

UNIVERSIDAD NACIONAL DE CÓRDOBA

FACULTAD DE MATEMÁTICA, FÍSICA Y ASTRONOMÍA



# REGISTRO DE IMÁGENES DE TELEDETECCIÓN ESPACIAL

**Autor: Marcos Javier Gómez**

**Director: Oscar H. Bustos**

2014

---

# Resumen

El registro de imágenes se ha transformado en una de los procesos de mayor interés e importancia en el procesamiento de las mismas. Podemos observar que gran parte del registro, se sigue realizando de manera manual, lo cual es una tarea tediosa, que puede producir resultados erróneos. En este trabajo, se estudian los diferentes métodos automáticos para el registro de imágenes. En particular se implementa el algoritmo automático de la transformada de Fourier. Se estudia su compartamiento y desempeño para el registro de imágenes satelitales.

Image registration has been transformed in one of the most interesting and important process on the field of Image Processing. In most cases, we can see that a lot of people choose manual registration, that is a difficult task, and can produce bad results. In this work, we study the different automatic methods for image registration. In particular we implement the automatic algorithm of Fast Fourier Transform (FFT). Finally, we study the behavior and performance of FFT for registering teledeteccion images.

**Clasificación:** 1.4.3 Registration (Image Processing and Computer Vision).

**Palabras Claves:** Teledetección, Imágenes satelitales, Registro de Imágenes, algoritmos de registro automáticos, métodos basados en intensidad, Transformada Rápida de Fourier, correlación, match.

---

# Agradecimientos

*Este trabajo no habría sido posible sin el apoyo, estímulo y acompañamiento de mi familia, novia y amigos.*

*Agradezco a Oscar Bustos por su dedicación, y siempre motivar mi desempeño en el desarrollo del trabajo. A la comunidad de FaMAF por todo que me han brindado a lo largo de la carrera.*

*Por último dedico mi trabajo final a mi Nono, con quien desearía poder estar compartiendo este momento.*

*Marcos Javier Gómez*

---

# Índice general

<b>Resumen</b>	<b>2</b>
<b>Agradecimientos</b>	<b>3</b>
<b>1. Introducción</b>	<b>9</b>
1.1. Objetivos de este trabajo . . . . .	9
1.2. Organización del trabajo . . . . .	10
<b>2. Teledetección Espacial</b>	<b>11</b>
2.1. Definición de la Teledetección . . . . .	11
2.1.1. Componentes de un sistema de teledetección espacial . . . . .	12
2.1.2. Principios físicos de la Teledetección . . . . .	14
2.1.3. Características de los Sensores . . . . .	16
2.1.4. Imágenes satelitales . . . . .	22
2.2. Historia de la teledetección . . . . .	23
2.3. Importancia de la teledetección espacial . . . . .	25
2.3.1. Beneficios de las Imágenes Satelitales . . . . .	27
<b>3. Registro de Imágenes</b>	<b>30</b>
3.1. Definición del Registro de Imágenes . . . . .	30
3.1.1. Elementos de los Algoritmos de Registro de Imágenes . . . . .	32
3.2. Un poco de historia del Registro de Imágenes . . . . .	35
3.3. Importancia del Registro de Imágenes . . . . .	39
3.4. Enfoques Generales del Registro de Imágenes . . . . .	42
3.4.1. Detección de Características . . . . .	43

---

3.4.2. Matcheo o alineación de características . . . . .	44
3.4.3. Estimación de Modelos de Transformación . . . . .	50
3.4.4. Transformaciones Radiométricas y Construcción de imagen ( <i>Re-sampling</i> ) . . . . .	52
<b>4. Algoritmo automático de la transformada rápida de Fourier</b>	<b>57</b>
4.1. Fundamentos teóricos . . . . .	57
4.1.1. Transformada de Fourier y sus propiedades . . . . .	57
4.1.2. Coeficiente de correlación en el dominio de Fourier . . . . .	60
4.1.3. Algoritmo y su implementación . . . . .	67
<b>5. Recolección de datos y Evaluación del Algoritmo</b>	<b>70</b>
5.1. Imágenes utilizadas para la evaluación . . . . .	70
5.1.1. Evaluación del algoritmo de con transformaciones individuales	72
5.1.2. Evaluación con múltiples transformaciones . . . . .	74
5.1.3. Obtención de una Imagen a partir de varios recortes . . . . .	76
<b>6. Conclusión y Trabajo Futuro</b>	<b>79</b>
<b>Bibliografía</b>	<b>81</b>

---

# Índice de figuras

2.1. Sistema de Teledetección . . . . .	14
2.2. Espectro Electromagnético . . . . .	16
2.3. Landsat 3 . . . . .	20
2.4. Spot 5 . . . . .	21
2.5. Palomas espaciales. . . . .	24
2.6. Primeros Satélites . . . . .	26
3.1. Proceso de Registro de Imágenes . . . . .	34
3.2. Patente cámara de enfoque de Becker . . . . .	37
3.3. Patente Dispositivo de Comparación de Dressler . . . . .	38
3.4. Patente Tubo Rayos Catódicos para Matcheo de Imágenes desarrollado por Enlace y SMITH . . . . .	40
3.5. Elección de valor de pixel, mediante Vecino más Cercano . . . . .	55
3.6. Elección de valor de pixel mediante Interpolación Bilineal . . . . .	56
3.7. Elección de valor de pixel mediante Convolución Cúbica . . . . .	56
5.1. Imagen completa de la ciudad de Córdoba (RGB) tomada por Lansdat TM . . . . .	71
5.2. Imagen completa de la ciudad de Córdoba (RGB) tomada por Lansdat ETM+ . . . . .	72
5.3. Recorte de la imagen (RGB) de la ciudad de Córdoba tomada por Lansdat ETM+ . . . . .	75
5.4. Aplicación de transformaciones a 5.3 . . . . .	75
5.5. Imagen generada por el registro entre 5.3 y 5.4 . . . . .	76

---

5.6. Recortes de la imagen (RGB) de la ciudad de Córdoba tomada por Landsat ETM+ . . . . .	77
5.7. Registro entre imágenes (5.6a) y (5.6b) . . . . .	78
5.8. Registro entre imágenes (5.7) y (5.6c) . . . . .	78

---

# Índice de cuadros

5.1. Resultados registro de imágenes afectadas por traslaciones . . . . .	73
5.2. Resultados registro de imágenes afectadas por rotaciones . . . . .	73
5.3. Resultados registro de imágenes afectadas por cambio de escala . . .	74

---

# Capítulo 1

## Introducción

El registro automático de imágenes, consta en alinear las mismas calculando un conjunto de parámetros para llevar a cabo las transformaciones matemáticas necesarias para generar la alineación. Esto parece ser un problema perfectamente definido, por lo que tendría una solución universal. El problema, es que estamos muy lejos de esa realidad.

El registro de imágenes hoy en día, es una tarea compleja. El aumento de categorías en las cuales se necesita del mismo, admite múltiples estrategias de solución. La gran disponibilidad de imágenes en teledetección, medicina, robótica, entre otras, junto a la diversidad de fuentes de imágenes, nuevos instrumentos donde el registro es aplicado, nuevos principios y algoritmos, generan un interés en la investigación en este campo.

### 1.1. Objetivos de este trabajo

1. Estudiar en profundidad los conceptos principales del registro de imágenes.
2. Desarrollar y evaluar algoritmo automático de la transformada de Fourier para el registro de imágenes.
3. Contribuir a la comunidad del software libre, brindando una opción para el registro de imágenes automático.

## 1.2. Organización del trabajo

Este trabajo está compuesto por cinco capítulos más, además del actual. Haremos una pequeña síntesis de lo que se estudiará en cada uno:

**Capítulo 2** : En este capítulo se describe la teledetección espacial. Hacemos referencia a su historia, elementos fundamentales e importancia de la misma.

**Capítulo 3**: Aquí describimos el proceso de registro de imágenes. Definimos el mismo, describimos sus características. Definimos los diferentes métodos que existen para llevar a cabo este proceso, como también los pasos necesarios.

**Capítulo 4**: Este capítulo se dedica de lleno al algoritmo que se ha implementado: *el algoritmo automático de la transformada rápida de Fourier*. Se enuncian propiedades teóricas las cuales son necesarias, para poder definir el algoritmo. Por último se explica su funcionamiento e implementación.

**Capítulo 5** : El penúltimo capítulo describe como fue evaluado el algoritmo. Se explica como se han recolectado los datos, necesarios para poder generar una conclusión sobre el desempeño del mismo.

**Capítulo 6** : En este último capítulo damos a conocer la conclusión final del trabajo. Además se proponen posibles trabajos futuros.

---

## Capítulo 2

# Teledetección Espacial

### 2.1. Definición de la Teledetección

La palabra **Teledetección** deriva del francés *téledétection*, traducido al inglés como *Remote Sensing*. Es complicado dar una definición exacta de lo que es la teledetección ya que no existe una definición universalmente aceptada, debido a que, dependiendo del interés con el cual se trabaje sobre la misma, cambia en parte su significado.

Lo que sí podemos hacer, es enunciar de manera muy general a la **Teledetección** como:

*La adquisición de información de un objeto a distancia, sin que exista contacto material entre el objeto o sistema estudiado, y el observador.*

En esta definición encontramos incluido el caso mas común y cotidiano de teledetección, el cual hace referencia al proceso de visión del ser humano (utilizando el ojo como *sensor*). Otras actividades como ecografías e imágenes detectadas por satélites, se identifican e incluyen en esta definición.

A pesar de haber múltiples definiciones de teledetección, podemos encontrar elementos en común en las mismas:

- *La adquisición de información a distancia, sin entrar en contacto con el objeto.*
- *Se realiza a través de sensores o receptores.*
- *Se registra radiación electromagnética.*

- *La transformación de los datos obtenidos para ser utilizados con diferentes objetivos.*

La teledetección parte del principio de una perturbación (energía electromagnética, campos gravitacionales, ondas sísmicas, etc) que el sistema bajo observación produce en el medio, la cual es registrada por el receptor, para luego ser utilizada con objetivos específicos.

Con respecto a la **teledetección espacial**, podemos definirla como:

*el conjunto de conocimientos y técnicas utilizadas para determinar características físicas y biológicas de objetos y superficies mediante mediciones a distancia, sin el contacto material con los mismos. La teledetección no engloba sólo los procesos que permiten obtener una imagen desde el aire o el espacio, sino también su posterior tratamiento en el contexto de una determinada aplicación. Por ende, la teledetección es la técnica de adquisición y posterior tratamiento de datos de la superficie terrestre desde sensores instalados en plataformas espaciales, en virtud de la interacción electromagnética existente entre la Tierra y el sensor, siendo la fuente de radiación proveniente del sol (teledetección pasiva) o del propio sensor (teledetección activa).*

### 2.1.1. Componentes de un sistema de teledetección espacial

Un sistema de teledetección espacial precisa de la *escena*, *sensores* y el *tratamiento de la información*, para poder llevar a cabo su objetivo. Además, los siguientes componentes, son fundamentales para la teledetección espacial: (2.1):[7]

#### (A) Fuente de energía o iluminación:

La fuente de energía que supone el origen del flujo energético detectado por el sensor. Puede tratarse de un foco externo al sensor, en cuyo caso se habla de *teledetección pasiva*, o de un haz energético emitido por éste, conocido como *teledetección activa*. La fuente de energía más importante es la energía solar.

