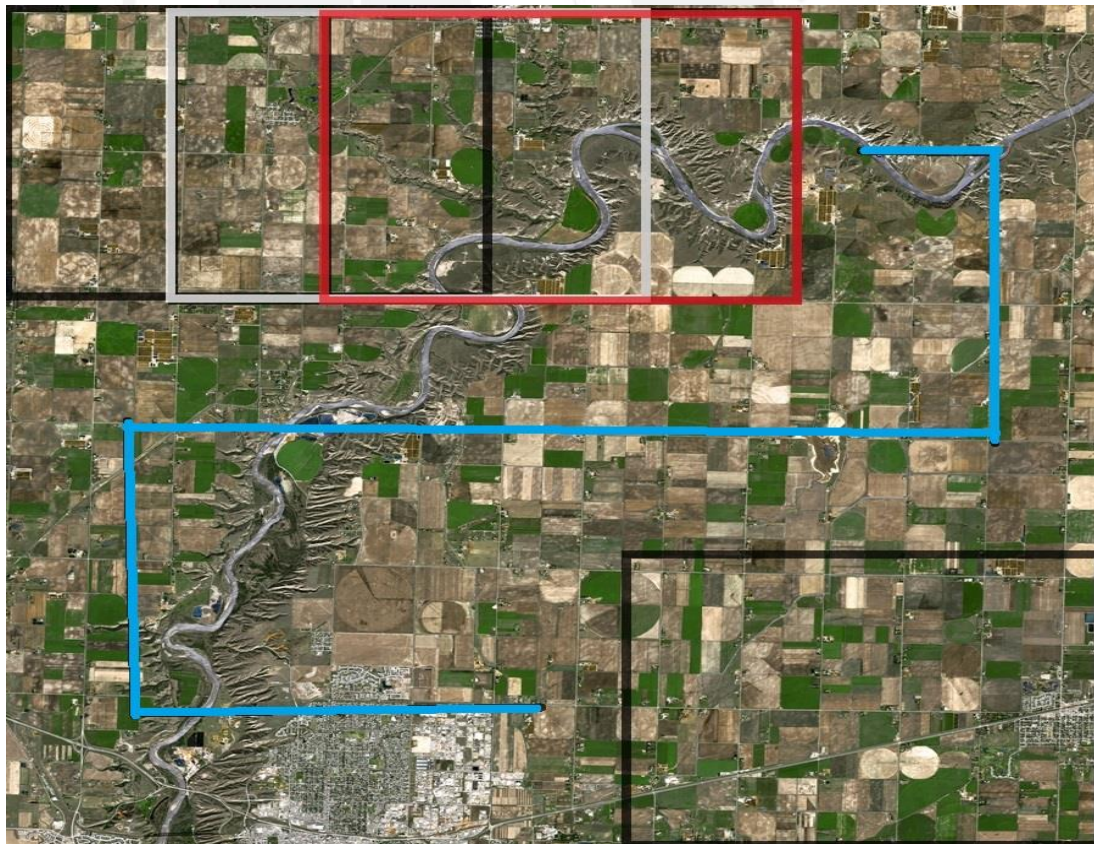


Figura 14: Cuadros de imágenes. Los cuadriláteros simbolizan las imágenes tomadas y la línea celeste la trayectoria del vuelo. La intersección de los cuadriláteros representa la aérea en común entre imágenes. [Alberta, Canada – Imagen satélite SPOT5]



El método usado en [10] para la generación de mosaicos consiste en tomar dos imágenes de las cuales una será la referencia y la otra será la imagen a superponer. Una vez generado ese mosaico, se toma dicho mosaico como imagen referencia y se superpone con una tercera imagen, luego se toma este mosaico generado como imagen referencia y se repite el proceso cuantas veces sea necesario. El déficit de este enfoque es que conforme el mosaico va creciendo el área en común con respecto a la nueva imagen a superponer decrece y se vuelve innecesario el hecho de computar todo el mosaico para unir una nueva imagen. Sin embargo, de la figura 14, se puede observar que la imagen obtenida en un instante dado siempre tendrá un área en común con la imagen tomada en el instante anterior. A partir de la característica del vuelo mencionada se propone construir un mosaico usando únicamente la relación obtenida con respecto a la imagen en el instante anterior y ya no usar de referencia el mosaico generado.

El enfoque propuesto, básicamente consistirá en utilizar de referencia la información de la imagen tomada un instante anterior, de esta manera se evitará procesar puntos innecesarios del resto del mosaico (ver figura 16), lo cual disminuirá la carga computacional. Cabe mencionar que el enfoque propuesto para el registro de imágenes solamente es factible para imágenes que están correctamente distribuidas y ordenadas ya que si las imágenes fueran tomadas en un orden aleatorio no habría correlación entre las mismas (imagen anterior y actual).

La ventaja del enfoque propuesto se muestra a continuación, sea el caso que se presenta en la figura 15, donde se tiene un mosaico de 15 imágenes y al cual le hace falta una imagen restante para completar la esquina vacía (espacio negro).



Figura 15: (Izq.) Mosaico de 15 imágenes con una esquina faltante. (Der.) Imagen de la esquina faltante en el mosaico

Mediante el enfoque propuesto en [10], se tendría que computar todo el mosaico en busca de puntos característicos para poder completar esa esquina faltante, mientras que con el enfoque propuesto en la presente tesis solo se buscarían puntos característicos en el espacio que tiene en común con la imagen anterior como se puede observar en la figura 16.

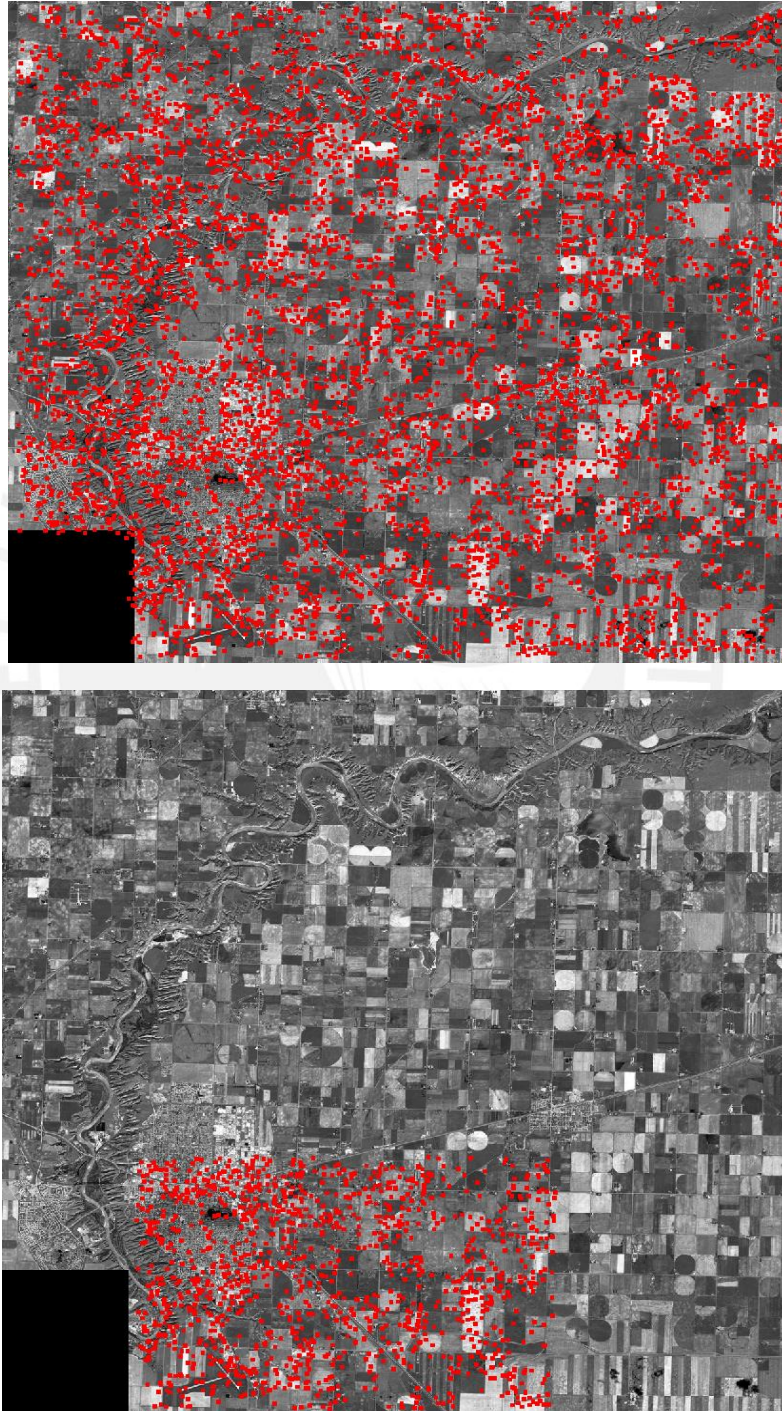


Figura 16: Relación de puntos evaluados en comparación con el mosaico total y utilizando la imagen anterior como referencia. (Arriba) Imagen con puntos evaluados (puntos rojos) en total para hallar la relación con la imagen que ocupa el espacio de la esquina. (Abajo) Imagen con puntos evaluados (puntos rojos) usando de referencia la imagen en el instante anterior.



### 3.3 Modelo del sistema

Para realizar el registro de imágenes se cuenta con una estructura básica (ver figura 17) que permite realizar el mosaico de 2 imágenes.

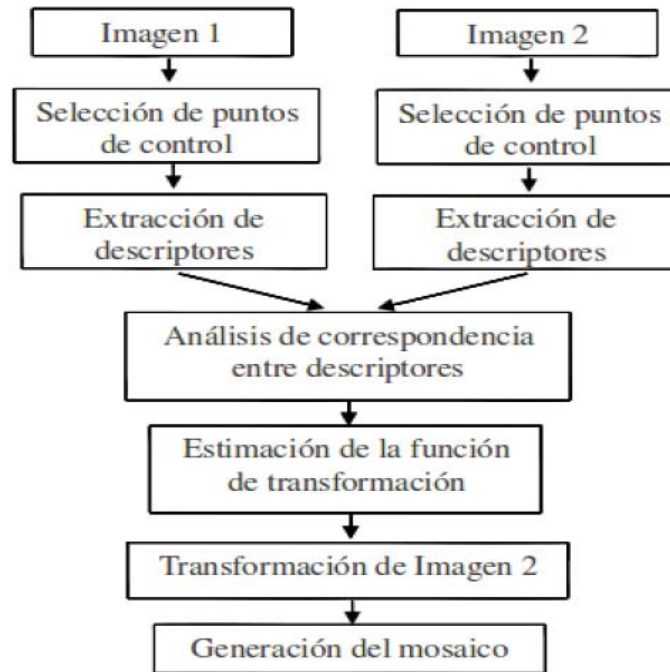


Figura 17: Estructura básica del registro de imágenes [10]

A partir de las características de vuelo mencionadas y del enfoque que se plantea utilizar, descrito en el punto anterior y partiendo de la estructura básica del registro de imágenes, se define el siguiente modelo para el sistema (ver figura 18).

De la figura 18, se puede observar que, a parte del enfoque que se plantea utilizar, se agregó un bloque previo al primer paso del registro de imágenes, el cual es la selección de puntos de control, con la finalidad de corregir la distorsión generada por los cambios de perspectiva en la imagen y además al hacer que la imagen este más alineada con la imagen referencia permitirá la facilidad en la correspondencia de los descriptores.

También se puede observar que el modelo propuesto es una repetición de la estructura básica del registro de imágenes en la cual la imagen referencia para la obtención de puntos deja de ser el mosaico y se convierte en la imagen previa debido a lo expuesto anteriormente. Cabe mencionar que según la lógica presentada la primera imagen obtenida se toma como referencia para la construcción del mosaico. En la sección 3.5 se describirán el funcionamiento de cada bloque de manera más detallada.

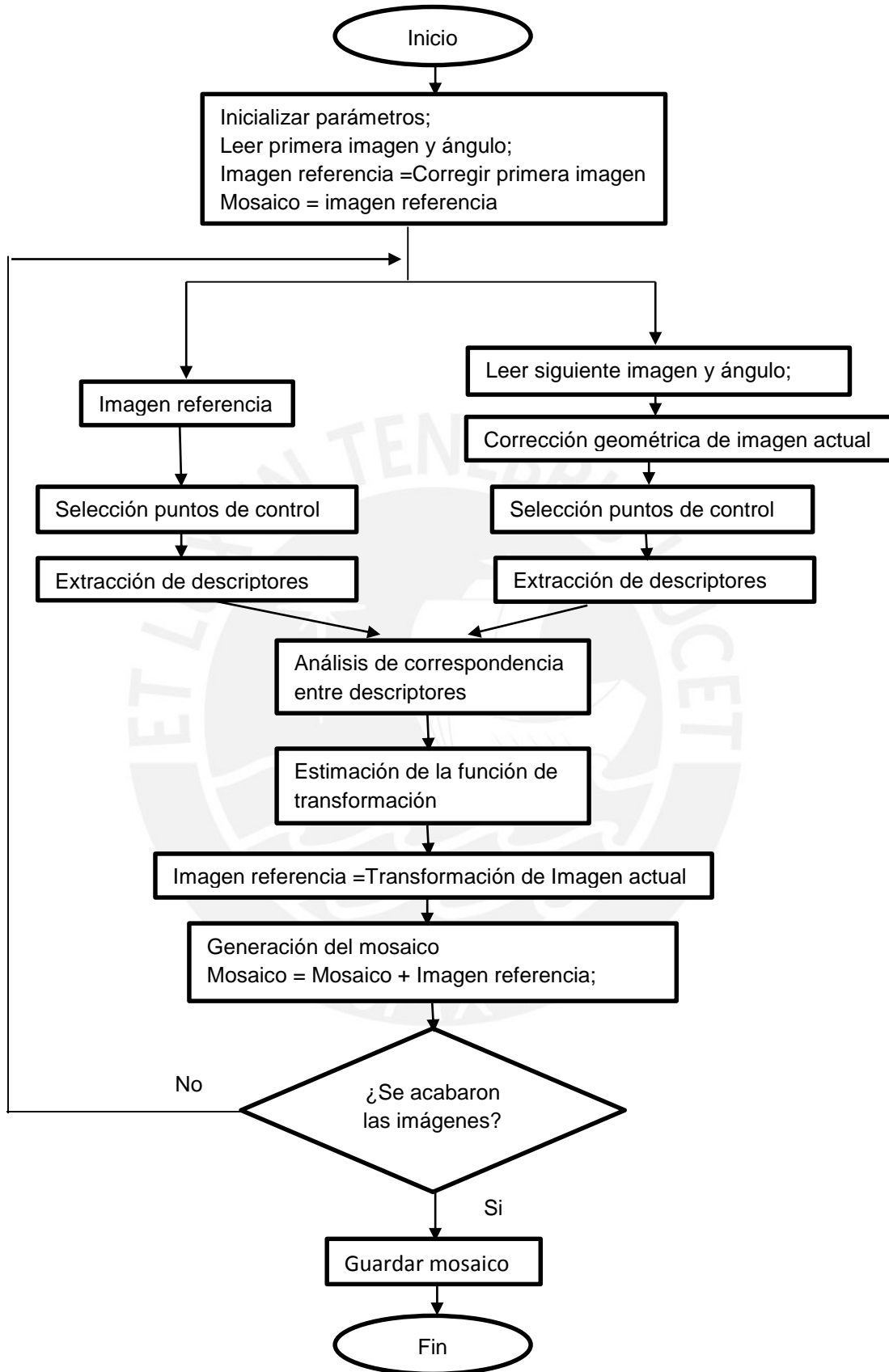


Figura 18: Diagrama de flujo del modelo propuesto

### 3.4 Selección de algoritmos

La selección de algoritmos para la etapa de registro de imágenes se hizo en base a la descripción de vuelo expuesto en el punto 3.2. Para el caso de la etapa de estimación de la función de transformación y correspondencia de puntos se utilizaron los algoritmos RANSAC y NNDR (Nearest Neighbor Distance Ratio) respectivamente, se optó por usar estos algoritmos ya que son métodos robustos y computacionalmente eficientes.

Por otro lado, para la etapa de selección de puntos de control y desarrollo de descriptores, la cual es la parte crítica del registro de imágenes, se utilizó el algoritmo SURF (Speed up robust Features) para el desarrollo de descriptores y dentro del cual se utiliza el algoritmo Fast-Hessian para la detección de puntos de control.

Se decidió optar por este algoritmo ya que tiene un menor costo computacional en comparación con un algoritmo del mismo tipo denominado SIFT (Scale-invariant feature transform) [11] y además el SURF presenta resultados similares a los de este algoritmo. La diferencia entre estos 2 algoritmos radica en el hecho de que el SIFT es capaz de detectar puntos invariantes ante efectos de rotación, traslación, escala y diferentes perspectivas, mientras que el SURF pierde efectividad antes cambios de perspectiva. Sin embargo, dado que se realizará el pre-procesamiento de las imágenes, el cual consistirá en la corrección de dichos cambios de perspectiva mediante el proceso de ortorectificación, se logrará obtener puntos confiables con el algoritmo SURF.

De esta manera nos queda que los algoritmos a utilizar para el registro de imágenes son los siguientes:

- Pre-Procesamiento: **Corrección Geométrica (Ortorectificación)**
- Selección de puntos de control: **Fast-Hessian**
- Extracción de descriptores: **SURF**
- Correspondencia de puntos de control: **NNDR**
- Estimación de función de transformación: **RANSAC**
- Modelo de Transformación establecido: **Transformación Proyectiva**
- Método de Interpolación : **Nearest Neighbor**

### 3.5 Descripción del funcionamiento del sistema

En esta sección del documento se darán los alcances de cómo funciona cada bloque descrito en el diagrama de flujo de la figura 18.

#### 3.5.1 Corrección de imágenes (Ortorectificación)

Para esta etapa se utilizó el algoritmo desarrollado en [8]. A continuación, se presentará el funcionamiento general de esta técnica. La ortorectificación consiste, como ya se explicó antes, en la corrección geométrica de las imágenes aéreas a partir de los ángulos a los que fueron tomadas dichas imágenes. Esta etapa se dividirá en 2 etapas: primero se realizará una corrección proyectiva proveniente del movimiento de alabeo y cabeceo y finalmente se corregirá la rotación generada por el movimiento de guiñada con una transformación afín.

- *Transformación proyectiva.*

En esta etapa, se utilizará la información de los ángulos de pitch y roll. Para esto se tendrá una imagen referencia, la cual estará completamente alineada al suelo (ver figura 19), y a partir de esta imagen se simulará el movimiento de la cámara en la dirección del ángulo respectivo ya sea pitch o roll; a partir de este movimiento se trazará el nuevo plano que visualizaría la cámara, una vez obtenido esto se proyecta el plano rotado al plano horizontal (ver figura 19).

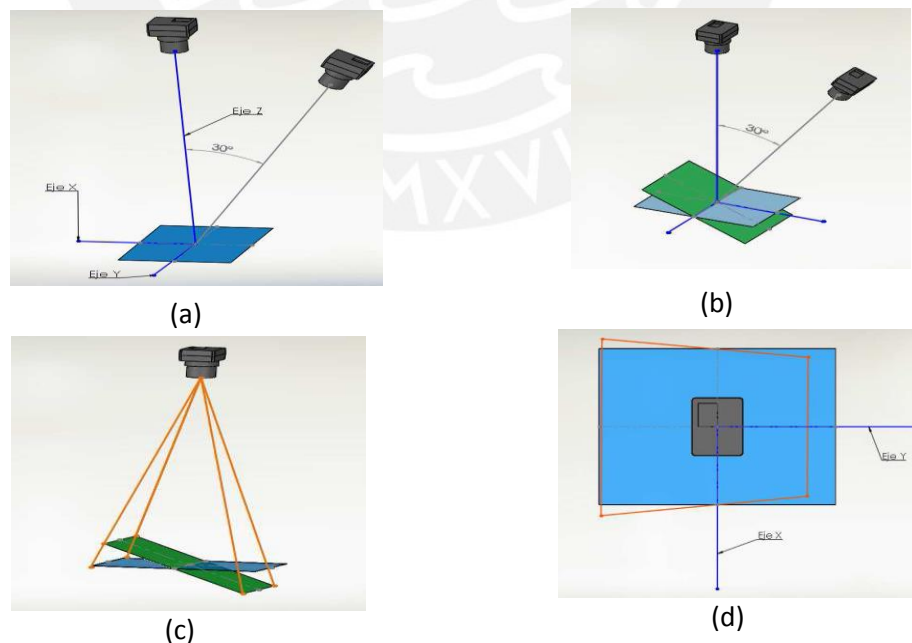


Figura 19: (a) Paso 1: proyección de un plano alineado al suelo. (b) Paso 2: proyección de un plano con respecto a la cámara rotada (plano verde). (c) Paso 3: proyección de la cámara rotada con respecto a la horizontal (líneas naranjas). (d) Paso 4: plano proyectado con respecto a la horizontal (plano naranja).

El proceso descrito líneas arriba se repite para el ángulo que falta corregir (pitch o roll). Una vez obtenida esta proyección se obtendrán los cuatro vértices que verdaderamente ocuparía la imagen y sobre la cual se realizará la corrección. Entonces lo que se hace con esta transformación es hacer que la imagen original (sin corrección) ocupe este nuevo plano proyectado.

- *Transformación afín*

Esta etapa es la más sencilla de la corrección, se basa simplemente en una corrección de rotación de la imagen a partir del ángulo yaw. Esta etapa se realiza luego de la transformación proyectiva y para ello se empleará la siguiente función de transformación afín de rotación (3.5.1) donde  $\theta$  indica el ángulo de rotación de la aeronave cuando la imagen fue tomada.

$$H = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3.5.1)$$

### 3.5.2 Selección de puntos de control [13]

La detección de puntos de control consiste en la detección de puntos invariantes en las imágenes. En este caso se utilizará el Fast-Hessian detector, el cual hace uso de la determinante de la matriz Hessiana (3.5.2). Sea  $X=(x, y)$  el punto en la imagen, la matriz Hessiana se define como  $H(x, y, \sigma)$  en el punto  $(x, y)$  y a escala  $\sigma$ .

$$H(x, y, \sigma) = \begin{bmatrix} L_{xx}(x, y, \sigma) & L_{xy}(x, y, \sigma) \\ L_{xy}(x, y, \sigma) & L_{yy}(x, y, \sigma) \end{bmatrix} \quad (3.5.2)$$

Donde los valores de  $L_{xx}$ ,  $L_{xy}$  y  $L_{yy}$  son los valores hallados mediante la convolución de la imagen con la segunda derivada gaussiana con desviación estándar de  $\sigma$ . Los kernels utilizados para hallar  $L_{xx}$ ,  $L_{xy}$  y  $L_{yy}$  se muestran en la figura 20. Sin embargo, para una mayor eficiencia computacional se utilizan aproximaciones de estos kernels (ver figura 20), a partir de esto se obtienen las aproximaciones  $D_{xx}$ ,  $D_{xy}$  y  $D_{yy}$ .

Una vez obtenido estos valores, se determina el valor de la determinante de la matriz Hessiana multiplicando por 0.9 los valores de las derivadas cruzadas (3.5.3). Dicha determinante representa la respuesta a detectar globos en la imagen alrededor del punto  $X$ .

$$\det(\mathcal{H}_{\text{approx}}) = D_{xx}D_{yy} - (0.9D_{xy})^2 \quad (3.5.3)$$

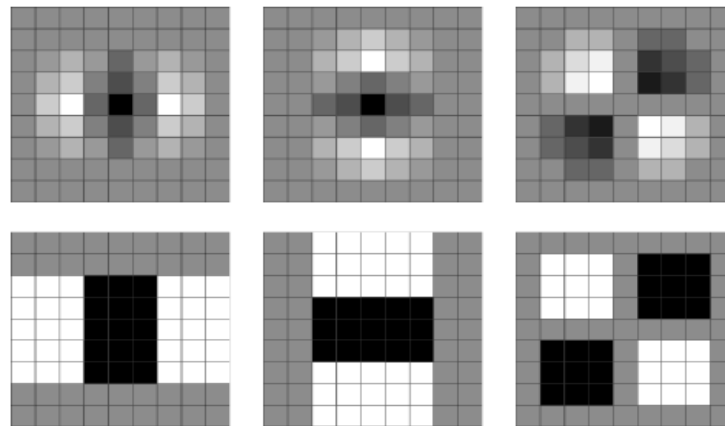


Figura 20: Kernels gaussianos para la obtención de la matriz Hessiana  
(Arriba) Segunda derivada gaussiana para obtener  $L_{xx}$ ,  $L_{xy}$  y  $L_{yy}$ .  
(Abajo) Aproximación utilizada para obtener  $D_{xx}$ ,  $D_{xy}$  y  $D_{yy}$ . [10]

La representación espacio-escala se implementa mediante pirámide de imágenes. La pirámide de imágenes consiste normalmente, como en SIFT, en filtrar la imagen con un filtro gaussiano y luego submuestrear la imagen y repetir el proceso. Cada una de estas repeticiones representa un nuevo nivel en la pirámide, sin embargo este algoritmo propone mantener la imagen a un tamaño constante y aumentar el tamaño de los filtros utilizados (ver figura 21). Donde cada nivel de la pirámide representa una octava y cada una de estas octavas se divide en niveles de escala.

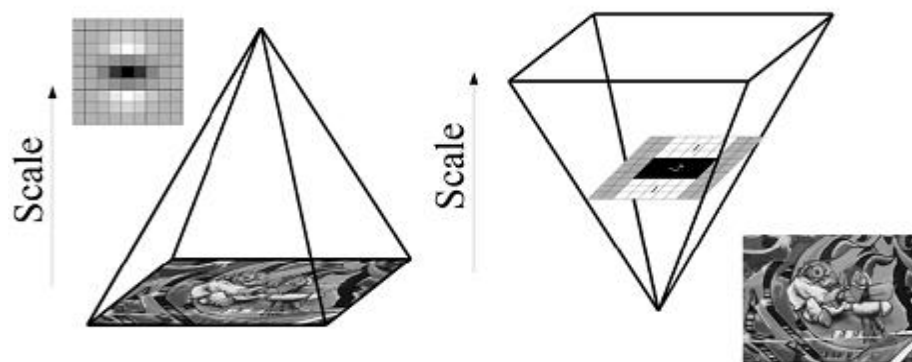


Figura 21: Representación espacio-escala. (Izq.) Enfoque SIFT variando tamaño de la imagen, (Der) Enfoque SURF variando tamaño del filtro. [13]

Para obtener los puntos de interés, primero se descartan todos los valores por debajo de un cierto umbral. Luego se eligen los valores máximos de la comparación de cada pixel con sus 26 vecinos, 8 de la misma capa, 9 de la capa superior y 9 de la capa inferior. Estas capas representan los niveles de escala descritos anteriormente y este proceso de detección de puntos se repite para todas las octavas.



### 3.5.3 Extracción de descriptores

**3.5.3.1 Asignación de orientación:** Una vez obtenido estos puntos de interés, el primer paso es asignar una orientación a cada punto de control con la finalidad de que el algoritmo sea invariante ante cambios de rotación. Para esto, primero se calcula la respuesta del wavelet Haar en la dirección x e y (ver figura 22) en una vecindad de radio  $6\sigma$  de la imagen (donde  $\sigma$  representa la escala en la que se encuentra el punto de interés), esta respuesta obtenida es promediada con un gaussiano centrado en el punto de interés y con un valor  $\sigma$  de 2,5. La orientación dominante se calcula a partir de la suma de todas las respuestas que se encuentren dentro de un sector circular de  $60^\circ$  alrededor del punto de interés (ver figura 23). El sector donde se encuentre el mayor vector será la orientación del descriptor [13].