#### (B) Radiación y la atmósfera:

Es el proceso y el medio en el que la energía electromagnética es propagada a través del espacio libre.

**(C) Interacción con la superficie terrestre:**

La cubierta terrestre, formada por distintas masas de vegetación, suelo, agua y construcciones, que reciben la señal energética procedente de la fuente de energía, y la reflejan o emiten de acuerdo a sus características físicas.

**(D) El sistema Sensor:**

Compuesto por el sensor propiamente dicho y la plataforma que lo sustenta. Tiene como misión captar la energía procedente de las cubiertas terrestres, codificarla y enviarla directamente al sistema de recepción.

**(E) Sistema de transmisión, recepción, y procesamiento:**

El Sistema de recepción recibe la información transmitida por el sensor. Esta información es grabada en un formato apropiado y es distribuido a los intérpretes para su tratamiento visual o digital.

**(F) Interpretación y análisis:**

En esta etapa el intérprete, analiza la información.

**(G) Aplicación:**

El último proceso es el de extraer la información necesaria para resolver un problema específico; esa información llega al usuario quien es el encargado de analizar el producto, así como dictaminar el impacto que de él se derive.

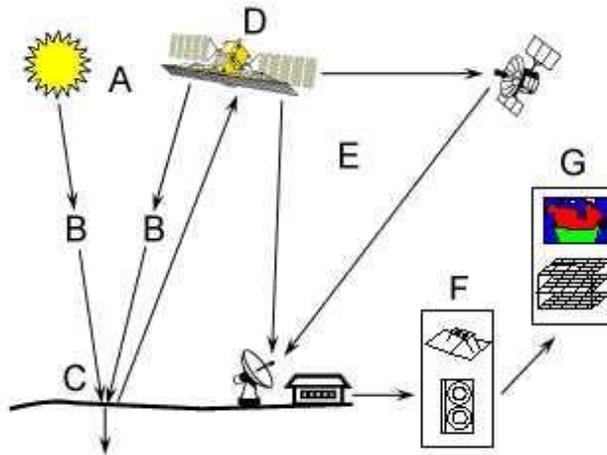


Figura 2.1: Sistema de Teledetección

### 2.1.2. Principios físicos de la Teledetección

Habitualmente se asume que la formación de imágenes en teledetección se basa en el uso de la **radiación electromagnética** para adquirir la información a distancia, esto es partiendo de sensores remotos. Las propiedades de la radiación electromagnética se han explicado históricamente a través de dos teorías: [7]

1. **Haz ondulatorio.** (*Huygens, Maxwell*)
2. **Sucesión de unidades discretas de energía, fotones, con masa igual a cero.** (*Planck, Einstein*)

En la actualidad se ha demostrado que la luz puede comportarse de ambas formas. Basándonos en la **teoría ondulatoria**, la energía electromagnética se transmite de un lugar a otro en forma de onda continua a la velocidad de la luz, conteniendo dos campos de fuerzas ortogonales entre sí: *campo eléctrico* y *campo magnético*. El flujo energético conforme a la teoría ondulatoria puede describirse por dos elementos:

- **Longitud de onda ( $\lambda$ ):** Hace referencia a la distancia entre dos picos sucesivos de una onda. Se mide en *micrómetros* ( $1\mu m = 10^{-6}m$ ).
- **Frecuencia ( $F$ ):** Designa el número de ciclos pasando en un punto fijo en una unidad de tiempo. Se mide en *Hertzios* (*ciclos por segundo*).

Estos dos elementos están inversamente relacionados por la siguiente ecuación (2.1):

$$c = \lambda * F \quad (2.1)$$

donde  $c$  es la velocidad de la luz,  $\lambda$  la longitud de onda y  $F$  la frecuencia.

### El espectro electromagnético

Aunque la sucesión de valores de longitud de onda es continua, suelen establecerse una serie de bandas en donde la radiación electromagnética manifiesta un comportamiento similar.

Desde el punto de vista de la teledetección, conviene destacar una serie de bandas espectrales, que son las más empleadas teniendo en cuenta la tecnología actual:

1. **Espectro visible** (0,4 a 0,7 $\mu m$ ): Se distinguen tres bandas elementales: *azul* (0,4 a 0,5 $\mu m$ ), *verde* (0,5 a 0,6 $\mu m$ ) y *roja* (0,6 a 0,7 $\mu m$ ).
2. **Infrarrojo próximo**(0,7 a 1,3 $\mu m$ ): Resulta de especial importancia por su capacidad para discriminar masas vegetales y concentraciones de humedad.
3. **Infrarrojo medio**(1,3 a 8 $\mu m$ ): Donde se entremezclan los procesos de reflexión de la luz solar y de emisión de la superficie terrestre. Resulta idóneo para estimar contenido de humedad en la vegetación y detección de focos de alta temperatura.
4. **Infrarrojo lejano o térmico**(8 a 14 $\mu m$ ): Incluye la porción emisiva del espectro terrestre, en donde se detecta el calor proveniente de la mayor parte de las cubiertas terrestres.
5. **Micro - ondas**(a partir de 1mm): Genera un gran interés por ser un tipo de energía bastante transparente a la cubierta nubosa.

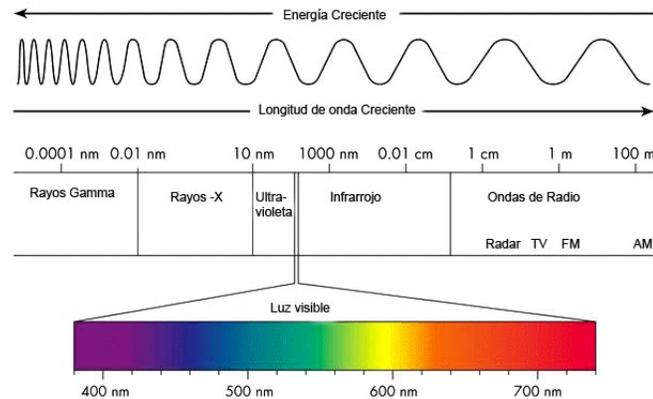


Figura 2.2: Espectro Electromagnético

### 2.1.3. Características de los Sensores

Para que un sensor pueda registrar energía reflejada o emitida desde un blanco, este sensor debe estar sobre una plataforma alejado del objetivo a observar. Estas plataformas podrían estar sobre la Tierra, sobre aviones, globos, así como sobre naves espaciales o satélites fuera de la atmósfera. Los sensores usados en tierra suelen cumplir la función de registrar información muy detallada de alguna superficie la cual, luego, es comparada con la obtenida desde aviones o satélites, por lo se los considera como validación terrestre para entender mejor la imagen satelital. Estos sensores terrestres pueden estar sobre la mano, sobre edificios o grúas.[7]

Cada sensor es diseñado para captar características específicas, basadas en resoluciones espaciales, espectrales, radiométricas y temporales. Cuando hablamos de **resolución**, nos referiremos a la unidad más pequeña de granularidad que puede ser medida por el sensor.

A continuación se describirá cada tipo de resolución y la importancia de las mismas para la teledetección espacial.

**La Resolución Espacial**, corresponde al sector del terreno del cual se tomará la reflectancia para computar el valor asignado a cada píxel.

Una imagen está formada por una matriz de elementos, o píxeles, los cuales constituyen las unidades más pequeñas de la imagen. Los píxeles de la imagen son generalmente cuadrados y representan cierta área de la imagen. Es impor-

tante distinguir entre el tamaño de píxel y la resolución espacial. Si un sensor tiene una resolución espacial de  $20m$  y la imagen es presentada en máxima resolución, cada píxel representa un área de  $20m \times 20m$  sobre la Tierra.

Hay que considerar que la resolución espacial de un sensor depende de varios factores, como son la altura orbital, velocidad de exploración y número de detectores.

**La Resolución Espectral**, define las bandas que se utilizarán del espectro electromagnético. Indica el número y ancho de las bandas espectrales que puede discriminar el sensor. Este será más eficiente cuanto mayor sea el número de bandas que ofrezca, pues facilita la caracterización espectral de las distintas cubiertas. También conviene que esas bandas sean suficientemente estrechas, con objeto de recoger la señal sobre regiones coherentes del espectro.

**La Resolución Radiométrica**, define el número de *bits* utilizados para guardar la energía correspondiente a la longitud de onda utilizada.

Hace mención a la sensibilidad del sensor, esto es, a su capacidad para detectar variaciones en la radiancia espectral que recibe. En el caso de los sistemas fotográficos, la resolución radiométrica del sensor se indica por el número de niveles de gris posibles de registrar en la película. Para los óptico - electrónicos, la imagen habitualmente se presenta en formato digital, por lo que la resolución radiométrica suele identificarse con el rango de valores que codifica el sensor. Cuando se trata de codificación digital, habitualmente se expresa esa resolución en el número de bits que precisa cada elemento de imagen para ser almacenado (256, 128, 64, etc.) Cuanto mayor sea la precisión radiométrica, tanto mejor podrá interpretarse la imagen. El número de niveles más adecuado depende del método que se siga en la interpretación. Si ésta es visual, conviene considerar que el ojo humano difícilmente percibe más de 64 niveles de gris, y no más de 200.000 tonalidades. En la interpretación digital, la sensibilidad radiométrica se puede aprovechar mejor porque la computadora es capaz de distinguir una escala numérica amplia.

**La Resolución Temporal**, determina la frecuencia de observación, generada por la órbita del satélite. Se refiere al período de tiempo que lleva el sensor en

completar un ciclo orbital completo.

Otra de las características principales de los sensores es la *Signal-to noise ratio* (SNR). Hacemos referencia a *ruido* como las variaciones de intensidad que son detectadas por el sensor y no son parte de las características usuales del objeto. Si el nivel de ruido es muy alto, los datos obentidos no serán una representación óptima.[6]

Dada una longitud de onda  $\lambda$ , utilizamos a  $SNR$  como una función de detección de calidad:

$$(S/N)_\lambda = D_\lambda \beta^2 \sqrt{H/V} \Delta_\lambda L_\lambda \quad (2.2)$$

donde,

$D_\lambda$  = capacidad de detección del sensor.

$\beta$  = campo de visión instantánea.

$H$  = altura de vuelo de la nave espacial.

$V$  = velocidad de la nave espacial.

$\Delta_\lambda$  = resolución espectral.

$L_\lambda$  = radianza espectral de las características del terreno.

Se puede observar en (2.2) que para mantener el SNR de un sensor, en un nivel aceptable, tendremos que negociar entre las diferentes características del mismo. Por ende, a pesar que pareciese ideal tener infinitas y bandas espectrales, la ecuación nos marca que es imposible aumentar las mismas sin sufrir deterioro en el SNR.

### Características de las órbitas satelitales

El camino seguido por un satélite es conocido como **órbita**. Diferentes trayectorias u órbitas deberan ser elegidas para un satélite, teniendo en cuenta varios requisitos, cómo, las características de los sensores, la frecuencia de toma de datos, el territorio y fenómenos a cubrir. La órbita elegida puede variar en altitud, y su orientación y rotación con respecto a la Tierra.