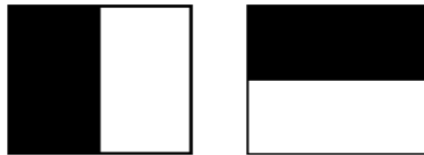


Figura 22: Haar wavelet para computar la respuesta en la dirección horizontal (Izq.) y en la dirección vertical (Der) [13]

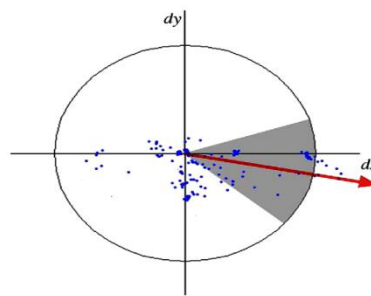


Figura 23: Asignación de orientación del vector dominante. Se toman los puntos dentro del sector circular sombreado [13].

**3.5.3.2 Extracción de componentes del descriptor:** Se genera un cuadrado de tamaño de  $20\sigma$  centrado alrededor del punto de interés y orientado a lo largo de la orientación del descriptor, este cuadrado es dividido en  $4 \times 4$  sub-regiones (ver figura 24). De cada una de estas subregiones se escogen 25 puntos regularmente espaciados y se computa la respuesta del wavelet Haar en sus componentes horizontal y vertical. A partir de estas respuestas se suman las componentes vertical y horizontal de cada punto así como también la sumatoria del valor absoluto de cada componente, obteniéndose un vector de la siguiente forma (3.5.4) donde dx y dy significan la respuesta al wavelet Haar en la dirección horizontal y vertical respectivamente.

$$v_{subregión} = \left[ \sum dx \quad \sum dy \quad \sum |dx| \quad \sum |dy| \right] \quad (3.5.4)$$

Por lo tanto como se tienen 16 subregiones en total y 4 componentes por cada subregión se tiene que cada descriptor tendrá 64 componentes. Este descriptor es invariante ante cambios de rotación y escala, sin embargo como ya se mencionó pierde efectividad ante cambios de perspectiva.

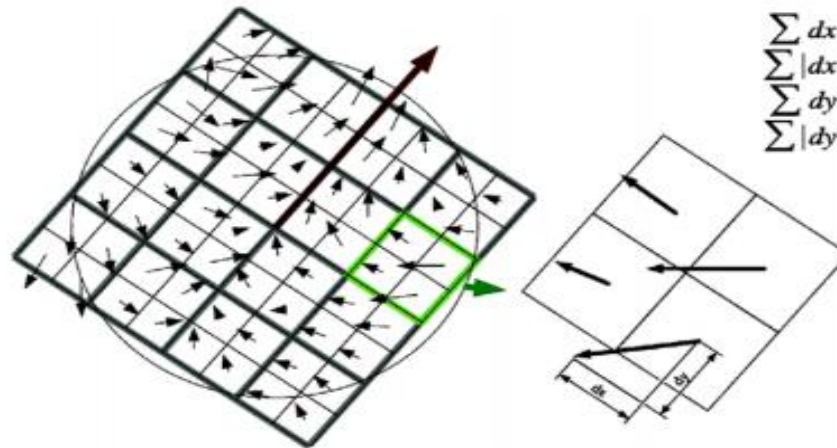


Figura 24: División del cuadrado formado en 16 regiones para obtener todos los componentes del descriptor [13]

#### 3.5.4 Correspondencia de puntos de control

Los puntos de control tendrán correspondencia si es que la distancia euclidiana entre su respectivos descriptores es la mínima de dicho conjunto y si es que la relación de dicha distancia con la segunda distancia mínima es menor a un determinado valor umbral [14]. La ecuación 3.5.5 representa la segunda condición descrita anteriormente, donde  $D_A$  es el descriptor a evaluar,  $D_C$  es la segunda mínima distancia y  $D_B$  es la nueva mínima distancia siempre y cuando cumpla la condición de que la relación entre dichas distancias sea menor al valor umbral (NNDR).

$$\frac{\|D_A - D_B\|}{\|D_A - D_C\|} \leq NNDR \quad (3.5.5)$$

#### 3.5.5 Estimación de la función de transformación

Para esta etapa se utiliza el algoritmo RANSAC (Random Sample Consensus), el cual es un método robusto para estimar los parámetros de la función de transformación. Este es un método iterativo el cual consiste en la estimación del modelo matemático que relaciona los pares de puntos de control encontrados en el paso anterior. Este método coge un grupo de 4 puntos aleatoriamente y calcula la matriz de homografía usando DLT (Direct Linear Transformation), luego se repite este proceso en base al número de iteraciones que se defina y se escoge la solución que incluya la mayor cantidad de puntos cercanos (inliers).

### 3.5.6 Transformación de la imagen y generación del mosaico

Para la realización de esta etapa se resuelve la ecuación 3.5.6. Donde H es la matriz de transformación obtenida del algoritmo RANSAC, (w, z) son las coordenadas horizontal y vertical respectivamente de la imagen sin transformar y (x, y) son las coordenadas horizontal y vertical de la imagen ya transformada. Una vez resuelta todas las ubicaciones de la imagen sin transformar, se ubica el valor de pixel trasladado al sistema coordenado transformado y se genera la imagen transformada.

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = H \begin{bmatrix} w \\ z \\ 1 \end{bmatrix} \quad (3.5.6)$$

Para la generación del mosaico también se parte de la matriz de transformación ya que de esta matriz se puede obtener los parámetros de traslación que indican donde se debe ubicar la imagen transformada con respecto a la imagen referencia. Una vez obtenidos dichos parámetros se superponen ambas imágenes y el valor de pixel que se superpone es el del vecino más cercano obtenido durante la transformación.

### 3.6 Funcionamiento del sistema con imágenes

Sean las imágenes de la figura 25; las imágenes a usar para generar el mosaico, donde la imagen de la izquierda será usada como la imagen referencia y la imagen de la derecha será la imagen a superponer/unir para formar el mosaico.

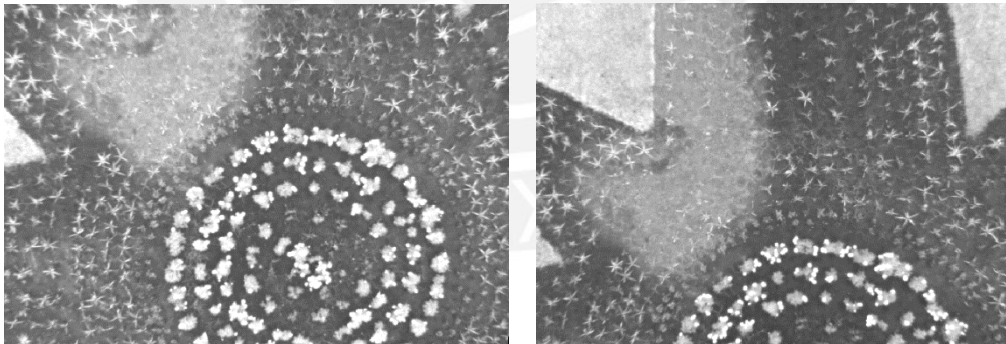


Figura 25: imágenes leídas. (Izq) imagen referencia. (Der) imagen a superponer

#### Paso1: Corrección geométrica de imágenes (Ortorectificación)

Una vez leídas las imágenes pasarán por el algoritmo de ortorectificación ya descrito antes y el cual nos dará un resultado equivalente al de la figura 26. Se puede apreciar un mejor alineamiento entre las imágenes y el efecto de la perspectiva generado por la cámara ha desaparecido. Los ángulos de corrección para la imagen de la izquierda son roll (r) = -21, yaw (y) = 0 y pitch (p)=1; para la imagen de la derecha sus ángulos son r=0, y=0 y p=-20.



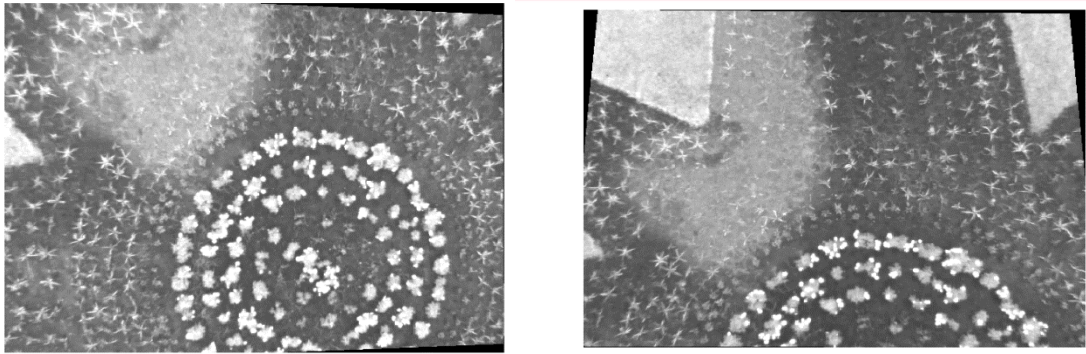


Figura 26: Imágenes corregidas

### Paso2: Selección de puntos de control

Del algoritmo Fast-Hessian desarrollado anteriormente, se obtendrá como resultado puntos característicos de ambas imágenes los cuales puede ser observado en la figura 27

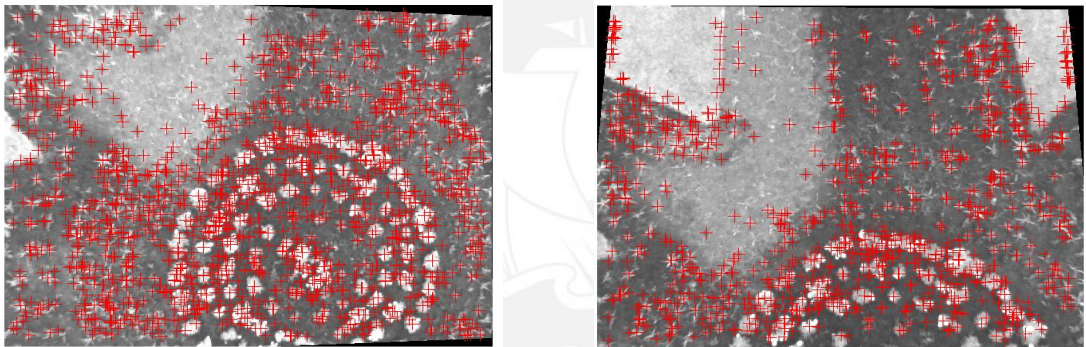


Figura 27: Imágenes con los puntos SURF. Los marcadores rojos indican donde se encuentran los puntos de control de cada imagen.

### Paso3: Desarrollo de descriptores

El algoritmo SURF descrito antes permite desarrollar descriptores de los puntos hallados en el paso previo. El resultado de este algoritmo se ve a continuación (ver figura 28).

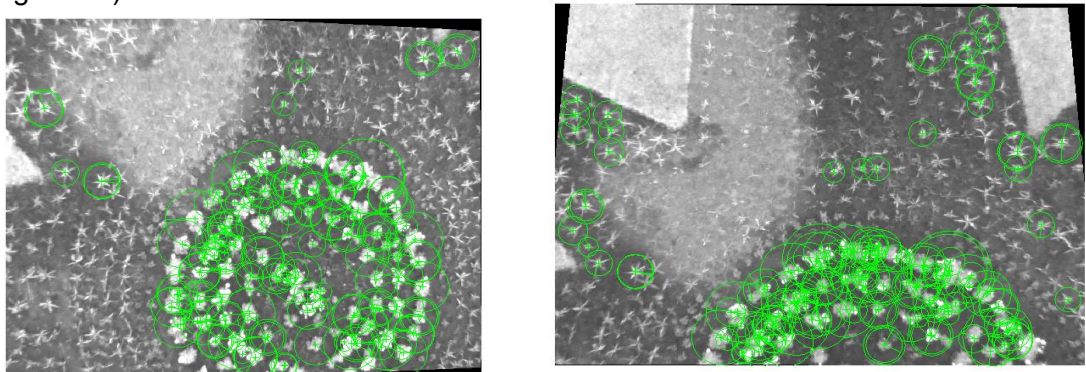


Figura 28: Imágenes con descriptor SURF



#### Paso4: Correspondencia de puntos de control

A partir de los puntos y descriptores hallados en los pasos previos, se realiza la correspondencia de los mismos, con la finalidad de hallar los puntos en común que hay entre ambas imágenes. De la figura 29, se puede observar esta correspondencia y se puede ver que los puntos se relacionan adecuadamente ya que los puntos hallados provienen de la zona superpuesta o en común que tienen ambas imágenes.

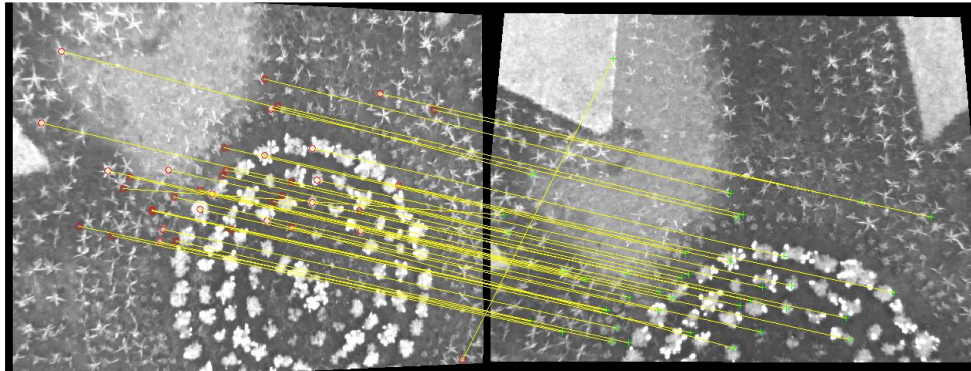


Figura 29: Correspondencia entre descriptores. Las líneas amarillas representan los puntos que corresponden de una imagen a la otra.

#### Paso5: Estimación de la función de transformación

A partir de los puntos en común encontrados en el paso4, se ejecuta el algoritmo RANSAC para estimar la función de transformación. Esta función es una matriz de 3\*3 la cual tiene la siguiente forma y valores para las imágenes que se están usando en este caso.

$$H = \begin{pmatrix} 0.8380 & -0.1232 & 7.9901 \\ -0.0828 & 0.6481 & 165.1467 \\ -0.0003 & -0.0004 & 1.0000 \end{pmatrix}$$

Matriz de transformación hallada

(3.5.7)

#### Paso 6: Transformación de la imagen y generación del mosaico

A partir de la matriz obtenida en el paso 5, se halla la imagen transformada. Para esto se resuelve el sistema de ecuaciones explicado anteriormente para las transformaciones del tipo proyectiva y finalmente se obtiene el resultado que puede ser observado en la figura 30. También a partir de esta matriz de transformación, se tiene la traslación de la imagen transformada lo cual nos permitirá ubicar exactamente donde se superpone con la imagen referencia, por lo cual una vez resuelta esta traslación se podrá unir ambas imágenes obteniendo el resultado de la figura 31.

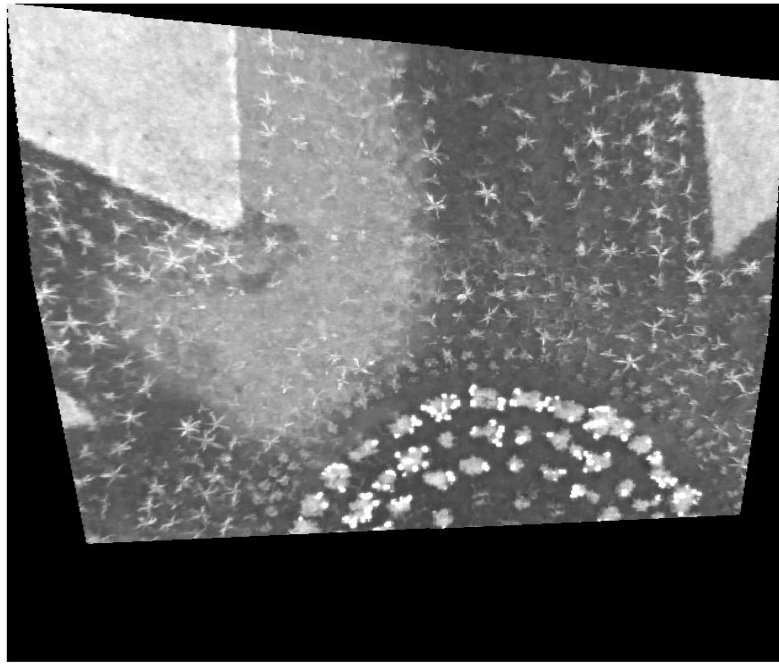


Figura 30: Imagen transformada. El espacio negro significa el traslado que sufre la imagen con respecto a la imagen referencia para formar el mosaico.

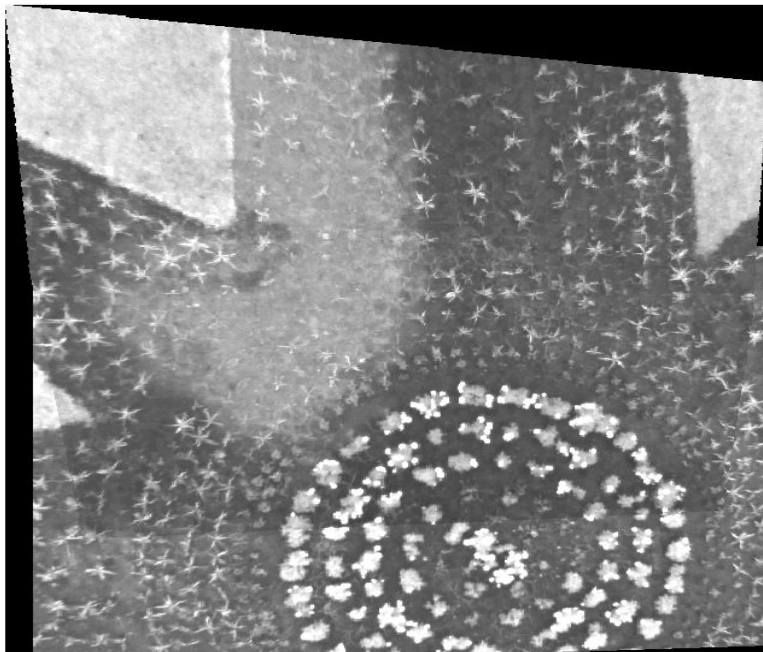


Figura 31: Mosaico generado a partir de las 2 imágenes iniciales. Cabe notar que hay un cambio de iluminación en la foto por lo que una parte del mosaico se ve más oscura que el resto. Sin embargo no afecta a la forma con la que se arma la imagen.

## **Capítulo 4: Ensayos y resultados**

La evaluación del modelo propuesto se dividirá en 2 partes. La primera parte consiste en la evaluación del algoritmo de generación de mosaicos utilizando un banco de imágenes de prueba [15] y las cuales servirán como referencia para saber que tan bien se está realizando el registro. La segunda parte consistirá en la evaluación del método propuesto utilizando imágenes aéreas.

### **4.1 Alcances del sistema**

- El algoritmo solo funciona para imágenes en escala de grises. Por lo cual, para las pruebas todas las imágenes se transformaron a escala de grises.
- Los parámetros utilizados para las pruebas fueron. NNDR = 0.8, número de iteraciones para RANSAC = 1000. El resto de parámetros son los utilizados por defectos en las distintas funciones empleadas. Los parámetros fueron elegidos de manera experimental.

### **4.2 Condiciones de pruebas con imágenes aéreas**

- Para la toma de imágenes se utilizó la aeronave de modelo DJIS800evo [19], el cual es un hexacóptero que cuenta con una tarjeta electrónica para el vuelo automático y distintos sensores como acelerómetros y giroscopios desde los cuales se obtuvieron los ángulos para el desarrollo de los algoritmos.
- Las imágenes aéreas de prueba son de resolución VGA y fueron obtenidas utilizando la cámara Firefly MV. El almacenamiento de las mismas se realizó utilizando una tarjeta electrónica llamada PCduino [20] en formato bmp.
- Para la programación del vuelo automático se utilizó el software misión planner, el cual fue mencionado en el capítulo 3.
- Los algoritmos se ejecutaron en la versión de Matlab 2013a en una PC con sistema operativo Windows 8 y con un procesador Intel core i5 de 3.1 GHz.
- La implementación en C fue utilizando la ayuda de las librerías de OpenCV y se ejecutaron en un sistema operativo GNU/Linux y con un procesador de las características mencionadas anteriormente
- Las imágenes aéreas de prueba usadas son del jardín de entrada de la PUCP. Para la obtención de las imágenes se programó un vuelo que cubriera toda la zona de la estrella que se encuentra en este jardín a una altura promedio de 50 metros. Se escogió esta zona ya que la figura de la estrella genera características resaltantes, con respecto al grass, que permitieron una mayor facilidad al momento de computar los descriptores.

### 4.3 Evaluación de los mosaicos

Se usaron 2 tipos de métricas en la formación de un mosaico. La primera es una métrica específica para evaluar el traslape entre imágenes del mosaico. La segunda métrica es una medida global de que tan bien se ha armado el panorama global del mosaico. Ambas métricas son explicadas a continuación.

La primera métrica, medirá la intensidad del total de pixeles que tienen en común las imágenes a traslapar y se tomará el valor absoluto de la diferencia entre dichos pixeles. Estas diferencias serán almacenadas en un vector y se realizará el promedio de las mismas (4.3.1), este parámetro nos indica el error promedio por pixel traslapado. Por ejemplo, si el error promedio fuera de 1 significa que hay una diferencia de intensidad de 1 por cada pixel traslapado. Hay que tomar en cuenta que se está tomando el valor absoluto de estas diferencias para evitar un error promedio igual a cero que nos daría una interpretación equivocada del error.

$$error\_promedio = \frac{\sum_1^{total\ pixeles} \text{diferencia intensidad por pixel}}{\# total de pixeles traslapados} \quad (4.3.1)$$

Este promedio será utilizado como umbral para definir el número de pixeles que se traslaparon correctamente; es decir, el número de pixeles cuya diferencia sea menor que el error promedio, se considerará un traslape correcto. Luego se hará una división entre los **pixeles correctos** y el **total de pixeles traslapados**, el cual nos dará una relación que significará el porcentaje de aciertos en el traslape (4.3.2).

$$\% \text{ aciertos} = \frac{\# \text{ pixeles correctos}}{\# total de pixeles traslapados} * 100\% \quad (4.3.2)$$

Para la segunda métrica, se evaluarán los descriptores del mosaico generado con respecto a una imagen referencia del mismo panorama pero que será adquirida desde una mayor altura. Para esto se relacionarán las distancias entre los puntos de control de ambas imágenes (el mosaico y la imagen referencia), esto debido a que las distancias entre dichos puntos deberían guardar una relación ya que como se puede ver en la figura 32 las distancias entre las esquinas de la estrella se mantienen en proporción. Por lo tanto, la métrica que se desarrolló consistió en computar los puntos de control de ambas imágenes y realizar la correspondencias de los mismos utilizando el método descrito anteriormente (NNDR = 0.6) para la generación del mosaico (ver figura 32).



Una vez obtenido los puntos correspondientes, se computa la distancia euclidiana de un mismo punto con respecto a los demás de la misma imagen, luego el siguiente punto con respecto a los que quedan y así sucesivamente (ver figura 32). Este proceso se repite hasta obtener todas las distancias posibles de estos puntos.

Estas distancias son almacenadas en dos vectores, uno para el mosaico y otro para la imagen referencia. Luego se realiza una división de elemento a elemento ya que cada elemento de los vectores representa la distancia entre los mismos puntos pero de las diferentes imágenes, por lo cual dicha división nos genera una relación entre todas las distancias de ambas imágenes. Se espera que dicha división tienda a ser una constante ya que si el mosaico se realizó de manera correcta se deberían conservar las características del panorama.

De este vector, al que llamaremos **relación de distancias**, se computará la media aritmética de estos valores y su desviación estándar. Se espera que la **media** sea un valor constante mayor a 1 y el cual nos indicará la **magnificación del mosaico** generado con respecto a la imagen referencia, se asume una magnificación porque la imagen referencia es de tamaño 640x480. La **desviación estándar** nos dará una idea de lo disperso que están estas distancias; es decir, cuanto error hay por cada punto mal ubicado. Idealmente se espera que esta desviación sea un valor pequeño cercano a cero; sin embargo, como todas las transformaciones geométricas son aproximaciones y la comparación se hará con una imagen de mucho menor tamaño se tomó en cuenta que para que el mosaico sea considerado correcto no debe sobrepasar de una desviación estándar correspondiente a 0,5.

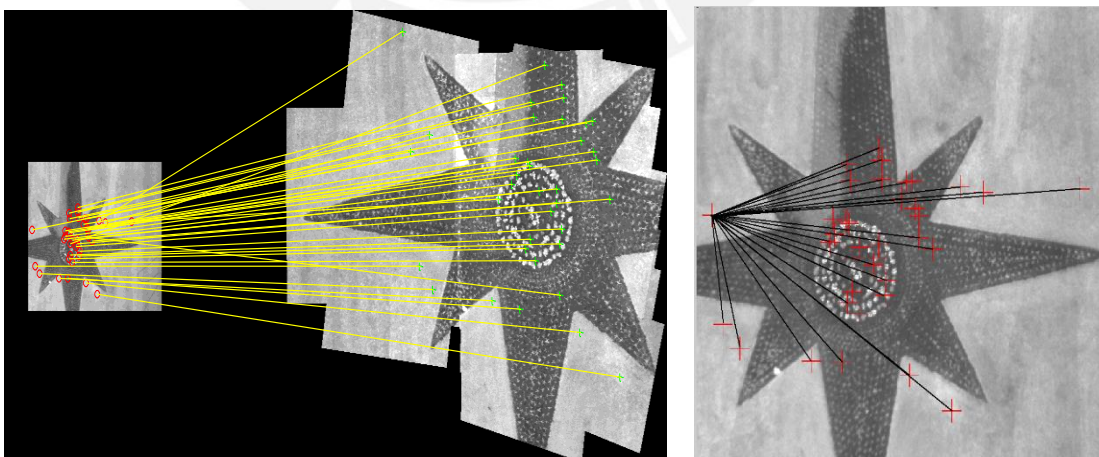


Figura 32: (Izq) Correspondencias entre imagen referencia y mosaico generado. Se puede apreciar la magnificación del mosaico generado. (Der) Ejemplo de la distancia de un punto con respecto a los demás (Líneas negras = distancias, cruces rojas = puntos de control)

#### 4.4 Resultados del banco de imágenes

Para la evaluación del algoritmo de generación de mosaico se utilizó un banco de imágenes de pruebas (ver figura 33), que cuentan con una matriz de transformación conocida y las cuales presentan diferentes efectos a evaluar. Se trabajó con 4 conjuntos de 6 imágenes cada uno y los diferentes efectos que presenta cada conjunto son cambios de desenfoque, cambios de perspectiva, cambios de iluminación y cambios de escala y rotación. De cada conjunto, la primera imagen, denotada como  $i_1$ , es la imagen referencia y las cinco imágenes restantes, denotadas como  $i_2$ ,  $i_3$ ,  $i_4$ ,  $i_5$  e  $i_6$ , son imágenes transformadas a partir de la primera imagen referencia y cuya matriz de transformación es de la forma (4.4.1). Las pruebas realizadas con este banco de imágenes son importantes ya que darán una confiabilidad al algoritmo de generación de mosaicos y se comprobará su eficiencia con respecto a los efectos mencionados. Este banco de imágenes puede ser encontrado en [15] y es un estándar de prueba en la literatura.