La mayoría de las plataformas de sensores remotos tienen órbitas (*norte- sur*) que en conjunción con el movimiento de la Tierra (*oeste-este*), permite cubrir toda la supercie terrestre cada cierto tiempo. Se denominan órbitas *casi polares* ya que tienen un pequeño ángulo con el eje terrestre. Muchos de éstos son además *heliosincrónicos*

de tal forma de cubrir cada área del mundo a una hora local solar constante. Esto es importante para asegurar condiciones de iluminación constantes tanto para análisis multitemporales como para la realización de mosaicos. La mayoría de las plataformas satelitales de teledetección de hoy tienen estas órbitas casi polares, lo cual significa que ellas viajan hacia el norte en la mitad del planeta y hacia el sur en la otra mitad de su órbita. Éstas se denominan *pasadas ascendentes y descendentes* respectivamente. Si la órbita además es heliosincrónica, la pasada ascendente suele ser del lado oscuro de la Tierra y la descendente es cuando hay iluminación solar disponible. Así los sensores de energía solar reflejada tomarán imágenes sólo en las pasadas descendentes y los activos y o térmicos también tomarán imágenes en las ascendentes. Cuando un satélite gira alrededor de la Tierra él ve sólo una parte de la superficie terrestre. El área mirada sobre la Tierra es denominado *campo de visión (swath)*. [7]

### Clasificación de los Sensores

Existe diversidad de sistemas de clasificación de sensores remotos. Los criterios empleados varían entre:

- **Banda del espectro electromagnético utilizada**
- **Fuente emisora de energía:** Se pueden distinguir dos tipos básicos de sensores:
  1. *Sensores Pasivos:* Detectan la radiación electromagnética emitida o reflejada de fuentes naturales.
  2. *Sensores Activos:* Poseen fuentes internas que generan artificialmente la radiación.
- **Ubicación espacial de la fuente emisora y receptora de energía:**
  1. *Sensores Mono-Estáticos:* Poseen tanto la fuente de emisión de energía como la fuente de recepción de energía. Es decir ambas fuentes, de emisión y recepción de energía, ocupan la misma posición espacial.
  2. *Sensores Bi-Estáticos,* Poseen solamente la fuente de recepción de energía. La fuente de emisión de energía (el sol) tiene una posición espacial diferente.

- **Información Registrada:**

1. *Sensores Fotográficos*, registran toda la información recibida en emulsión fotosensible.
2. *Sensores No Fotográficos*, proyectan la información que reciben sobre una pantalla, en forma gráfica o digital. Tienen la ventaja de detectar la energía correspondiente a una banda muy angosta del espectro (radiómetro, espectrómetro).

- **Mecanismo utilizados para captar la información:**

1. **Sensores Fotográficos**

Sensores que se encuentran compuestos por cámara y película.

2. **Sensores Electro-ópticos formadores de imagen**

Este tipo de sensor, en base a los sistemas utilizados para captar información, es clasificado en tres grandes grupos:

- **De formato a cuadro:**

Es un sensor autosuficiente, porque no necesita movimiento de barrido para captar imágenes de un área.

El primer sensor de este tipo, fue el sistema de televisión. Se emplea pantalla fotosensitiva con un haz electrónico de barrido, que se coloca en el foco del sistema óptico gran angular. Un sistema de este tipo fue usado en los programas LANDSAT 1 a 3.



Figura 2.3: Landsat 3

- **Barredores estáticos**

Es un sensor que posee movimiento de barrido para captar imágenes de un área. Barre en una dirección por medio de lecturas electrónicas y requiere de un sistema auxiliar de movimiento o inclinación para barrer en el otro sentido. Para el desplazamiento sobre áreas relativamente grandes, necesita de una fuente externa para desplazarse en otro sentido. Este tipo de desplazamiento es proporcionado por la plataforma. Un sistema de este tipo fue usado en los programas SPOT 2.4 y MAPSAT.



Figura 2.4: Spot 5

- **Barredores mecánicos**

Es un sensor que posee movimiento de barrido para captar imágenes de un área. Consta de una amplia variedad de sistema de barrido doble o sencillo. Barre de derecha a izquierda y requiere una fuente externa para desplazarse en otro sentido. Normalmente, esta última forma de movimiento, lo proporciona la plataforma portadora. Este tipo de barredor, ha sido ampliamente utilizado en estudios de recursos naturales. Es transportado por plataforma aérea como por plataforma espacial.

### 3. Sensores Electro-ópticos no formadores de imagen

Este tipo de sensor, no forma imagen a partir de la información captada, sino que produce curvas, grupos de números o números simples, que identifica la energía electromagnética recibida. Esta energía recibida es aquella

emitida desde, reflejada por, o transmitida a través de una superficie o región del espacio. La mayor aplicación de este tipo de sensor se encuentra en la medición de la cantidad óptica tal como el flujo radiante. Aquí se mide la densidad y la radiancia; las mismas que describen la intensidad del campo de radiación o las propiedades ópticas de una superficie o región del espectro. El sensor electro-óptico no formador de imagen, es el *Radiómetro*.

#### 4. Sensores microonadas

De acuerdo con el modo de operación, son sistemas activos o pasivos.

##### 1. Sistema Activo Formador de Imágenes:

- **Sistema activo formador de imágenes**

El más utilizado en Teledetección es el *RADAR (Radio Detection and Ranging)*, que opera según sea el sistema de apertura real o apertura sintética.

- **Sistema pasivo no formador de imágenes**

#### 2.1.4. Imágenes satelitales

Las imágenes satelitales representan las formas de captura indirecta de la información espacial. Estas son almacenadas como matrices (cada matriz corresponde a una banda del sensor), donde el elemento de la imagen es llamado píxel (acrónimo del inglés picture element) y tienen un valor proporcional a la energía reflejada o emitida de la superficie terrestre para el área de la imagen. El píxel es la unidad visual más pequeña que se puede representar en una imagen digital. El sensor detecta la radiación media de un objeto o elemento sobre la superficie correspondiente al tamaño del píxel. Este valor es traducido por el sensor a un valor numérico entero denominado Nivel Digital (ND), también conocido como valor de brillo o escala de grises; a partir de este se realiza el procesamiento digital de las imágenes. El arreglo de varios píxeles genera la imagen que se puede interpretar como una matriz n-dimensional, donde cada una de estas dimensiones representa la cantidad de bandas espectrales que contenga el sensor.

## 2.2. Historia de la teledetección

Se dice que la teledetección surge y se desarrolla en el siglo XX, en dos etapas básicas: la **teledetección aérea** y la **teledetección espacial**. Los años 60 con el desarrollo de los satélites artificiales, marcan el comienzo de la famosa era espacial.

A pesar de que la fotografía comienza a principios de siglo XIX (*Niépce y Daguerre producen las primeras **heliografías***), se tendría que esperar hasta inicios del siglo pasado donde se dan tres hechos que hacen posible la teledetección:

1. **La mejora de las cámaras fotográficas.**
2. **La aviación.**
3. **Los comienzos de la estereoscopía<sup>1</sup> y la aplicación de fotografías aéreas con fines topográficos**

### Etapa Precursora

Las primeras fotografías tomadas desde un avión datan de alrededor de 1909, cuando el piloto Wilbur Wright llevó a un pasajero, el cual tomó fotografías cerca de Roma. Aunque en realidad las primeras fotografías aéreas se realizaron con anterioridad con métodos más rudimentarios. Por ejemplo, a finales de siglo XIX se registran imágenes de París, las cuales fueron tomadas por cámaras que iban atadas a palomas 2.5. Estas fotografías eran **incontroladas**. Las primeras fotografías controladas se obtuvieron desde globos aerostáticos.

Todos estos avances fueron eclipsados por la fotografía aérea, generando el nacimiento de la **teledetección aérea**, la cual utilizaba el avión como plataforma. Su crecimiento está ligado al desarrollo de aplicaciones con fines militares en torno a la primera guerra mundial.

### Desarrollo de la teledetección espacial

Tras el acontecimiento bélico, durante el período de entreguerras (1920 - 1939), tienen lugar las primeras aplicaciones científicas y prácticas de la fotografía aérea.

---

<sup>1</sup>La estereoscopía es cualquier técnica capaz de recoger información visual tridimensional y/o crear la ilusión de profundidad mediante una imagen estereográfica, un estereograma, o una imagen 3D (tridimensional).

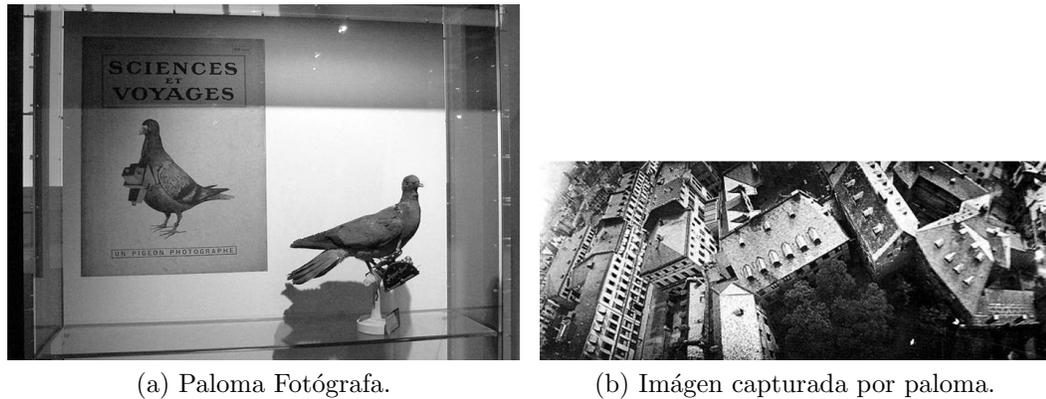


Figura 2.5: Palomas espaciales.

Surge un número importante de publicaciones científicas sobre el uso de la fotografía aérea, y comienzan las aplicaciones a escala. La fotografía aérea se dedica a la prospección geológica y minera, realización de inventarios forestales y a la cartografía de zonas poca conocidas.

Durante la segunda guerra mundial el control de la investigación y desarrollo de la fotografía aérea quedó a cargo de los estados dedicados a la práctica militar. Se generaron lentes y nuevos sensores (radar).

La época post guerra se puede considerar como el período de aplicación generalizada de la foto aérea y técnicas fotogramétricas para la cartografía. Se observa un desarrollo científico en este campo, plasmado en el número de publicaciones relacionados con la teledetección, el nacimiento de revistas y sociedades especializadas en el tema.

Durante las próximas décadas la teledetección aérea alcanzan su etapa de madurez. Al mismo tiempo nace una nueva forma de observar la tierra: **la teledetección espacial**. La guerra fría es la principal impulsora de este nacimiento.

La primera foto espacial se tomó en 1946 a bordo del cohete espacial V-2 (dentro de la atmósfera). En 1957 se lanza el primer satélite artificial: el **Sputnik 1** 2.6a. En 1959 se obtiene la primer foto espacial de la Tierra transmitida por el **Explorer 6**.