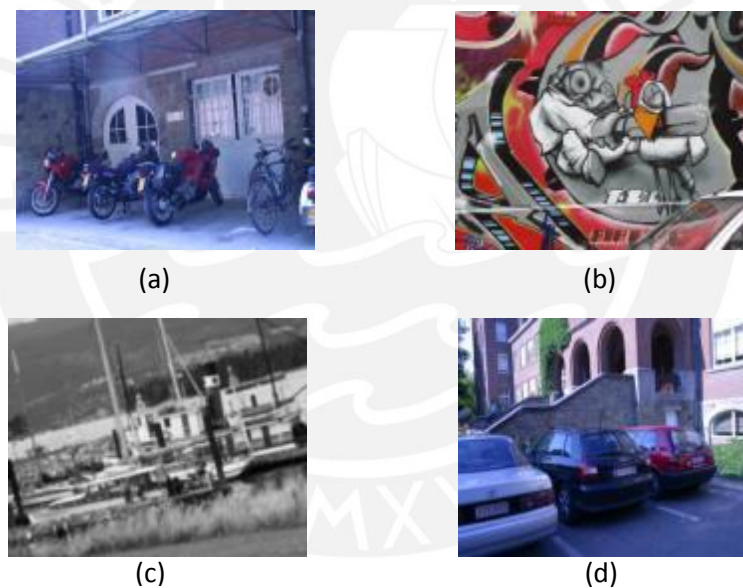


Figura 33: Banco de imágenes de prueba. (a) Imagen de prueba para cambios de desenfoque. (b) Imagen de prueba para cambios de punto de vista. (c) Imagen de prueba para cambios de escala y rotación. (d) Imagen de prueba para cambios de iluminación. [15]

$$H = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{pmatrix} \quad (4.4.1)$$

La evaluación consistirá en la diferencia de cada uno de los parámetros (4.4.1) de la matriz de transformación real con respecto a la matriz hallada. También se evaluarán parámetros importantes de cada imagen como número de puntos de control hallados, número de correspondencias halladas, porcentaje de aciertos, error promedio y tiempo de procesamiento. La notación  $i_1$ - $i_2$  corresponde al análisis entre las

imágenes 1 y 2, la notación i1-i3 corresponde al análisis entre las imágenes 1 y 3, y así sucesivamente. Los resultados obtenidos del programa realizado en C no se agregaron en este documento debido a que los resultados fueron similares a los obtenidos en Matlab; sin embargo, la tabla de resultados se encuentra en el anexo2.

#### 4.4.1 Cambios ante desenfoque

Para esta prueba los cambios en la matriz de transformación hallada con respecto a la real (*tabla1*) son mínimos por lo cual se puede afirmar que la transformación de la imagen se está haciendo de manera correcta y que el mosaico generado es correcto. Se puede observar que conforme va aumentando el grado de desenfoque con cada imagen, el número de correspondencias disminuye (*tabla 2*), lo cual tiene sentido ya que los cambios de intensidades entre imagen e imagen son muy notorios; sin embargo se mantienen las características de la imagen que es lo que el algoritmo detecta y con esa información puede obtener una matriz de transformación cercana a la real. A pesar de que el mosaico final se torne borroso no significa que se esté transformando mal sino que al realizarse una superposición se arrastra error de la imagen borrosa.

Tabla1: Diferencia entre los parámetros de la matriz hallada y la matriz real

	I1-I2	I1-I3	I1-I4	I1-I5	I1-I6
a	-0.029 E-01	-0.028 E-01	-0.026 E-01	0.048 E-01	0.103 E 01
b	-0.019 E-01	-0.026 E-01	-0.015 E-01	-4.693 E-04	0.150 E-01
c	0.1319 E01	0.1552 E01	7.957 E01	-7.941 E-01	-1.465 E-01
d	-8.542 E-04	-5.442 E-04	-4.158 E-04	0.013 E-01	-0.090 E-01
e	-0.028 E-01	-0.030 E-01	-0.038 E-01	0.055 E-01	0.103 E 01
f	9.641 E-01	8.519 E-01	9.769 E-01	-1.966 E-01	-4.586 E-01
g	-1.348 E-06	-1.182 E-06	-7.533 E-04	4.685 E-06	-4.062 E-06
h	-2.514 E-06	-3.272 E-06	-3.775 E-06	1.230 E-06	1.553 E-05

Tabla2: Información del funcionamiento del algoritmo ante cambios de enfoque

	Descriptores		Descriptor en común	% de Aciertos	Error promedio	Tiempo de procesamiento
	Imagen 1	Imagen 2				
I1-I2	1414	993	544	70.60	5.24	3.5
I1-I3	1414	869	411	71.84	6.03	3.6
I1-I4	1414	640	230	69.01	7.54	3.7
I1-I5	1414	515	161	68.78	8.52	4.1
I1-I6	1414	385	101	68.54	9.47	4.5

#### 4.4.2 Cambios de escala y rotación

Al igual que en la prueba anterior, se puede observar que los cambios en la matriz de transformación hallada con respecto a la real (*tabla 3*) son mínimos por lo cual se puede afirmar que la transformación de la imagen se está haciendo de manera correcta y que el mosaico generado es correcto. Se puede observar que conforme va aumentando el efecto de rotación y escala con cada imagen, el número de correspondencias disminuye (*tabla 4*), lo cual tiene sentido ya que ocurren cambios en la distribución de las intensidades de la imagen; sin embargo se mantienen las características de la imagen que es lo que el algoritmo detecta y con esa información puede obtener una matriz de transformación cercana a la real. También se puede observar que el tiempo de procesamiento aumenta a pesar que se computen menos descriptores de correspondencia, esto es debido a que durante la generación del mosaico la imagen transformada aumenta de tamaño por lo cual se computa una imagen de casi el doble de dimensiones, lo cual genera un costo extra computacionalmente hablando.

Tabla3: Diferencia entre los parámetros de la matriz hallada y la matriz real

	I1-I2	I1-I3	I1-I4	I1-I5	I1-I6
a	-8.901 E-05	0.072 E-01	-0.053 E-01	0.114 E-01	-0.083 E-01
b	-0.020 E-01	0.111 E-01	-0.124 E-01	0.154 E-01	0.215 E-01
c	6.267 E-01	-0.332 E-01	0.145 E 01	-0.446 E 01	-0.389 E01
d	-1.164 E-05	-0.015 E-01	0.014 E-01	0.052 E-01	-0.503 E-01
e	-0.025 E-01	0.103 E-01	-0.097 E-01	0.171 E-01	0.671 E-01
f	3.205 E-01	-8.847 E-01	-0.151 E01	-0.333 E 01	7.359 E-01
g	1.042 E-06	6.038 E-06	-9.525 E-06	1.222 E-05	5.169 E-05
h	-2.743 E-06	1.604 E-05	-1.793 E-05	2.878 E-05	7.654 E-05

Tabla4: Información del funcionamiento del algoritmo ante cambios de escala y rotación

	Descriptores		Descriptor en común	% de Aciertos	Error promedio	Tiempo de procesamiento
	Imagen 1	Imagen 2				
I1-I2	4004	3836	796	70.63	14.4	14.3
I1-I3	4004	2634	423	70.77	13.6	12.9
I1-I4	4004	1892	281	65.77	23.9	13.4
I1-I5	4004	1747	194	67.26	22.8	15.5
I1-I6	4004	1269	51	65.16	45.21	31.5



#### 4.4.3 Cambios de perspectiva

Para los cambios de perspectiva hasta la imagen 4 se tienen resultados cercanos a la matriz real, esto se puede apreciar en la *tabla5* en la cual para la generación del mosaico I1-I5 e I1-I6 se puede observar cambios en los valores de la matriz mayores a 200 lo cual nos arrastrará un error muy grande al momento de la transformación. Es por esta razón que entre más pronunciado es el cambio de perspectiva, más problemas tienen el detector para hallar puntos de control útiles, por lo cual la cantidad de descriptores en común va decreciendo rápidamente y además esta correspondencia no es acertada, esto impide hallar una matriz de transformación adecuada ya que no se cuentan con suficientes puntos confiables (inliers). Los valores de tiempo que no se ponen en la tabla 6 son debido a que el programa no termino de armar el mosaico ya que la matriz de transformación estaba muy lejos de la real.

Tabla5: Diferencia entre los parámetros de la matriz hallada y la matriz real

	I1-I2	I1-I3	I1-I4	I1-I5	I1-I6
a	-0.090 E-01	2.461 E-05	-0.063 E-01	5.585 E 00	0.478 E 00
b	-0.069 E-01	-0.046 E-01	-0.432 E-01	-0.124 E00	0.728 E-01
c	0.287 E 01	0.141 E 01	0.819 E 01	-260.1 E00	215.5 E00
d	-0.032 E-01	0.019 E-01	0.166 E-01	6.904 E00	0.288 E00
e	-0.095 E-01	0.022 E-01	-0.347 E-01	0.206 E00	2.995 E00
f	0.105 E 01	-0.157 E 01	-0.313 E 01	-400.9 E00	-545.1 E00
g	-8.964 E-06	8.350 E-06	-1.836 E-06	0.0169 E00	2.082 E-04
h	-1.153 E-05	-7.518 E-06	-5.414 E-05	-0.026 E-01	0.036 E-01

Tabla6: Información del funcionamiento del algoritmo ante cambios de perspectiva

	Descriptores		Descriptor en común	% de Aciertos	Error promedio	Tiempo de procesamiento
	Imagen 1	Imagen 2				
I1-I2	2065	2397	398	79.33	13.5	9.9
I1-I3	2065	2503	195	76.09	21	10
I1-I4	2065	2209	77	64.15	40.1	11
I1-I5	2065	2551	67	-	-	-
I1-I6	2065	1955	47	-	-	-

Nota: Los mosaicos generados a partir del banco de imágenes podrán ser ubicados en el anexo 1.

#### 4.4.4 Cambios de iluminación

Al igual que el resto de pruebas, se obtuvo un resultado cercano de la matriz de transformación hallada con respecto a la original (Tabla7). Y conforme se va acentuando el efecto de iluminación (más oscura) se disminuye el número de descriptores y correspondencias pero eso es lo esperado; sin embargo igual se encuentran puntos que definen las características propias de la imagen lo cual es suficiente para obtener una matriz de transformación correcta. A pesar de que el mosaico final se torne más oscuro no significa que se esté transformando mal sino que al realizarse la superposición se coloca la imagen más oscura encima generando esta tonalidad más oscura con respecto a la original. Se puede ver que las métricas de porcentaje de acierto y error promedio no presentan resultados coherentes a pesar de estimarse bien la matriz de transformación, esto debido a que ocurre una diferencia de intensidad en todos los píxeles. Es por esta razón que mientras más decrece el grado de iluminación mayor es el error promedio y por ende el porcentaje de aciertos disminuye a casi el 50 % de aciertos.

Tabla7: Diferencia entre los parámetros de la matriz hallada y la matriz real

	I1-I2	I1-I3	I1-I4	I1-I5	I1-I6
a	-4.212 E-01	-0.425 E 00	-4.228 E-01	-1.580 E 00	-0.421 E00
b	1.507 E-04	-0.024 E-01	0.061 E-01	-0.173 E-01	-0.091 E-01
c	-0.192 E 01	-1.905 E00	-4.143 E 00	-6.268 E-01	0.485 E00
d	-0.016 E-01	3.889 E-04	-9.448 E-04	0.025 E-01	-0.014 E-01
e	-4.221 E-01	-4.261 E-01	-0.423 E 00	-1.587 E00	-4.250 E-01
f	0.131 E 01	1.788 E00	3.720 E00	12.16 E00	7.385 E00
g	1.257 E-06	2.036 E-06	2.547 E-06	5.743 E-06	6.113 E-07
h	-2.073 E-06	-4.091 E-06	2.586 E-06	-2.577 E-05	-1.804 E-05

Tabla8: Información del funcionamiento del algoritmo ante cambios de iluminación

	Descriptores		Descriptor en común	% de Aciertos	Error promedio	Tiempo de procesamiento
	Imagen 1	Imagen 2				
I1-I2	1566	1135	564	46.58	29.72	3.1
I1-I3	1566	906	365	48.15	41.8	3.2
I1-I4	1566	763	265	51.4	51.41	3.1
I1-I5	1566	620	171	52.8	59.66	2.99
I1-I6	1566	499	105	55.5	66.01	3.01

#### 4.5 Discusión de resultados de imágenes de prueba

De los resultados obtenidos del banco de imágenes de prueba se puede observar que el detector y descriptor empleado tienen un buen comportamiento ante cambios de enfoque, luminosidad, escala y rotación. Sin embargo, pierde efectividad ante cambios de perspectiva pronunciados. Dado que el algoritmo diseñado contará con una etapa de corrección de perspectiva previa antes de la generación del mosaico, se puede afirmar que el registro de imágenes debería ser realizado de una manera más adecuada.

Adicionalmente se observa que la generación de mosaico ante cambios de perspectiva y rotación son las que más tiempo demoran. Esto debido a un mayor tamaño de la imagen a transformar, la complejidad al momento de estimar la función así como también el hecho de que la imagen a transformar no este alineada con la imagen referencia, lo cual genera un mayor tiempo de procesamiento para la correspondencia y estimación de la función de transformación.

Se observó que las métricas propuestas dan resultados coherentes para todas las pruebas. Cabe mencionar que para la prueba de iluminación, toda la imagen a superponer se torna más oscura generando así un error muy alto por pixel superpuesto; sin embargo, esto era de esperarse, por lo cual el porcentaje de aciertos es 50% en lugar de cero ya que esta diferencia de intensidades es tomado en cuenta al calcular el umbral. Sin embargo, lo que se espera es que el error promedio sea pequeño, idealmente cero, por lo cual una de las consideraciones para las imágenes aéreas es que las tomas de las imágenes sean bajo una misma iluminación. En el caso de la prueba de escala y rotación, así como la de cambios de perspectiva, el error promedio se incrementó de manera abrupta debido a que la imagen a traslapar presenta información o partes de imágenes que no contiene la imagen referencia, por lo cual a pesar que la transformación se realice de manera correcta habrán incongruencias entre las mismas imágenes generando un error por pixel alto.

#### 4.6 Análisis de imágenes aéreas

##### 4.6.1 Análisis por pares de imágenes

Dado que los errores en la generación del mosaico son bastante pequeños se puede afirmar que este algoritmo funciona de una manera correcta. En esta etapa se agregará la etapa de corrección de las imágenes aéreas y se formará un mosaico con las mismas. La evaluación de este mosaico será utilizando un criterio que se basa en la diferencia de intensidad de los pixeles superpuestos o en común que

tengan las imágenes que forman el mosaico. Así mismo se evaluará la ventaja, si es que la hubiera, de usar la ortorectificación de imágenes.

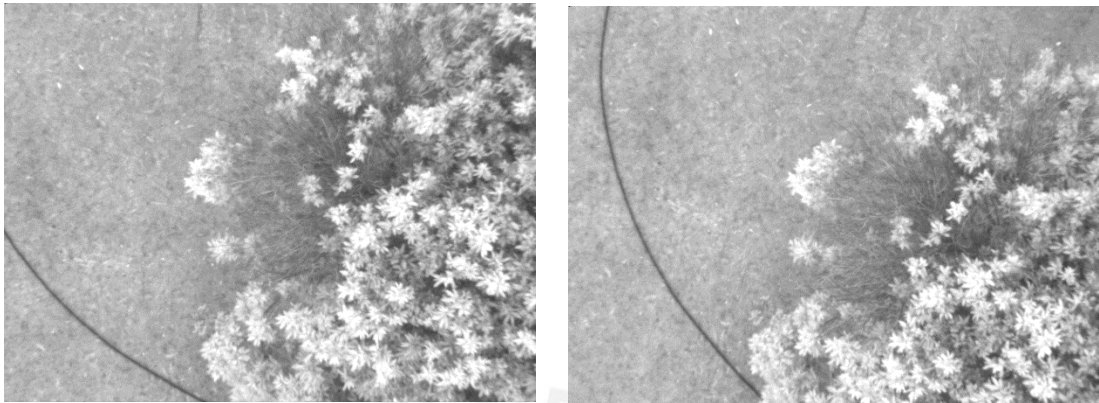


Figura 34: (Izq.) Imagen original usada como referencia. Ángulos de la imagen yaw= 0, pitch =4, roll =2. (Der) imagen original a superponer. Ángulos de la imagen yaw= 331, pitch = 5 roll=2.

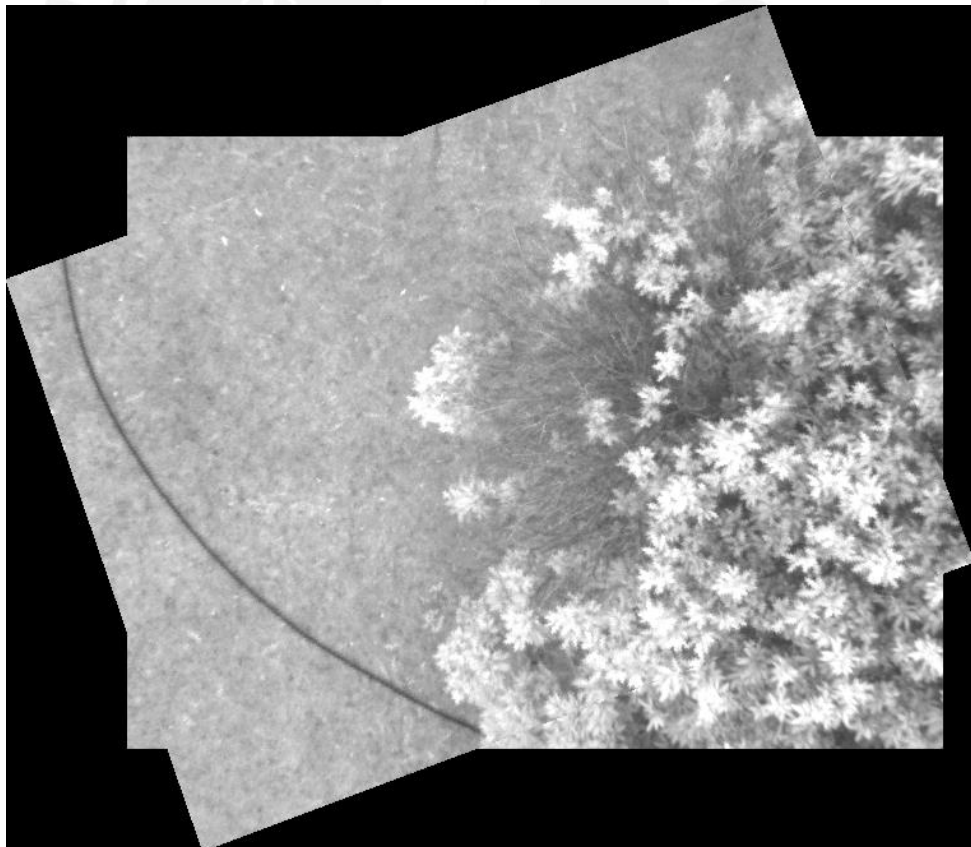


Figura 35: Mosaico generado a partir de las imágenes sin corrección



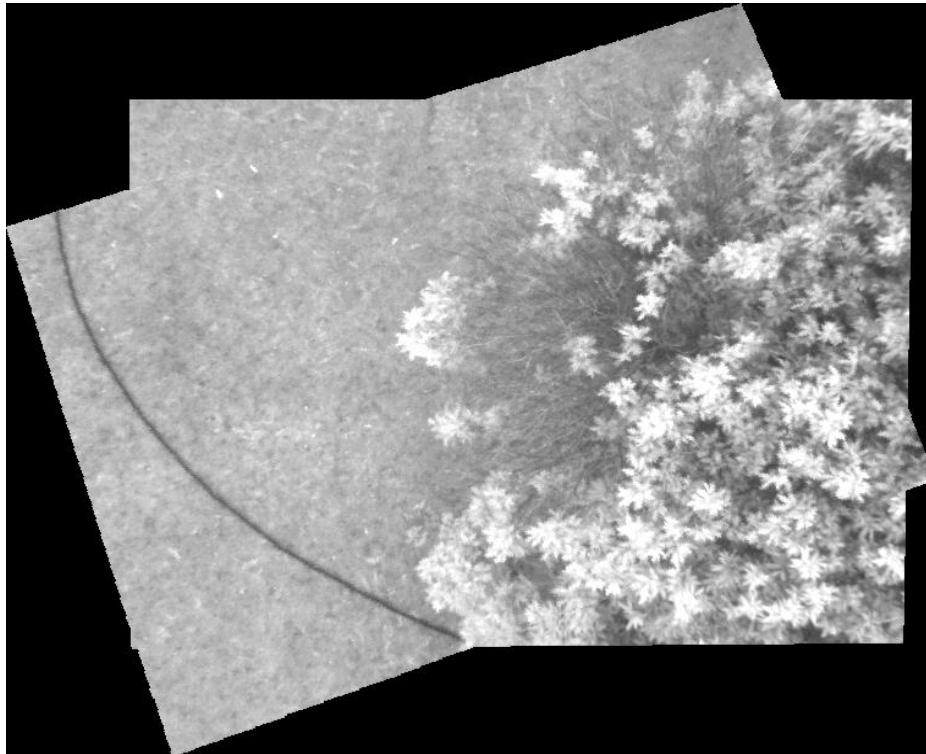


Figura 36: Mosaico generado a partir de las imágenes corregidas

Tabla9: Información del funcionamiento del algoritmo con las imágenes propuestas

Parámetros	Algoritmo en Matlab	
	Sin corrección	Con corrección
# Descriptores Imagen 1	1134	1126
# Descriptores Imagen 2	909	1074
Numero de correspondencias	176	203
Total de pixeles superpuestos	248750	250489
# pixeles correctos	161292	160549
% Aciertos	63.31	64.1
Error promedio	16.4	15.4
Tiempo de procesamiento (s)	5.0	5.3

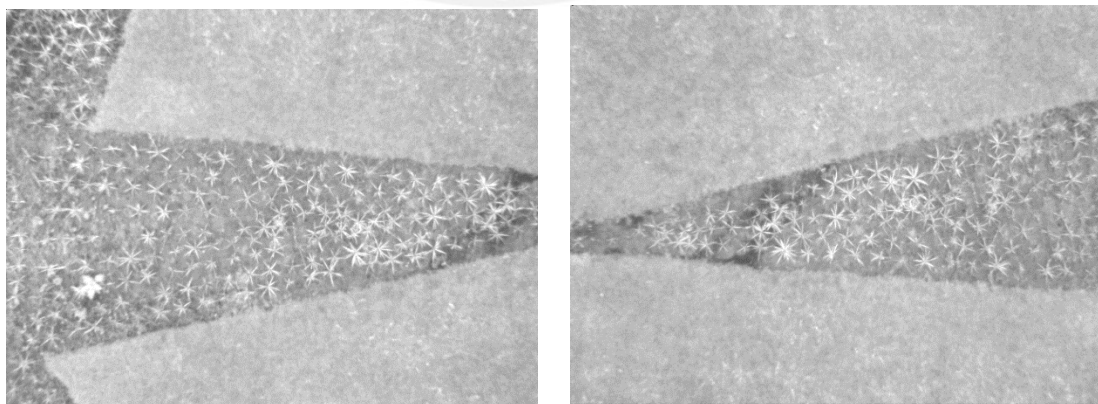


Figura 37: (Izq.) Imagen original usada como referencia. Ángulos de la imagen yaw= 0, pitch =0, roll =3. (Der) imagen original a superponer. Ángulos de la imagen yaw= 183, pitch = 2 roll=1.

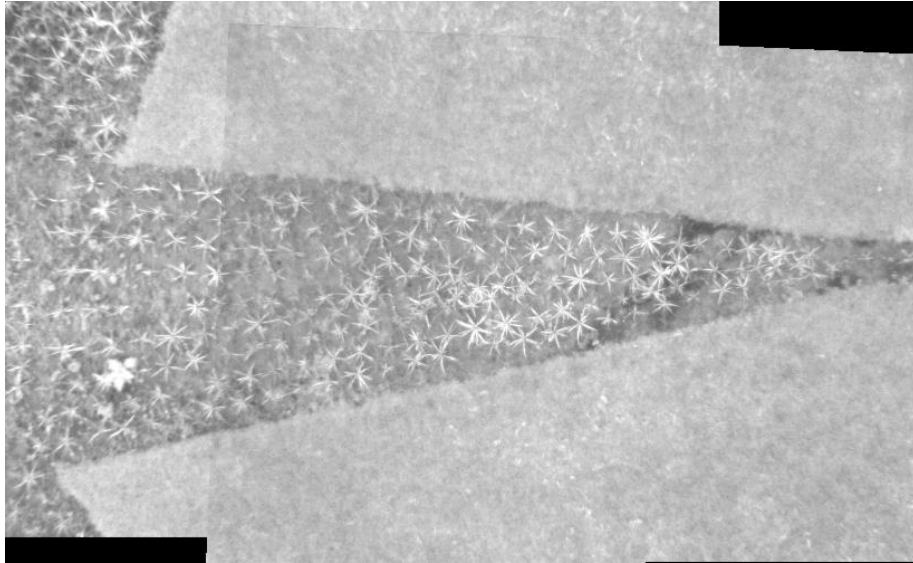


Figura 38: Mosaico generado a partir de las imágenes sin corrección

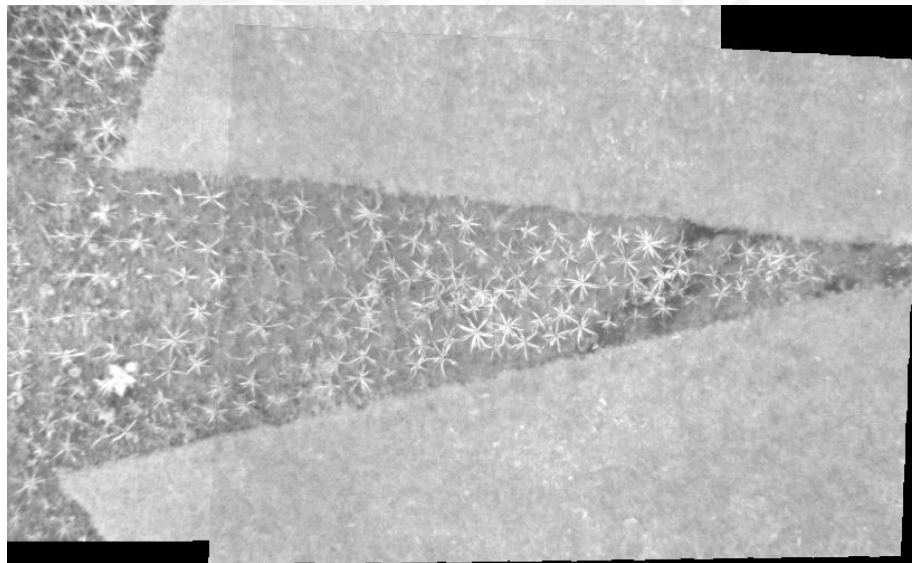


Figura 39: Mosaico generado a partir de las imágenes corregidas

Tabla10: Información del funcionamiento del algoritmo con las imágenes propuestas

Parámetros	Algoritmo en Matlab	
	Sin corrección	Con corrección
# Descriptores Imagen 1	706	700
# Descriptores Imagen 2	370	503
Numero de correspondencias	86	71
Total de pixeles superpuestos	202588	200796
# pixeles correctos	123630	120719
% Aciertos	61.03	60.12
Error promedio	9.27	9.5
Tiempo de procesamiento (s)	2.31	2.5

De las imágenes en las figuras 35, 36, 38 y 39 se puede observar que el mosaico se arma coherentemente a partir de las imágenes propuestas, también se observa que los resultados utilizando la corrección son muy similares, subjetivamente hablando, a los resultados sin utilizar este paso previo.