En los años 60, la NASA pone en marcha los primeros programas satelitales meteorológicos con el lanzamiento del **TIROS 1** (Television and Infrared Observation Satellite), el cual obtuvo 23.000 imágenes de diversas partes del planeta. Dentro de la década del 60 hay proyectos y misiones importantes como: el proyecto **Mercury**

(1960) y las misiones **Gemini** (1965) y **Apolo** (1969). Estos proporcionaron imágenes espaciales para el análisis de recursos terrestres. Esto motivo la creación, en 1966, del programa **EROS**(**E**arth **R**esources **O**bservation **S**ystem **P**rogram), el cual se materializa con la puesta en órbita, el día 29 de Julio de 1972, del primer satélite diseñado para el estudio de los recursos naturales: el ERTS-1 (Earth Resources Technology Satellite) rebautizado como **Landsat 1**. La puesta en órbita del ERTS-1 es el inicio de una nueva etapa en la teledetección: las fotografías aéreas se centran en las zonas visibles y del infrarrojo próximo del espectro electromagnético, mientras que los sensores a bordo de satélites extienden las bandas de infrarrojo térmico y próximo.

En la década del 70 el programa Landsat sigue en desarrollo (se lo considera uno de los más fructíferos de la historia de la teledetección espacial). Se destaca en esta década el lanzamiento de el laboratorio espacial tripulado **Skylab** (1973), el satélite meteorológico **GOES** (1974) y el satélite oceanográfico y meteorológico **Nimbus-7** (1978)

En la década del 80 se destacan: el programa francés con colaboración de diversos países europeos **SPOT** (Satellite Probatoire d'Observation de la Terre) en 1986, el satélite japonés **MOS-1** (1987) y el primer satélite de la serie hindu **IRS-1** (1988).

En los años noventa se ponen en órbita los primeros satélites con radares o sensores que trabajan sobre el rango espectral de las microondas.[10]

Los nuevos sensores están abocados a la obtención de imágenes de mayor precisión y específicas.

### 2.3. Importancia de la teledetección espacial

Las imágenes obtenidas por los satélites ofrecen una perspectiva única de la Tierra, sus recursos y el impacto que sobre ella ejercen los seres humanos. La teledetección espacial es una fuente de información para numerosas y diferentes actividades como: **el cuidado del medio ambiente, la urbanización, la explotación minera y petrolífera, entre otras.**

Las imágenes satelitales y la información que obtenemos de las mismas, ofrecen una visión global de objetos y detalles de la superficie terrestre, permitiendo com-



(a) Sputnik 1 en construcción



(b) TIROS 1 expuesto en museo



(c) Explorer 6



(d) Lanzamiento Mercury 6



(e) Gemini 6



(f) Apollo 1

Figura 2.6: Primeros Satélites

prender relaciones que no pueden verse claramente desde la Tierra. Muestran mucho más de lo que el ojo humano puede observar, detalles que estarían fuera de su alcance sin la participación de la teledetección espacial.

El valor práctico y la multiplicidad de aplicaciones de las imágenes continúan aumentando a medida que se lanzan nuevos satélites. Al haber más satélites se dispone de imágenes de diferentes tamaños, resoluciones espectrales, etc. Estos nuevos sensores hacen que las imágenes sean más útiles y específicas.

En la actualidad existen muchos satélites en órbita dedicados a captar imágenes de tipo muy específico. Casi todos los satélites comerciales de teledetección captan imágenes utilizando sensores digitales que funcionan según los mismos principios que las cámaras digitales. Los sensores satelitales cuentan con miles de detectores diminutos que miden la cantidad de radiación electromagnética que refleja la superficie de la Tierra y los objetos que hay en ella. Estas mediciones se denominan espectrales. Cada valor de reflectancia espectral se registra como un número digital. Estos números se transmiten de nuevo a la Tierra donde son transformados como colores o variaciones del gris para crear una imagen.

### 2.3.1. Beneficios de las Imágenes Satelitales

Las imágenes satelitales son más rápidas, mejores y menos costosas, que cualquier otra fuente de recolección de datos geográficos (estudios sobre terreno, fotografías aéreas, etc). Estas ventajas las podemos caracterizar de la siguiente manera:

**Son Digitales:** La gran mayoría de las imágenes satelitales se adquieren digitalmente, por lo que no debemos complicarnos con conversiones de datos. Con una preparación mínima, las imágenes están listas para ser procesadas para obtener información que en otras fuentes sería inaccesible.

**Se obtienen rápidamente y se transmiten de manera inmediata:** Un satélite es capaz de captar zonas enormes en muy poco tiempo. Como los mismos cuentan con órbitas estables, rara vez tardan más de una semana en obtener las imágenes necesarias.

Los sistemas de teledetección graban las imágenes en formato digital, lo que permite transmitir las imágenes a las estaciones terrestres en tiempo real. Esta dispo-

nibilidad inmediata resulta más que importante en casos de situaciones de emergencia, por lo que se considera una de las características más destacadas de la teledetección en la gestión de catástrofes.

**Son Económicas:** Para cubrir zonas extensas, las imágenes satelitales son mucho mas económicas que fotografías aéreas o campañas topográficas.

**Cobertura global, sinóptica, perspectiva panorámica y multiescala:** Los sensores ubicados en plataformas espaciales cubren prácticamente la totalidad de la superficie del planeta, por ello, nos proporcionan datos de zonas remotas y de difícil acceso, como océanos o regiones cubiertas de hielo donde, de no ser por esta técnica, la información sería difícil o casi imposible de conseguir.

La visión sinóptica de amplias zonas del planeta permite la observación de fenómenos (relieve, estructuras geológicas, dinámica litoral y marina, distribución de coberturas de suelo) que no serían identificables trabajando a escalas más detalladas. Estas ventajas son validas no solo para territorios geográficamente amplios, sino también para pequeñas aéreas de varios  $km^2$ , dado que se dispone cada día de mas imágenes de alta resolución.

La altura orbital del satélite le permite detectar grandes espacios, proporcionando una visión amplia de los hechos geográficos. Una fotografía aérea a su altura máxima con una escala de 1 : 30,000 capta en una sola imagen una superficie aproximada de  $49km^2$ . Una imagen del sensor Landsat-TM permite contemplar  $34,000km^2$  en una sola adquisición.

**Actualización y frecuencia:** En el mundo actual necesitamos información actualizada para tomar decisiones críticas para nuestros proyectos. Se puede disponer de una imagen satelital un par de días después de su toma.

Los satélites proporcionan datos de forma continua con una alta repetitividad. Los satélites geoestacionarios meteorológicos nos proporcionan una imagen cada 30 minutos, la serie NOAA cuatro imágenes diarias, los satélites de recursos SPOT y Landsat una imagen cada 26 y 16 días. La frecuencia temporal permite el estudio de fenómenos dinámicos(atmósfera, océano) y la detección o análisis de cambios de paisajes(incendios, inundaciones, desertificación, cambios en los

usos del suelo, evolución de cultivos, etc.), así como la actualización continua de cartografía existente. Del mismo modo, cada día más, la disponibilidad de series temporales largas de imágenes de satélites hace factible la realización de estudios estadísticos (evolución de la temperatura o la vegetación a escala global) o utilización de las imágenes como documento objetivo del pasado. El valor temporal de las imágenes cobra cada día más valor entre las singularidades que ofrece la teledetección espacial. Los archivos de imágenes pueden ser fácilmente consultados desde distintos lugares gracias al crecimiento de los sistemas de información y al desarrollo de Internet, mejorando sensiblemente el acceso al usuario.

**Homogeneidad y precisión de los datos:** Una imagen satelital en bruto se crea sin intervención humana, por lo que, la información que contiene es una representación precisa, objetiva e imparcial de los objetos y detalles de la superficie terrestre.

Los datos tomados por un sensor en todo el globo son datos especialmente consistentes, es decir, registrados por un mismo instrumento, mientras que los datos tomados sobre el terreno a menudo proceden de sistemas de medición distintos y no siempre son comparables. La homogeneidad de los datos tiene gran valor para la cartografía de zonas extensas y para los estudios de evolución que tratan series temporales largas.

---

## Capítulo 3

# Registro de Imágenes

### 3.1. Definición del Registro de Imágenes

Podemos definir de manera general al **Registro de Imágenes** como el proceso de alineación de dos o más imágenes, o de una o más imágenes con algún otro conjunto de datos (mapas por ejemplo).

Una imagen, se define de manera simple cómo una matrix de píxeles. La alineación entre varias imágenes se realiza mediante transformaciones matemáticas. Para poder ser registradas, las dos imágenes deberían reproducir mismos sectores para que puedan ser superpuestos. En el caso básico una imagen debería ser trasladada o, rotada y trasladada para poder ser alineada.

El proceso de registro de imágenes incluye localizar y alinear regiones similares en dos imágenes.

Teniendo en cuenta el motivo por el cual precisemos registrar las imágenes, podemos clasificar el registro como:

**Registro Multimodo:** Permite la integración de datos de diferentes sensores.

**Registro Temporal:** Integra datos obtenidos por distintos sensores en un período de tiempo determinado.

**Registro de Puntos de Vistas:** Integra datos generados por distintos puntos de vista. El objetivo es obtener representaciones 2D o 3D más grandes de la escena escaneada.

**Registro de Modelos:** Busca la correspondencia entre datos nuevos y conjuntos de datos previamente obtenidos o generados.

Dependiendo de quién lleve a cabo estas tareas, podemos dividir al registro de imágenes en dos grandes categorías:[6]

**Registro manual:** Una persona se encarga del proceso. Utiliza un software interactivo. Los usuarios extraen de ambas imágenes puntos distintivos llamados *Puntos de Control (PC)*. Luego los PC correspondientes son utilizados para calcular los parámetros de las transformaciones geométricas.

**Registro automático:** Algoritmos autónomos llevan a cabo las tareas.

En teledetección, los procedimientos automatizados no siempre ofrecen la fiabilidad y la precisión necesaria, por lo que el registro manual se utiliza con frecuencia. Los usuarios extraen de ambas imágenes puntos distintivos, llamados puntos de control o puntos de referencia. En primer lugar, los puntos de control de ambas imágenes se *matchean* interactivamente de a pares para lograr la correspondencia. Luego estos puntos que se corresponden se utilizan para calcular los parámetros de las transformaciones geométrica a utilizar. La selección manual de estos PC es una tarea repetitiva y requiere mucho tiempo, por lo que resulta casi imposible para cantidades grandes de datos. Esta elección manual en imágenes satelitales genera puntos de control inexactos o mal distribuidos, lo que podría dar lugar a grandes errores de registro.

Por ende el objetivo principal de la investigación en el registro de imágenes es mejorar eficiencia, precisión y robustez de los algoritmos autónomos.

Generalmente, los algoritmos automáticos de registro de imágenes incluyen tres pasos principales:[1]

1. Extracción de regiones o puntos distintivos para ser alineados.
2. Matcheo o alineación de los puntos, buscando la transformación geométrica correspondiente.
3. Construcción de una nueva imagen, con la alineación generada entre las imágenes a registrar.

Los enfoques automatizadores, varían en como tratan cada uno de estos pasos. Algunos algoritmos usan técnicas simples para la extracción de características, pero usan estrategias complejas para el matcheo correspondientes; otros algoritmos son más complejos para la selección de los puntos característicos, pero los algoritmos de matcheo son mucho más simples.