De los resultados obtenidos se puede observar en la tabla 9 que utilizar la corrección mejora bastante el rendimiento del algoritmo ya que permite un mayor porcentaje de aciertos y un menor error promedio, adicionalmente la corrección logra que las imágenes estén alineadas lo cual permite una mayor cantidad de correspondencias que es útil al momento de estimar la función de transformación. En la tabla 10, no se aprecia una diferencia notoria debido a que los ángulos dan a entender que la imagen simplemente esta invertida por lo cual el tiempo de corregirla no se compensa.

En este caso, las imágenes han sido tomadas con una iluminación muy parecida, por lo cual, el parámetro de porcentaje de aciertos resulta ser confiable. Ya que si las imágenes fueran tomadas a diferentes intensidades luminosas, el valor de los pixeles variaría mucho, a pesar de mantener las características de la imagen, y al momento de realizar el traslape y realizar la diferencia de pixeles, los mismos darían un valor elevado y lo cual significaría un mal traslape a pesar de no ser ese el caso.

De estas pruebas, sin embargo, se puede asegurar que la corrección genera una mejora para la realización de mosaicos. Además debería generar una mejor realización de panoramas más grandes ya que el error que se arrastraría por mosaico sería menor. Adicionalmente el hecho de alinear las imágenes (corrección geométrica) debería reducir el tiempo a nivel global al generar el panorama.

#### **4.6.2 Análisis de mosaicos de panoramas completos.**

Para esta segunda parte, además de la evaluación de intensidad de pixeles, también se evaluará el mosaico obtenido con respecto a una imagen de referencia. Es decir, se obtendrá una imagen global del mismo panorama de la estrella, obtenida con la misma cámara y a condiciones similares solo que a mayor altura, y se comparará con el mosaico obtenido mediante la comparación de los descriptores de ambas imágenes como se describió en 4.3. Al igual que en los resultados en 4.4 se evaluará los efectos que ocasione la ortorectificación y además se evaluará la eficiencia del modelo propuesto (*referencia imagen anterior*) con respecto al modelo en [10] (*referencia mosaico*). Esta prueba se realizó utilizando 16 imágenes obtenidas del campus de la PUCP. La imagen referencia así como las imágenes usadas para generar el mosaico pueden ser encontradas en el CD de anexos.



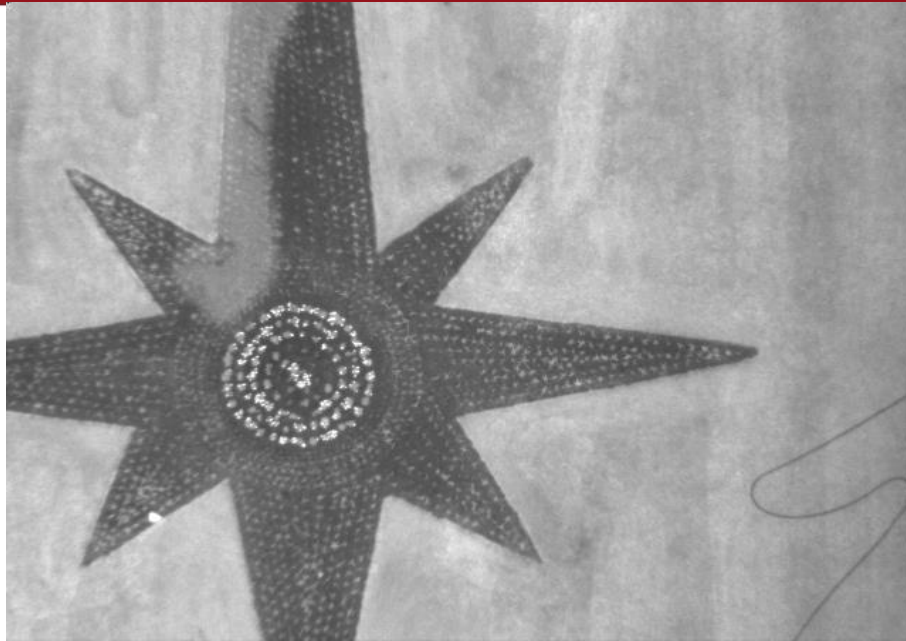


Figura 40: Imagen referencia para la comparación del mosaico generado. (Yaw = 0; Pitch= 0; Roll = 1). Se tomó esta imagen como referencia ya que era la que presentaba menos distorsiones al momento de la captura de la imagen.

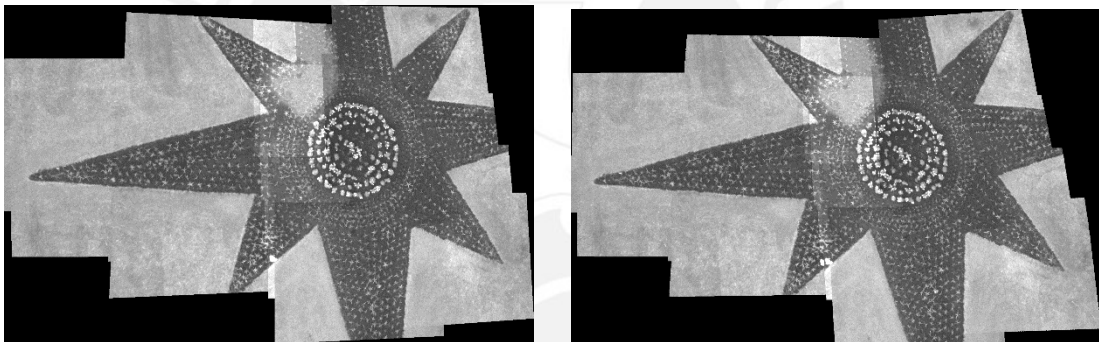


Figura 41: Mosaicos generados a partir de 16 imágenes aéreas. (Izq) Mosaico generado a partir de imágenes sin corregir. (Der) Mosaico generado con imágenes corregidas.

Tabla 11: Evaluación de parámetros de los mosaicos generados.

Parámetros	Algoritmo en Matlab			
	Sin corrección		Con corrección	
	Modo 1	Modo 2	Modo 1	Modo 2
% Aciertos en promedio	63.74	63.22	65.1	64.09
Error Promedio	11.3	12.15	10.5	11.2
Media de la relación de descriptores	2.72	2.589	2.911	2.9312
Desviación estándar de la relación de descriptores	0.1624	0.23	0.1478	0.1385
Tamaño imagen	1670x1068	1625x1069	1812x1146	1747x1158
Tiempo de procesamiento (segundos)	27.9	34.5	28.1	34.7

Nota: modo 1 significa el algoritmo propuesto (referencia imagen anterior) y modo 2 significa el algoritmo definido en [10] utilizando de referencia el mosaico generado.



Se observó que el modo propuesto obtiene resultados similares al del otro método ya que los parámetros obtenidos en la tabla 11, tanto para imágenes corregidas como las no corregidas, son similares. Adicionalmente, se ve una mejora del tiempo de procesamiento usando el método propuesto, el cual no es muy notorio debido a que solamente se procesaron 16 imágenes, pero en campos de cultivo más grandes donde se trabajen con cientos de imágenes, la diferencia en tiempos será mucho mayor.

Adicionalmente, se puede afirmar que la corrección de imágenes ayuda a obtener un mejor mosaico ya que la desviación estándar cuando se usan las imágenes corregidas es menor en comparación con las otras simulaciones (ver tabla 10).

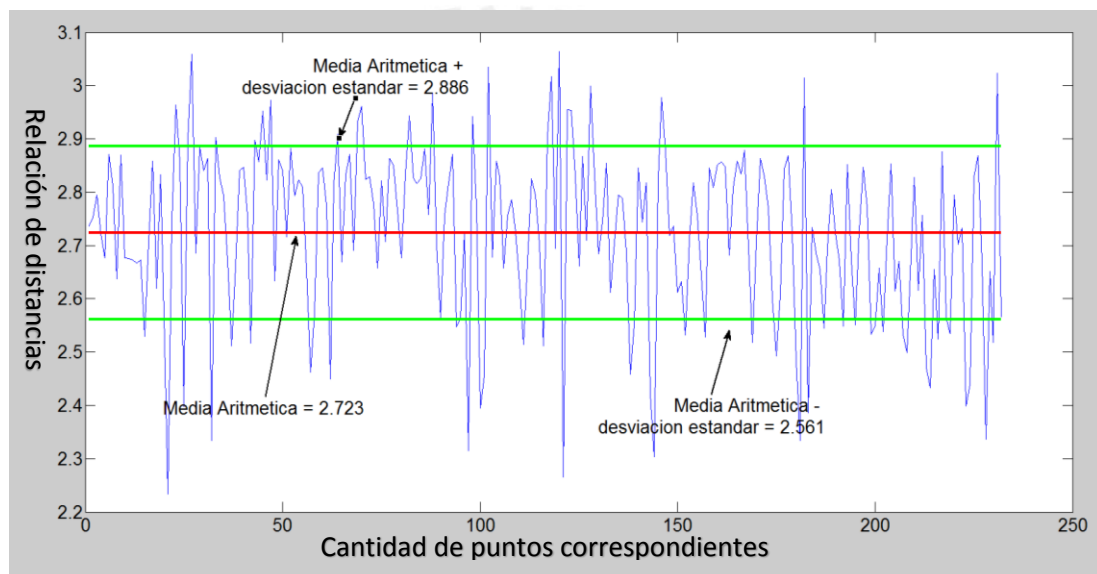


Figura 42: Relación geométrica de los puntos de interés entre la imagen de referencia y el mosaico generado a partir del modo 1 y sin imágenes corregidas.

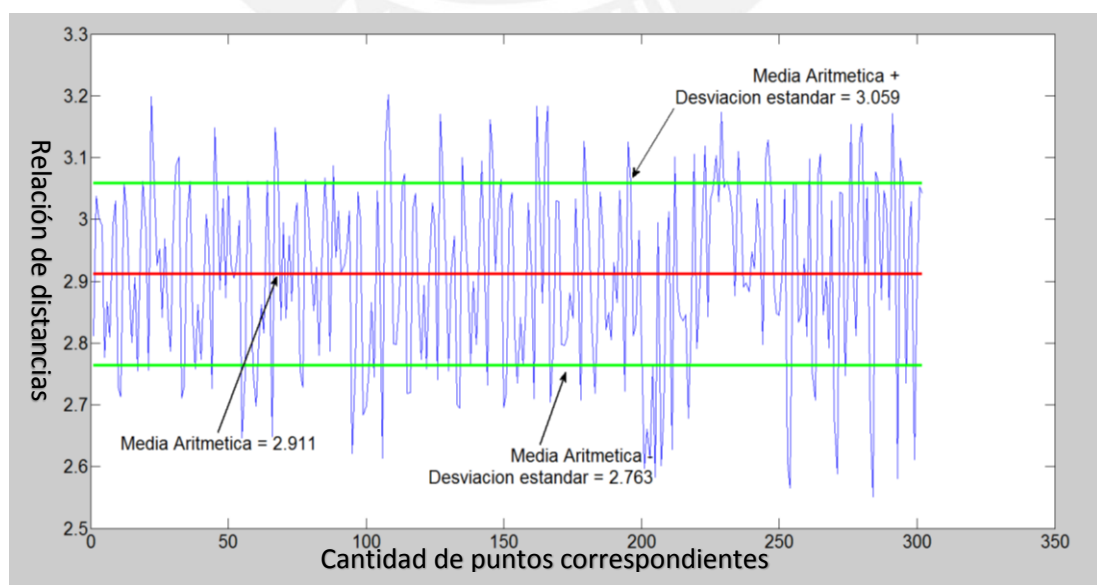


Figura 43: Relación geométrica de los puntos de interés entre la imagen de referencia y el mosaico generado a partir del modo 1 y con imágenes corregidas.

En las figuras 42, 43, 44 y 45 se observa la relación geométrica de los puntos de interés y del cual se esperó que sea una línea recta. De estas curvas cabe resaltar que los picos más pronunciados que se observan se debe a que se realizó una mala correspondencia entre algunos de los puntos debido a la semejanza de algunas partes de la imagen, por ejemplo el hecho de que la imagen tenga la forma de una estrella implica una gran cantidad de semejanzas entre los brazos de la estrella, lo cual arrastra un error al momento de realizar la correspondencia (ver figura 40). Este ejemplo mencionado también es un problema en el enfoque propuesto en [10] ya que al analizar todo el mosaico podría aparecer duplicidad de puntos en zonas similares de la imagen como en el caso de los brazos de la estrella mencionada, esto; sin embargo, es una de las ventajas del método propuesto ya que se evitan la duplicidad de puntos.

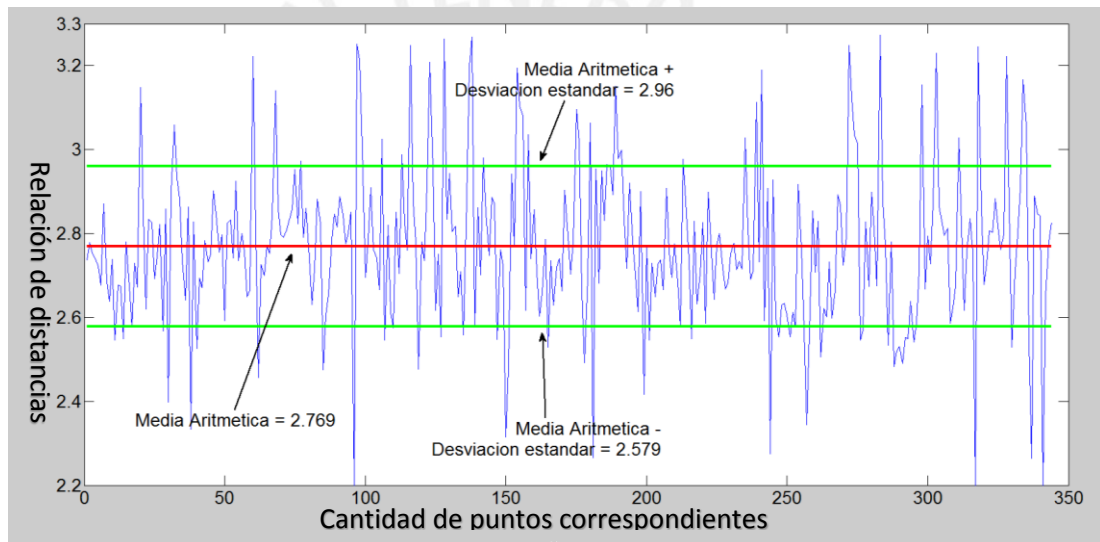


Figura 44: Relación geométrica de los puntos de interés entre la imagen de referencia y el mosaico generado a partir del modo 2 y sin imágenes corregidas.

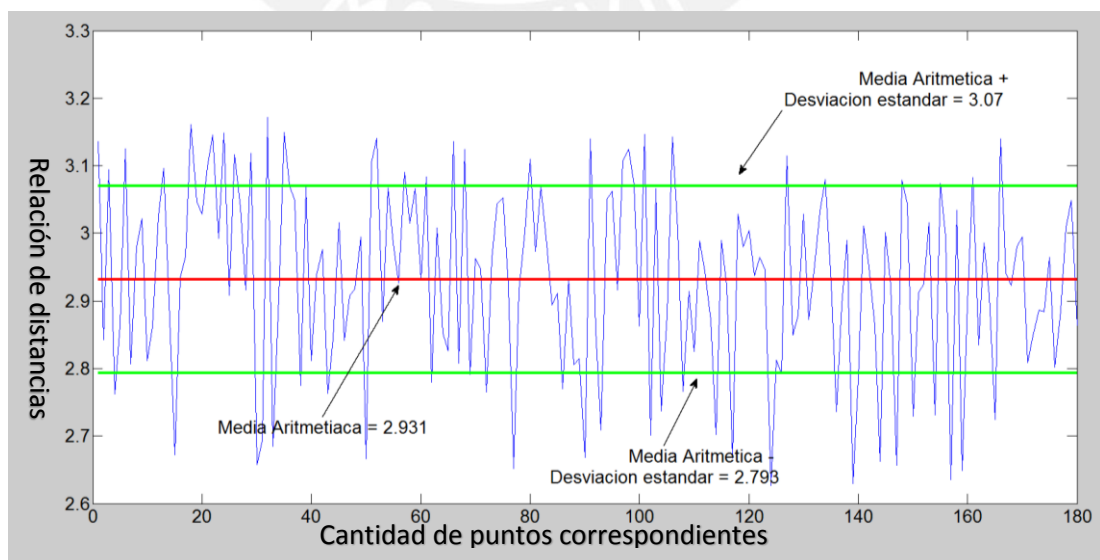
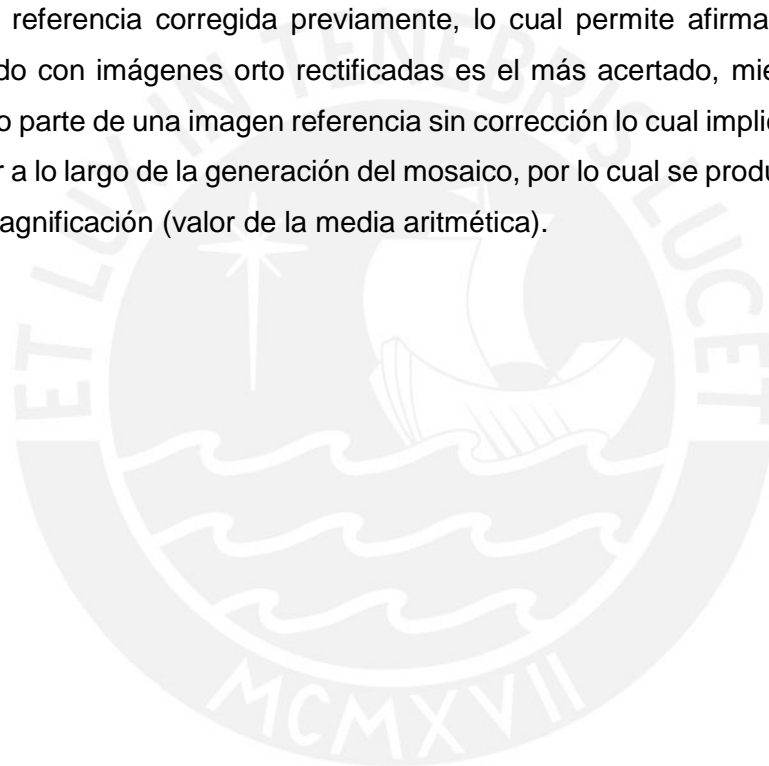


Figura 45: Relación geométrica de los puntos de interés entre la imagen de referencia y el mosaico generado a partir del modo 2 y con imágenes corregidas.

En esta prueba, en el resultado de los mosaicos, se puede observar una diferente iluminación en diferentes partes del mosaico (ver figura 41), el cual nos genera un error promedio elevado; sin embargo se puede observar una continuidad en la forma de la estrella que indica que el registro se realizó de manera correcta y lo cual es confirmado con el elevado número de porcentajes de aciertos hallados.

La media aritmética del vector de relación geométrica obtenida para el mosaico con imágenes sin corregir nos muestra que la magnificación con respecto a la imagen original es de 2.6 en promedio, mientras que para el mosaico con imágenes corregidas nos muestra una magnificación de 2.9 en promedio. Esta diferencia se debe a que el mosaico generado a partir de imágenes corregidas parte de una imagen referencia corregida previamente, lo cual permite afirmar que el mosaico generado con imágenes orto rectificadas es el más acertado, mientras que el otro mosaico parte de una imagen referencia sin corrección lo cual implica que se arrastra un error a lo largo de la generación del mosaico, por lo cual se produce esa diferencia en la magnificación (valor de la media aritmética).



## Conclusiones

1.- El sistema de registro de imágenes desarrollado a partir de los algoritmos seleccionados permite establecer relaciones geométricas con un bajo margen de error para imágenes con diferentes tipos de escena caracterizados por cambios de orientación de hasta  $75^\circ$  (peor caso del conjunto de imágenes de rotación), cambios de escala con un factor de hasta el cuádruple (peor caso del conjunto de imágenes escaladas), también permite relacionar imágenes con cambios de iluminación y cambios de enfoque. Por otro lado, el algoritmo encuentra bastantes dificultades ante las imágenes con cambios de perspectiva; sin embargo, esto era lo esperado, por lo cual se incluyó el proceso de corrección de perspectiva para compensar este efecto.

2.- Añadir el paso previo de la corrección geométrica de imágenes así como el uso del método propuesto (referencia imagen anterior), permitió obtener un mejor registro de las imágenes y también mejorar el tiempo de procesamiento del mismo, lo cual se comprobó con los parámetros obtenidos en la sección 4.5. Adicionalmente, se notó que las métricas propuestas para la evaluación de mosaico por pares es útil siempre y cuando las imágenes hayan sido tomadas a una misma iluminación, mientras que la medida global por el hecho de respaldar su métrica en los descriptores que caracterizan a la imagen nos permite tener una medida más confiable del panorama generado. Cabe mencionar que el mosaico generado no debería ser mayor del triple de tamaño que la imagen referencia ya que el número de correspondencias disminuirá y la métrica no contará con suficientes puntos para resultar confiable.

3.- El algoritmo se logró desarrollar en C y la ventaja de esta implementación es una mejora en el tiempo de procesamiento de aproximadamente la mitad, mientras que las métricas de los mosaicos son muy similares a los obtenidos en Matlab. Sin embargo, estos resultados no se incluyen en el documento ya que lo que se busca en ese documento es una manera de obtener un sistema capaz de desarrollar mosaicos de panoramas totales y una métrica confiable de los mismos, mas no una mejora en el tiempo de procesamiento.

4.- Finalmente el algoritmo está orientado a la realización de mosaicos a partir de imágenes de campos de cultivo planos, sin embargo, uno de los grandes problemas del algoritmo es encontrar puntos de control resaltantes ya que en los campos de cultivo, donde la mayoría es tierra y pasto, no se tienen muchas áreas características resaltantes, por lo cual la mayoría de imágenes aéreas que se escogieron para las pruebas tenían vegetación densa o cambios notorios para permitir una correcta generación del mosaico.



### Recomendaciones

Para obtener un registro de imágenes correcto se recomienda que al momento de realizar las tomas se asegure un traslape mayor al 50% entre imágenes. Esto permitirá un número confiable de descriptores comunes entre imágenes. Esto puede ser asegurado al momento de la toma de muestras utilizando el vuelto automático del aeromodelo.

Como ya se mencionó antes, un gran inconveniente del problema propuesto es encontrar puntos característicos en zonas de baja vegetación por lo que se recomienda poner “marcas” que permitan reflejar la luz y así al momento de adquirir las imágenes estas marcas sean captadas y sean puntos característicos para el detector del algoritmo.

Uno de los principales problemas son las características a las que son tomadas las imágenes, ya sea altura del aeromodelo, la iluminación del terreno, el traslape de las imágenes entre otros. Por esto se recomienda un protocolo de simulación para definir las características óptimas a las que se deberían adquirir las imágenes y así tener una muestra de datos confiables con los que trabajar.

Como se mencionó antes el objetivo de este documento no era obtener un sistema que realice el registro de imágenes de la manera más veloz sino mantener la calidad del registro. Para mejorar el tiempo de procesamiento del método propuesto se puede recurrir a la implementación del sistema en un lenguaje de programación más especializado o la implementación del mismo en una plataforma embebida (CUDA).

## Bibliografía

- [1] V. Torres Lozada.  
La agricultura peruana en los tiempos del TLC. *Documento de trabajo para el X congreso nacional*, 2004.
- [2] M. Marote.  
Agricultura de Precisión, 2010
- [3] R. Bongiovanni.  
Rentabilidad de la Agricultura de Precisión, 2004
- [4] B. Zitova y J. Flusser.  
2003, "Image registration methods: a survey". *Image and Vision Computing* 21, páginas 977-1000
- [5] A. Goshtasby.  
2005, "2-D and 3-D Image Registration for Medical, Remote Sensing, and Industrial Applications" Wiley & Sons.
- [6] Firefly® MV Technical Reference Manual. Point Grey, Richmond, BC, 2011
- [7] M0814-MP2, 2/3" 8mm f1.4 w/locking Iris & Focus, Computar.  
< [http://computarganz.com/product\\_view.cfm?product\\_id=552](http://computarganz.com/product_view.cfm?product_id=552) >  
[Consulta 27/11/2014]
- [8] R. Tupac Yupanqui Fernandez.  
2013, "Rectificación geométrica de imágenes mediante el empleo de transformaciones proyectivas y un sistema de medición inercial". PUCP
- [9] R. Rojas.  
2009, "Diseño de un sistema de registro de imágenes orientado a la agricultura de precisión". PUCP
- [10] J.Barba.  
2012, "Registro de imágenes para agricultura de precisión usando lenguaje C". PUCP
- [11] D. Lowe.  
2004, Distinctive Image Features from Scale-Invariant Keyoints. *International Journal of Computer Vision*, Volumen 60, Número 2, Páginas 91-110.

- [12] D. Lowe.  
1999, Object Recognition from Local Scale-Invariant Features, IEEE International Conference on In Computer Vision, páginas 1150-1157.
- [13] H. Bay, A. Ess, T. Tuytelaars y L. Van Gool.  
2008, Speeded-Up Robust Features (SURF). Computer Vision and Image Understanding 110, páginas 346-259.
- [14] K. Mikolajczyk y C. Schmid.  
2005, A performance evaluation of local descriptors. IEEE Transactions on Pattern Analysis and Machine intelligence.
- [15] K. Mikolajczyk  
2007, "Affine covariant features".  
<http://www.robots.ox.ac.uk/~vgg/research/affine/> [Consultado 27/11/2014]
- [16] Gonzalez, Rafael C.  
2008, Digital image processing. 3rd Edition, Pearson Education
- [17] R. Hartley and A. Zisserman,  
2003, Multiple View Geometry in Computer Vision, 2nd edition, Cambridge University Press.
- [18] K. Rane, M. Limbachiya, S. Pandit y S. Bhirud.  
20013, Mosaic Evaluation: An Efficient and Robust Method base don Maximun Information Retrieval. 2013
- [19] Spreading Wings S800 EVO User Manual  
[http://download.djiinnovations.com/downloads/s800evo/en/S800\\_EVO\\_User\\_Manual\\_v1.10\\_en.pdf](http://download.djiinnovations.com/downloads/s800evo/en/S800_EVO_User_Manual_v1.10_en.pdf) [Consultado 27/11/2014]
- [20] <http://www.pcdduino.com/pcduino-lite/> [Consultado 27/11/2014]
- [21] G. Bradsky y A. Kaehler.  
2008 Learning OpenCV: Computer Vision with the OpenCV Library. O'Reilly.
- [22] Imagen referencia  
<http://api.ning.com/files/dY7SoS5WqNE8hxrwPZXP6GiOaCSMFjKaob8LYJ2uUqHtChjyFKcTqoKJJ0xzWIR61jhQ42o3fzql-Moja9HRNUT3mFU5kXpO/GridV2.jpg>  
[Consultado 27/11/2014]