### 3.1.1. Elementos de los Algoritmos de Registro de Imágenes

Como se enunció anteriormente, una de las instancias del problema de registro de imagen consta en alinear una imagen, *Imagen Detectada*, con respecto a una segunda imagen, *Imagen de Referencia*. En particular, podemos definir el problema de registro de imagen teniendo en cuenta cinco elementos:[6]

**Imagen de Referencia:**  $I_1(x, y)$ , generalmente no se modifica.

**Imagen Detectada:**  $I_2(x, y)$ , es transformada para poder ser matcheada o alineada con la imagen de referencia.

**Transformaciones Geométricas:** Son el conjunto de transformaciones matemáticas utilizadas para transformar las posiciones espaciales de la imagen detectada, y poder alinearla con la imagen de referencia. Serán representadas por  $f$ .

**Transformaciones Radiométricas:** Representadas con la letra  $g$ , conjunto de funciones que transforman los valores de intensidad de una imagen en los valores de la otra.

**Término de Ruido:**  $n(x, y)$ , modela el ruido que pueda haber sufrido la imagen.

Basándonos en los elementos definidos anteriormente, podemos determinar una relación entre la imagen de referencia y la imagen detectada. Suponiendo que  $(x, y)$  representa el sistema de coordenadas de la imagen de referencia, y  $(u, v)$  el de la imagen detectada, podemos decir que una posición de la imagen de referencia se **corresponde** con una posición en la imagen detectada (*Puntos Correspondidos*) si:

$$I_1(x, y) = g(I_2(f_x(u, v), f_y(u, v))) + n(x, y) \quad (3.1)$$

Como podemos ver en (3.1), el problema de registro de imágenes consta en calcular las transformaciones geométricas y radiométricas para que los puntos puedan ser alineados. De manera más general, podemos decir que el el registro de imágenes consta en desarrollar un algoritmo que tome un par de imágenes, y calcule las transformaciones *óptimas* para generar el matcheo correspondiente. La definición óptima va a depender fundamentalmente de las **medidas de similitud** escogidas.

Debido a la diversidad de imágenes a registrar y la variedad de aplicaciones, es imposible diseñar un método universal. Cada transformación geométrica o radiométrica depende de la precisión necesaria, de las características y de los datos del problema a resolver.

Sin embargo, la mayoría de los métodos de registro se basan en los pasos siguientes: [4]

**Detección de Características:** Objetos sobresalientes y distintivos (regiones cerradas o límites, bordes, curvas, intersecciones de líneas, esquinas, etc ) son elegidos de forma manual o detectados automáticamente. Para el tratamiento posterior, estas características pueden ser representados por sus puntos representantes (centros de gravedad , los finales de líneas, puntos distintivos) llamados puntos de control.

**Matcheo de Características:** En este paso, se establece la correspondencia entre las características elegidas en la imagen detectada y de referencia. Varios incluyen descriptores y medidas de similitud para llevar a cabo este paso.

**Estimación de Modelos de Transformación:** Tanto el tipo y los parámetros de las denominadas funciones de mapeo, como la alineación de la imagen detectada con la imagen de referencia, son estimados. Los parámetros de las funciones de mapeo se calculan por medio de la función de correspondencia establecida.

**Transaformaciones Radiométricas y Construcción de Imagen:** La imagen detectada es transformada por medio de las funciones de mapeo. Los valores de la imagen en coordenadas se calculan por la técnica de interpolación apropiada, para de esta manera generar una nueva imagen, la cual es el resultado de alinear las imágenes de referencia y detectada.

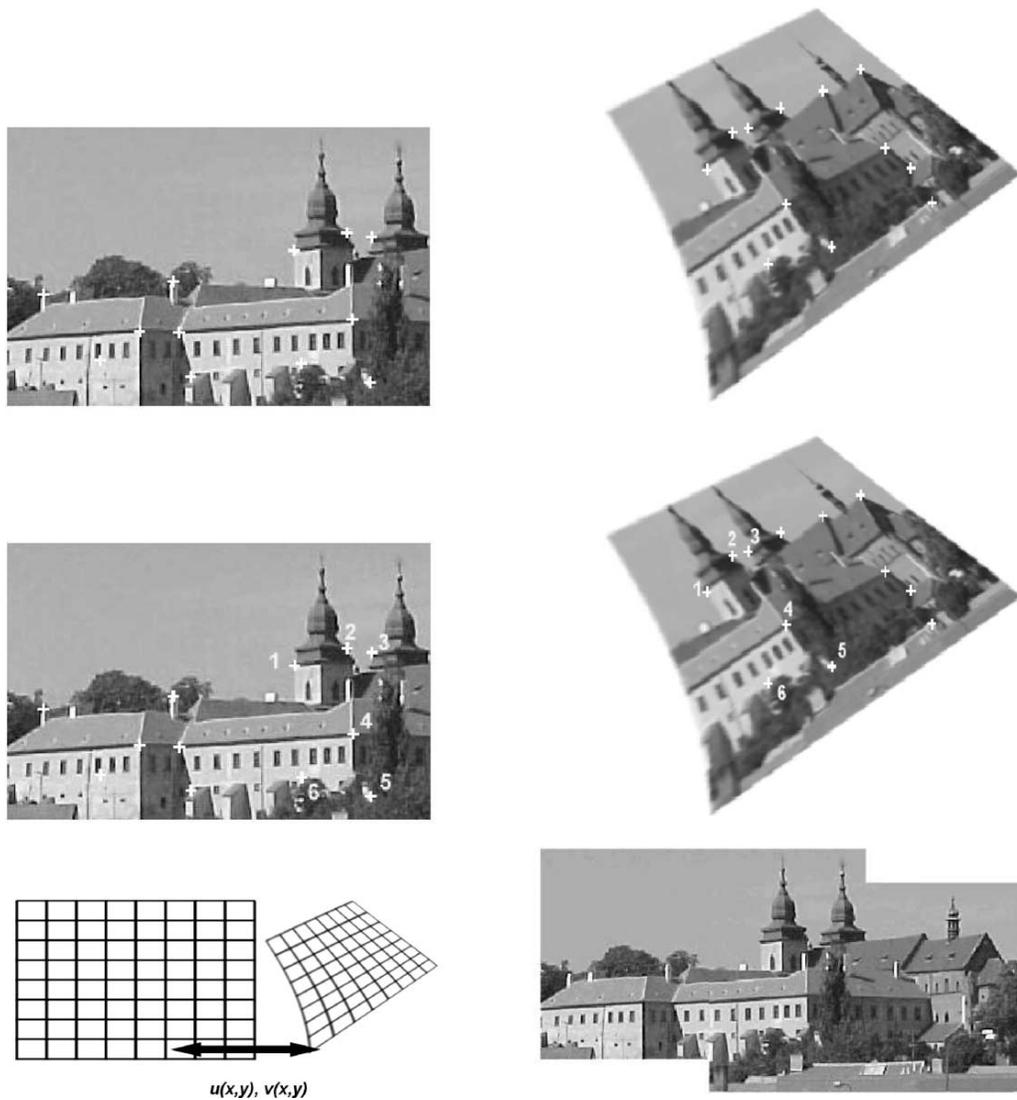


Figura 3.1: Proceso de Registro de Imágenes

La implementación de cada uno de estos pasos tiene sus problemas típicos e incidencias en pasos posteriores. Primero, tenemos que elegir cuales serán las características adecuadas para la tarea que debemos llevar a cabo. Estas características, deberían de ser objetos distintivos, esparcidos frecuentemente sobre las imágenes, y fácilmente detectables. Los conjuntos de características de la imagen de referencia y la imagen detectada deben tener suficientes elementos en común, también en situaciones donde las imágenes no cubran la misma escena en totalidad o cuando hay

obstrucciones de objetos o cambios inesperados.

Los métodos de detección deben tener una buena precisión de localización y no deben ser sensibles a la degradación de la imagen. **En un caso ideal, el algoritmo debe ser capaz de detectar las mismas características en todas las proyecciones de la escena, independientemente de la deformación de la imagen.**

En el paso de mapeo de características, cualquier problema causado para una detección de objetos distintivos errónea o por degradaciones de imágenes puede afectar el desempeño. Físicamente, puntos característicos correspondientes, pueden ser diferentes bajo condiciones distintas de las imágenes o por la sensibilidad espectral de los sensores. Por lo que, la elección del descriptor y la medida de similitud debe considerar estos factores. Los *descriptores de características* deberían ser invariantes a las degradaciones asumidas. Al mismo tiempo, tienen que ser capaces de distinguir entre diferentes características, así como suficientemente estable para no ser influido por ligeras variaciones generadas por puntos característicos inesperados o por el ruido generado en la imagen.

El tipo de las funciones de mapeo debería ser escogido dependiendo de la información previa que tengamos del proceso de adquisición de la imagen y de la degradación esperada. En el caso de no contar con información previa, el modelo debería de ser flexible y lo suficientemente general como para manejar todas las posibles degradaciones que pudiesen ocurrir. Debemos de considerar la precisión del método de detección de características, la confiabilidad del estimador de correspondencia y el error de aproximación.

## 3.2. Un poco de historia del Registro de Imágenes

La palabra **Registro** se remonta al año 1900 patentada por Becker en los Estados Unidos. En esta patente, Becker desarrollo una cámara de enfoque 3.2 que utiliza un medio-espejo y un espejo completo para crear dos imágenes de una escena en la retina del espectador. Al cambiar la orientación de un espejo con respecto a la otra, las imágenes de los dos espejos quedan alineadas. El mecanismo que cambia la orientación de un espejo con respecto al otro, también cambia la distancia del plano

de la imagen a la lente. Está diseñado de tal manera que cuando las imágenes de los dos espejos se alinean perfectamente en la retina del espectador, la película también se mueva a la distancia correcta de la lente, lo que permite la creación de una imagen nítida de la escena en la película.

El siguiente avance en el registro de imágenes se observa en la industria del cine, en la producción de una película de color de doble capa. En un aparato descrito por Kelley y Mason patentado en EE.UU. en 1917, una película de color se crea mediante el registro de la banda azul en un lado de la película y una combinación de bandas de color rojo y verde en el lado opuesto de la película. El reto en la creación de una película de color estaba dado por imprimir en la película dos imágenes, de la misma escena, en perfecta alineación. El aparato descrito logra la alineación de diferentes bandas de color, con alta precisión, utilizando dos pernos de alineación que encajan en las perforaciones en los bordes de las películas.

El siguiente uso de registro de imágenes se dió en la impresión. Seaman patentó en los EE.UU. en 1937, un aparato para la impresión generada por la superposición de imágenes de una escena, capturadas por separado. La invención no sólo permitió la creación de un mosaico de imágenes a partir de dos o más imágenes superpuestas, sino también generó un mecanismo de que hizo posible añadir datos u objetos capturados de una escena a otra totalmente distinta.

El registro de imagen como lo conocemos hoy en día, surgió como un dispositivo electrónico de comparación de imágenes 3.3. Dressler en una patente en 1956, da a conocer un dispositivo electrónico para la comparación y la combinación de dos imágenes superpuestas. Las imágenes eran proporcionadas a través de películas. A través de la utilización de un medio espejo y un espejo completo, las imágenes se proyectan a un circuito de sustracción de imágenes que produce una salida con una amplitud proporcional al grado de coincidencia entre las imágenes. Al proporcionar los medios para traducir y girar una película con respecto a la otra, y mediante la observación de la salida de los circuitos, la capacidad de alinear y combinar imágenes a partir de dos películas en una imagen más grande fue proporcionado.

La idea de la comparación de imágenes de manera electrónica se extendió más tarde a un correlacionador de imágenes. Enlace y Smith, patentaron en 1962, un tubo electrónico 3.4 que podría almacenar una imagen de referencia y correlacionarlo con

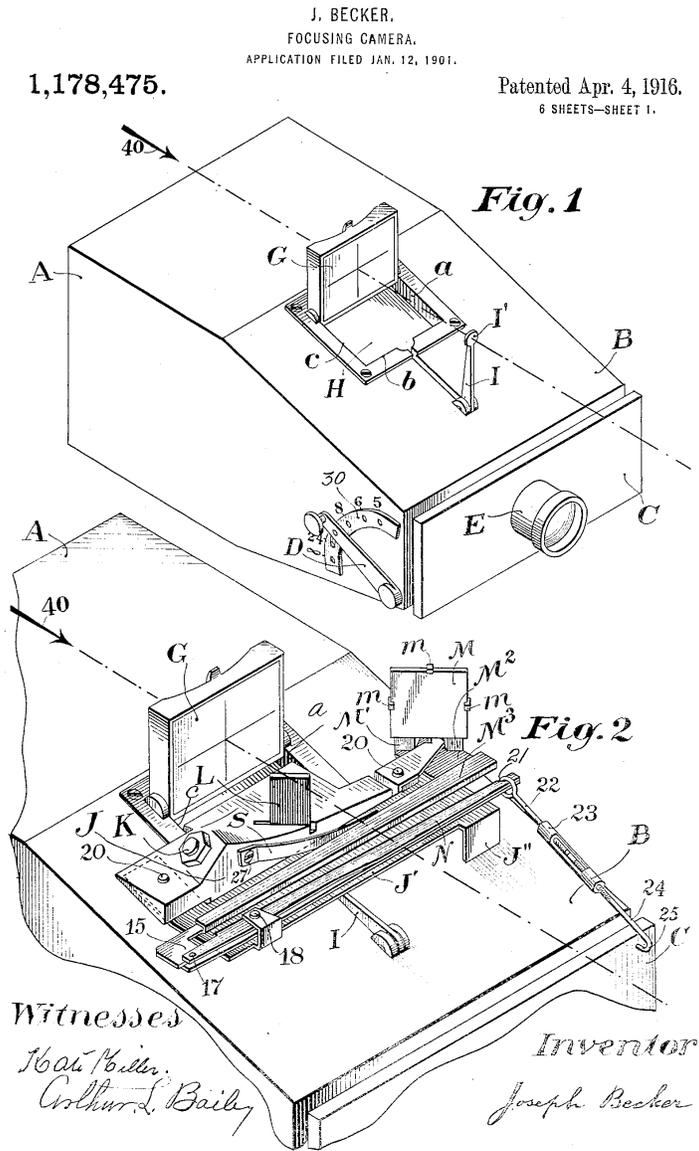


Figura 3.2: Patente cámara de enfoque de Becker

June 27, 1961

R. DRESSLER

2,989,890

IMAGE MATCHING APPARATUS

Filed Nov. 13, 1956

8 Sheets-Sheet 4

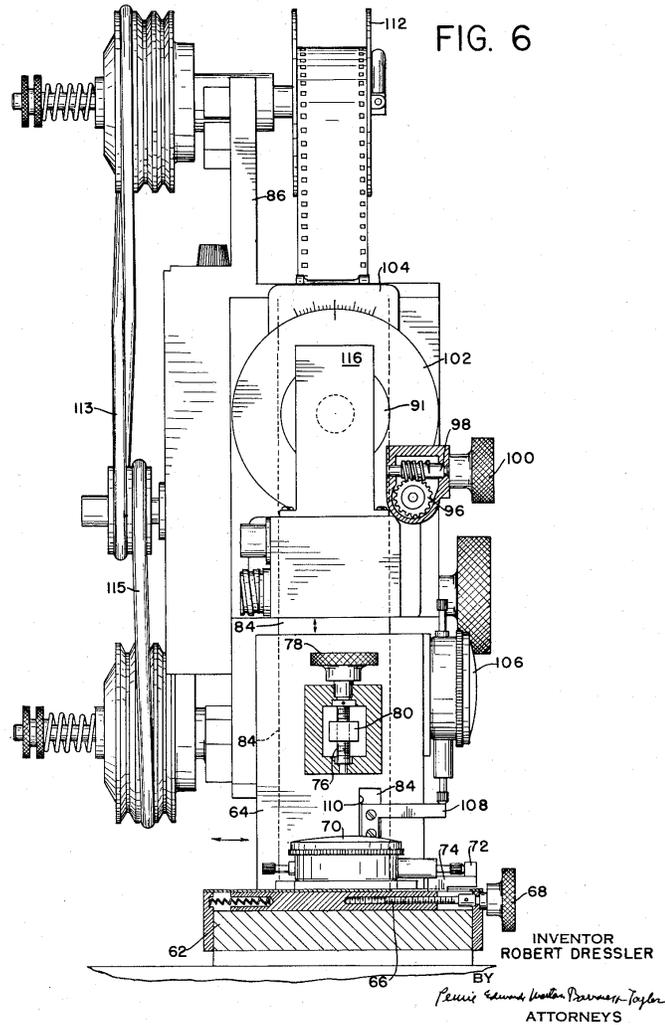


Figura 3.3: Patente Dispositivo de Comparación de Dressler

una imagen detectada, produciendo una señal de salida con su resistencia en función de la correlación entre las imágenes. Steiner en 1965, avanzó aún más el tubo para que sea posible mover la imagen detectada con respecto a la imagen de referencia y alinear las imágenes. Johnson, en 1969, mejora el aparato para normalizar la medida de correlación de manera que la amplitud de salida sea independiente de la amplitud de la imagen detectada, midiendo de ese modo el grado de correspondencia entre las imágenes referencia y detectada con más precisión.

El primer ejemplo de registro de imágenes digitales se remonta al trabajo de Roberts en el MIT en 1963. Mediante la alineación de las proyecciones de los bordes de poliedros con bordes de la imagen, Roberts desarrolló un método de cálculo para la localización y el reconocimiento de poliedros predefinidos en una imagen.

El registro de imágenes digitales apareció por primera vez en la literatura de tele-detección. Si bien las actividades de registro durante la década de 1970 se centraron en la alineación de imágenes satelitales, el aumento de la producción de imágenes aéreas de baja altitud durante la década de 1980, fue la fuerza impulsora detrás de la invención de métodos que podrían alinear espacialmente imágenes con diferencias geométricas locales.

Debido al incremento en el uso de imágenes médicas en 1990 y la necesidad de alinear imágenes multimodales, se lograron avances considerables en la formulación de la teoría de información y con respecto a las medidas de similitud (disimilitud) que podrían comparar y registrar este tipo de imágenes.

Durante la última década, debido al aumento del uso de los vídeos en la vigilancia y otras aplicaciones en tiempo real, el registro de imágenes se convirtió en un paso necesario en el análisis de los vídeos de una o varias fuentes. Los avances en la tecnología de imágenes para aumentar la resolución y la calidad de las imágenes, también aumentó la complejidad de los métodos que pueden registrar este tipo de imágenes. [4]

### **3.3. Importancia del Registro de Imágenes**

El registro de imágenes es crucial para todas aquellas tareas de análisis de imágenes que precisan de la información generada por la combinación de diferentes imágenes.

Dec. 6, 1966

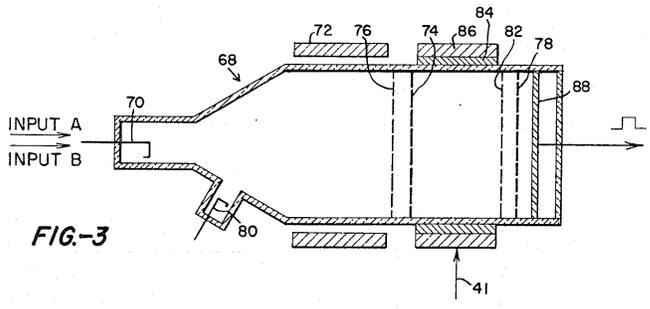
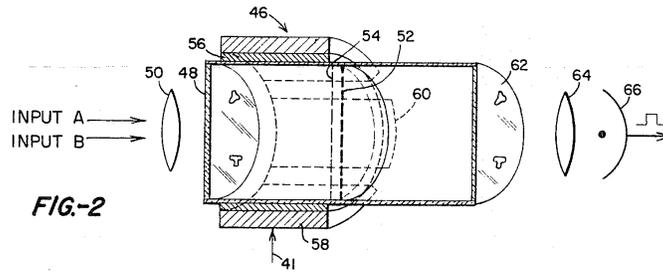
T. V. LINK ETAL

3,290,546

CATHODE RAY TUBE IMAGE MATCHING APPARATUS

Filed Oct. 25, 1962

2 Sheets-Sheet 2



INVENTORS:  
THOMAS V. LINK  
RICHARD H. SMITH  
BY *[Signature]*  
ATTORNEY

Figura 3.4: Patente Tubo Rayos Catódicos para Matcheo de Imágenes desarrollado por Enlace y SMITH

nes. Es utilizado en medicina, teledetección, cartografía, robótica, entre muchas otras disciplinas. Dependiendo el problema, depende el enfoque y técnicas que se utilizan para llevar a cabo este proceso.

En medicina se utiliza para combinar tomografías con NMR para obtener información más completa sobre el paciente, monitorear crecimiento de tumores, verificación de tratamientos, comparación de datos del paciente con atlas de anatomía, etc.

Los estudios científicos de la tierra a menudo tratan de cuestiones tales como predecir el rendimiento del cultivo, evaluar el cambio climático, la localización de las tierras de cultivo y de fuentes de agua, monitoreo de la contaminación, y hasta la comprensión del impacto de la actividad humana sobre los principales Ecosistemas de la Tierra.

Para abordar estas cuestiones, los científicos utilizan las mediciones globales proporcionadas por una amplia variedad de sistemas de teledetección por satélite (no estarían disponibles con equipos terrestres o sensores aéreos). Estas mediciones globales a menudo necesitan ser complementadas por medidas locales o regionales para generar una más completa investigación de los fenómenos observados.

En los últimos años, el *registro de imágenes* se ha transformado en una herramienta de vital importancia en teledetección y el procesamiento de imágenes. Esto se debe principalmente a que la teledetección comenzó a tener aplicaciones importantes a **nivel social** además de científico. Por ej:

- Manejo de desastres y recursos naturales,
- Asesoramiento para cambios climáticos,
- Preservación y Cuidado del medio ambiente,
- Planificación Urbana.

Como podemos ver, son aplicaciones que ameritan un monitoreo constante de la superficie terrestre, por lo que el énfasis en el registro de imágenes.

Además, ha aumentado la disponibilidad de imágenes con diferentes características, debido a la velocidad de rastreo de los satélites, distintas modalidades de obtención de datos y la evolución de los sensores.

Muchas de las aplicaciones del registro de imágenes afectan nuestra vida cotidiana, por lo que necesitamos que el Registro sea *Preciso*. Tengamos en cuenta que afecta temas como, la predicción de lluvias, evaluación del cambio climático en grandes escalas temporales, localización de fuentes de agua, etc.

Esta tarea se ha vuelto crítica por varias razones:[6]

- Juega un papel esencial en la calibración espacial y radiométrica de mediciones multitemporales para la obtención de grandes conjuntos de datos integrados para el seguimiento a largo plazo de diversos fenómenos.
- La detección de cambios en el tiempo o escala sólo es posible si los datos obtenidos por múltiples sensores y a través del tiempo están precisamente integrados.
- La capacidad resultante de la extrapolación de los datos a lo largo de varias escalas.

### 3.4. Enfoques Generales del Registro de Imágenes

Los métodos de registro de imágenes pueden clasificarse dependiendo de muchos criterios: área de aplicación, dimensión de datos, tipo y complejidad de las imágenes, costo computacional y las ideas esenciales del algoritmo de registro. En este caso nos basaremos en las ideas esenciales del algoritmo, haciendo hincapié en cada uno de los pasos.

Los algoritmos de registro de imágenes generalmente son clasificados como:

**Basados en Intensidad:** En estos métodos, las áreas o regiones de la imagen original son alienados con un mínimo pre procesamiento, el cual preserva la mayoría de la imagen. Estos algoritmos pueden calcular las diferencias de los valores de los píxeles de la imagen cruda, o utilizar los valores de píxel para calcular una representación completa intermedia como los coeficientes de Fourier. Los métodos basados en intensidad a menudo son llamados de **Menor Correspondencia de Matcheo**, ya que toda una zona puede ser matcheada sin la necesidad de construir una correspondencia explícita entre los puntos de las dos imágenes.

**Basados en Características:** Las imágenes originales son pre procesadas para obtener información distintiva y características que serán utilizadas en el proceso de matcheo. Estos algoritmos extraen puntos de control, detectan bordes o contornos, etc.

### 3.4.1. Detección de Características

Como dijimos anteriormente, los **Métodos Basados en Intensidad** no realizan pre procesamiento de las imágenes, por ende no detectan puntos característicos, por lo que este primer paso no forma parte en los métodos basados en intensidad.

Por otra lado, para los **Métodos Basados en Características**, este primer paso es fundamental. Regiones significantes (bosques, lagos), líneas (límites de región, rutas, ríos) o puntos (esquinas, intersecciones de líneas, curvas) son entendidos como *características*. Las características seleccionadas deben: poder distinguirse, estar esparcidos en toda la imagen y ser fácilmente detectables, tanto en la imagen de referencia como en la detectada.

La comparabilidad de los conjuntos de características de las imágenes detectadas y de referencia está asegurada por la invariancia y la precisión del detector de características y por el criterio de superposición. En otras palabras: el número de elementos comunes de los cambios geométricos de las imágenes, las condiciones radiométricas, la presencia de ruido aditivo, y de los cambios de la escena detectada. Los métodos basados en características no funcionan directamente con los valores de intensidad de la imagen. Las características representan la información en un nivel superior. Esta propiedad hace que los métodos basados en características sean adecuados para situaciones en las que se esperan cambios de iluminación o el análisis multisensor es requerido.

**Regiones:** Pueden ser las proyecciones de alto contraste de regiones de fronteras cerradas de un tamaño adecuado, embalses y lagos, los edificios, los bosques, las zonas urbanas o sombras. El criterio de las regiones de fronteras cerradas es muy utilizado. Las regiones son representadas por sus centros de gravedad, ya que son invariantes con respecto a la rotación, la escala y estables bajo ruido aleatorio y la variación de niveles de gris.

Las características como regiones son detectadas mediante *Métodos de Segmentación*. Estos métodos dividen a una imagen en varias partes.

**Líneas:** Pueden ser representaciones de segmentos lineales, contornos de objetos, líneas costeras, caminos, o en el caso de las imágenes médicas, estructuras anatómicas. La correspondencia de línea se expresa con los pares de fin de líneas o los puntos medios de las mismas.

Los métodos estandar como el detector de Canny o el Laplaciano de Gauss son utilizados para captar estas líneas como características.

**Puntos:** Los grupos de puntos característicos consisten en métodos que trabajan con intersecciones de líneas, cruces de rutas, centros de regiones acuáticas, petróleo o gas, puntos con alta varianza, esquinas, etc.

### 3.4.2. Matcheo o alineación de características

Las características detectadas tanto en la imagen de referencia como en la imagen detectada pueden ser matcheadas mediante la intensidad de los valores de la imagen en los vecinos más cercanos, distribuciones espaciales de la característica o descripciones simbólicas de las mismas.

Recordemos que los **Métodos Basados en Intensidad** trabajan con las imágenes sin procesamiento previo, es decir, no se detectan puntos, líneas ni regiones características. A estos métodos se los suele llamar de *Correlación*. Ventanas de tamaños predefinidos o la imagen completa son utilizadas para estimar la correspondencia en este segundo paso de registro.

Estos métodos tienen limitaciones generadas por su idea básica. La *Ventana Rectangular* (la más utilizada), permite registrar únicamente aquellas imágenes que difieren espacialmente por traslaciones. Si las imágenes tiene diferencias generadas por transformaciones mucho más complejas, este tipo de ventana no será capaz de cubrir de la misma manera ambas imágenes, por lo que se recomienda alguna otra forma. Se propone el uso de una ventana circular para imágenes que se diferencien entre sí por no solo traslaciones sino también por rotaciones.

Otra desventaja de los métodos basados en intensidad es la alta probabilidad que una ventana contenga un área borrosa sin detalles de importancia, y de la misma

manera sea matcheada con regiones borrosas de la imagen de referencia, que nada tengan que ver.

Los métodos clásicos basados en intensidad realizan el matcheo utilizando directamente las intensidades de las imágenes, sin ningún análisis previo. Por ende, son sensibles a los cambios de intensidad introducidos por ruido, variaciones de iluminación o uso de diferentes sensores.

Podemos observar que para realizar el matcheo precisamos de medidas de similitud o disimilitud, para comparar las regiones de la imagen detectada con la de referencia. Estas medidas también pueden ser utilizadas en los métodos basados en características.

## Medidas de Similitud Y Disimilitud

En el registro de imágenes la precisión de los resultados depende en gran medida de la medida de similitud empleada. Para el registro de imágenes se emplean muchas medidas de similitud. Estas medidas pueden ser calculadas sobre imágenes completas o partes de las mismas.

Tradicionalmente, la limitada información sobre el comportamiento de una medida de similitud era obtenida estudiando la calidad del registro final o dibujando gráficas de similitud de los valores de medida obtenidos por rotación o translación de la imagen. Hoy día estamos asistiendo a procedimientos de optimización mediante la evaluación estadística de las medidas de similitud. Estos tipos de técnicas estiman entre otras cosas el rango de captura, el número y localización de un óptimo local y la precisión de un óptimo global, para obtener el mejor resultado. No hay una única medida de similitud que pueda producir un mejor resultado que otra para todo tipo de imágenes, de tal modo que la búsqueda de factores de similitud constituye un área de investigación relevante en el mundo del procesado de la señal.

A grandes rasgos, las medidas de similitud están divididas en dos tipos: **medidas basadas en métodos geométricos**, que utilizan para ello la extracción de distintas características, y **medidas basadas en niveles de intensidad**. Éstas últimas a su vez también se encuentran divididas en dos tipos: las de promedios y las de tipo estadístico. La división efectuada está directamente relacionada con el tipo de algoritmo de registro llevado a cabo.

En el pasado, los métodos de registro se han basado mayoritariamente en emparejar las características correspondientes de las imágenes a registrar, pero más recientemente, el interés se ha dirigido a las medidas de la correspondencia global obtenida directamente a partir de las intensidades de la imagen.

Los registros de imágenes basados en medidas de similitud ajustan los parámetros de una transformación espacial apropiada, hasta que la medida en cuestión alcanza un óptimo. Dadas las imágenes a ser registradas y el modelo de transformación espacial, el resultado de la tarea de registro principalmente depende de la medida de similitud y del método de optimización.

**Definición 1.** *Formalmente definimos, dadas dos secuencias de mediciones:  $X = \{x_i : i = 1, \dots, N\}$  e  $Y = \{y_i : i = 1, \dots, N\}$ , la similitud (disimilitud) entre estas variables es una medida que cuantifica la dependencia (independencia) entre las mismas.  $X$  e  $Y$  pueden representar mediciones a partir de dos objetos o fenómenos. En este caso representarán imágenes y  $x_i$  y  $y_i$  son las intensidades de los píxeles correspondientes de las mismas.*[4]

**Definición 2.** *Una medida de similitud  $S$  se considera una métrica si a medida que la dependencia entre las variables aumenta, el valor generado por la medida es mayor.*

Sea  $S$ , una métrica de similitud, entonces satisface:

**1. Rango Limitado:**

$$S(X, Y) \leq S_0, \text{ para algún } S_0.$$

**2. Reflexividad:**

$$S(X, Y) = S_0 \text{ si y solo si } X = Y.$$

**3. Simetría:**

$$S(X, Y) = S(Y, X).$$

**4. Desigualdad Triangular:**

$$S(X, Y)S(Y, Z) \leq [Z(X, Y) + S(Y, Z)]S(X, Z).$$

$S_0$  es el mayor valor posible entre  $X$  e  $Y$ .

Una medida de disimilitud  $\mathbf{D}$  se considera métrica, si a medida que disminuye la dependencia entre variables, disminuye el valor generado por la medida.

Sea  $D$  una métrica de disimilitud entonces:

1. **No Negativa:**  $D(X, Y) \geq 0$
2. **Reflexividad:**  $D(X, Y) = 0$  si y solo si  $X = Y$
3. **Simetría:**  $D(X, Y) = D(Y, X)$
4. **Desigualdad Triangular:**  $D(X, Y) + D(Y, Z) \geq D(X, Z)$

**Destaquemos que no es necesario que una medida cumpla con ser métrica para poder satisfacer los requisitos en el registro de imágenes.**

### Clasificación Algoritmos de Matcheo Basados en Intensidad

Podemos clasificar los algoritmos de matcheo de los métodos basados en intensidad como:

#### Métodos basados en Correlación:

Calculan directamente la medida de similitud (disimilitud) de las regiones de las imágenes mediante la comparación de la intensidad de los píxeles. Los mismos son denominados también como matcheo de templates. Una región de una imagen es trasladada sobre la otra imagen, para poder encontrar la alineación que optimice la medida de similitud (disimilitud) utilizada. A pesar de ser fundamentales en el registro de imágenes, tienen una serie de inconvenientes por lo que necesitan ser optimizados para diferentes aplicaciones.

Primero que nada, consideremos que el enfoque de fuerza bruta (comparar las ventanas escogidas alrededor de toda la imagen de referencia) tiene un alto valor computacional. Suponiendo que el tamaño de la imagen de referencia es  $N$  y el de la ventana escogida es  $M$  tenemos una complejidad de  $O(N^2 M^2)$ . Para ello, a lo mejor es conveniente aplicarlo en dominios de frecuencia. También se puede optimizar usando calculos parciales, paralelización, optimizaciones numéricas, entre otros.

### Métodos basados en la transformada de Fourier:

Los métodos basados en el dominio de la frecuencia, son la primera alternativa ante el cálculo rápido de las medidas de similitud. Una *Transformación de Frecuencia*, convierte una señal digital o imagen, en una colección de coeficientes de frecuencia que representan la intensidad de esa frecuencia en la señal original. Los métodos de Fourier se basan en el **Teorema de Desplazamiento de Fourier**, el cual establece que sea  $g(x, y)$  la versión trasladada de la señal  $f(x, y)$  en el dominio espacial  $xy$ , luego las transformadas de Fourier correspondientes,  $G(u, v)$  y  $F(u, v)$  en el dominio de frecuencia  $uv$  están relacionadas mediante un desplazamiento de fase que puede ser recuperado eficientemente. Para el proceso de registro de imagen, las dos imágenes son transformadas al dominio de frecuencia mediante *Transformada Rápida de Fourier*. Estas transformaciones son multiplicadas eficientemente en el dominio de frecuencia, mientras que su producto puede ser transformado al dominio espacial (mediante la Transformada Rápida de Fourier Inversa) para la recuperación de la traslación.

### Información Mutua y enfoques basados en distribución:

Que uno de los requisitos sea la igualdad entre las intensidades de las imágenes es una restricción fuerte, la cual no es válida en muchos casos. Estos métodos proponen evitar esta restricción; las medidas de similitud, se basan en comparar las intensidades locales de las distribuciones en vez de los valores de los píxeles.

La medida de similitud que es más efectiva bajo esta propuesta ha sido la *Información Mutua (MI)* y sus posibles variantes. El registro basado en Información Mutua fue introducido inicialmente para las imágenes médicas.

Estos métodos tienen precondiciones débiles sobre la relación entre las intensidades de los píxeles de ambas imágenes, y no requieren de un modelo monótono o funcional. La Información Mutua se encarga de medir la dispersión de los valores en un histograma conjunto, representado mediante una matriz de dos dimensiones. Por lo que, si las dos imágenes coinciden perfectamente, sus intensidades se deberían agrupar alrededor de la diagonal de la matriz del histograma conjunto; en caso contrario los valores deberían estar dispersos

de la diagonal.

Considerando el conjunto de intensidades de las imágenes, representados por las variables  $X$  e  $Y$ , definimos la información mutua entre ambas como:

$$MI(X, Y) = \sum_{x \in X} \sum_{y \in Y} p_{X,Y}(x, y) \log \left( \frac{p_{X,Y}(x, y)}{p_X(x)p_Y(y)} \right) \quad (3.2)$$

donde:

- $p_X(x)$  y  $p_Y(y)$  representan la *La Función de Densidad de Probabilidad (FDP)*, de  $X$  e  $Y$ .
- $p_{X,Y}(x, y)$  representa la Función de Densidad de Probabilidad Conjunta de  $X$  e  $Y$ .

Recordemos que la **FDP** de una variable aleatoria continua describe la probabilidad relativa según la cual dicha variable aleatoria tomará determinado valor.

### Clasificación Algoritmos de Matcheo Basados en Características

Asumiendo que los conjuntos de características de las imágenes de referencia y detectada han sido seleccionadas como puntos de control, pasamos a la descripción de los diferentes **Métodos Basados en Características** para el matcheo. Podemos clasificarlos como:

#### Basados en Relaciones Espaciales:

Son usualmente utilizados si las características detectadas son ambiguas. La información de la distancia entre los Puntos de Control y su distribución espacial es usada. Siguiendo este alineamiento, encontramos métodos que utilizan algoritmos de matcheo de grafos. En los mismos es evalúa el número de características de la imagen detectada, que luego de una transformación cae en algún grupo

#### Basados en Descriptores Invariantes:

La correspondencia de características se puede estimar usando su descripción, preferentemente invariante a la deformación esperada de la imagen. Estos descriptores deben de satisfacer estas propiedades: *invariancia* (las descripciones de las características correspondientes de las imágenes referencia y detectadas tiene que ser la misma ), la *singularidad* (dos características diferentes deben tener diferentes descripciones), *estabilidad*(la descripción de una característica que se deforma debe estar cerca de la descripción de la función original), e *independencia* (si la descripción de la función es un vector, los elementos que deben ser funcionalmente independientes). Por lo general no todas estas condiciones deben (o pueden) ser satisfechas al mismo tiempo y es necesario encontrar un equilibrio adecuado.

Las características de las imágenes de referencia y detectada, con la descripción invariante más similar, son emparejadas. La elección del tipo de descriptor invariante depende las características seleccionadas de las imágenes como la deformación geométrica de las mismas.

### 3.4.3. Estimación de Modelos de Transformación

Después que la correspondencia de características se ha establecido se construye la función de mapeo. Se debe transformar la imagen detectada para superponer encima de la imagen de referencia. La correspondencia de los Puntos de Control de las imágenes de referencia y detectada como el hecho de que los pares de PC correspondientes deben estar tan cerca como sea posible después de la transformación de la imagen detectada, son utilizadas en el diseño de la función de mapeo.

Para resolver esta tarea debemos escoger el tipo de función de mapeo y la estimación de sus parámetros. El tipo de esta función, va a depender de la deformación geométrica asumida de la imagen detectada, el medio de donde se consiguió la misma y la precisión requerida para el registro. Si tenemos el conocimiento previo suficiente sobre el modelo del sensor que obtuvo la imagen, distorsiones geométricas generadas por la órbita satelital y la variación de altitud, efectos atmosféricos, podemos construir un modelo de transformación preciso y un algoritmo apropiado para el mismo podría ser elegido. En el caso en que no contáramos con toda esta información podríamos utilizar un modelo empírico estimado.

Consideremos las coordenadas de  $n$  puntos correspondientes:

$$\{(x_i, y_i), (X_i, Y_i) : i = 1, \dots, n\}$$

queremos determinar la función de transformación con componentes  $f_x$  y  $f_y$  tal que:

$$X_i \approx f_x(x_i, y_i) \quad (3.3)$$

$$Y_i \approx f_y(x_i, y_i) \quad (3.4)$$

Teniendo en cuenta las diferencias entre la imagen de referencia con la imagen detectada podemos definir estas funciones de transformación como:

### 1. Rotación, Escala y Traslación:

Estas transformaciones cuentan con cuatro parámetros, y suelen ser un subconjunto útil para las transformaciones afines.

Dependiendo de las diferencias entre las imágenes de referencia y detectada, pueden tener que llevarse a cabo solo traslaciones, rotaciones o modificaciones de escala. En este caso vamos a considerar cuando sean necesarias estas tres transformaciones sobre la imagen detectada. Diremos que las coordenadas de los puntos correspondientes estarán relacionadas mediante una *Transformación de Similitud*, definida como:

$$X = xs \cos \theta - ys \sin \theta + x_0 \quad (3.5)$$

$$Y = xs \sin \theta + ys \cos \theta + y_0 \quad (3.6)$$

donde  $s$  define la escala,  $\theta$  la orientación y  $(x, y)$  la localización del sistema de coordenadas de la imagen detectada para relacionarlo con la imagen de referencia.

Los parámetros  $s$ ,  $\theta$ ,  $x_0$  y  $y_0$  precisan de al menos dos puntos correspondientes entre las imágenes para poder ser determinados. Si contamos con mas dos

puntos correspondientes entre las imágenes los parámetros serán determinados mediante estimadores robustos.

## 2. Transformaciones Afines:

Estas transformaciones cuentan con seis parámetros, y podemos escribirla como la combinación de transformaciones lineales y traslaciones:

$$X = a_1x + a_2y + a_3 \quad (3.7)$$

$$Y = a_4x + a_5y + a_6 \quad (3.8)$$

Estos modelos a pesar de ser más generales, continúan siendo lineales como se puede observar. Para definir el mismo, precisamos de tres puntos correspondientes no co-lineales. Es utilizado para el registro de imágenes obtenidas desde diferentes puntos de vista.

Dependiendo de la *Cantidad de Datos* de la imagen que utilicemos podemos dividirlos en dos grandes categorías:

- **Modelos Globales:**

Usan todos los **PC** para poder estimar el conjunto de parámetros válidos para toda la imagen, de la función de mapeo.

- **Modelos Locales:**

Las funciones de mapeo local, toman a la imagen como composición de diferentes partes, y los parámetros de las mismas dependen del sector de la imagen.

### 3.4.4. Transformaciones Radiométricas y Construcción de imagen (*Resampling*)

Bajo las transformaciones geométricas correctas, dos puntos de imagen correspondientes deberían de tener valores de intensidad similares. Si asumimos que estos puntos representan los mismos sectores o características del terreno, la interpretación radiométrica de los mismos debería de ser *consistente*.

Esta consistencia mezcla múltiples factores, como lo son las respuestas radiométricas de dos sensores distintos, la parte del día y diferentes factores del ambiente que hayan influido en el momento en que han sido detectadas las imágenes, el ángulo de vista, entre otros.

Si podemos tener un recuento de todos estos factores, podemos construir un modelo radiométrico o una función  $g$  que logre mapear las intensidades de una imagen sobre la otra, usando el mismo para el algoritmo de registro. Sin embargo, pocos algoritmos de registro usan relaciones radiométricas para el proceso de matcheo. Esto se debe a que son relaciones difíciles de medir y generalizar ya que al algoritmo muchas veces será aplicado a varios sensores por lo que, precisaremos de varias relaciones. Además los factores del ambiente pueden dificultar demasiado la elección de las mismas, generando una gran complejidad en la elección y creación de estos modelos. Por lo tanto, la mayoría de los algoritmos de registro de imagen basan sus medidas de similitud, con pre condiciones sobre su relación radiométrica, las cuales pueden ser **explícitas** o **implícitas**.

Según Roche(2000) las asuncpciones sobre el modelo radiométrico a utilizar puede ser clasificada como:

1. **Relación de Identidad:** Asume que la intensidad de imagen es invariante. Esta asunción es explícita en métodos que se basan en medidas de similitud, las cuales suman directamente el cuadrado de las diferencias de los valores absolutos de la diferencia entre las intensidades.
2. **Relaciones Afines:** Asumen que los intensidades difieren por alguna función lineal. La asunción siguiente es implícita en métodos basados en coeficientes de correlación o minimización de los cuadrados menores.
3. **Relaciones Funcionales:** Asumen que las intensidades difieren por una función general. La misma es implícita en varias medidas de similitud.
4. **Relaciones Estadísticas:** Asumen que puntos correspondientes pueden diferir en sus intensidades, pero los neighborhoods de los cuales forman parte deben de tener la misma distribución estadística. Las mismas son implícitas en información mutua y enfoques similares.

Existen algoritmos que integran relaciones radiométricas explícitas en el proceso de matching.

Desde que las imágenes digitales son representadas por valores en una matriz, sabemos que una imagen es geoméricamente transformada para ser matcheada con la imagen de referencia. Ños valores deben ser localizadas en nuevas posiciones de la matriz o grilla. Esto se realiza con dos propósitos:

1. Obtener imagen final luego del proceso de registro de imágenes para otras aplicaciones futuras.
2. Producir imágenes intermedias durante el proceso de registro de imágenes, en el caso de utilizar métodos iterativos por ejemplo.

Un sistema de registro de imágenes debe utilizar múltiples técnicas para la construcción (*Resampling*) de la imagen final. En los dos casos mencionados anteriormente, es importante evitar el remuestreo de elementos que degraden el conjunto de datos o la precisión del registro.

El remuestreo se realiza comúnmente mediante muestreo inverso de valores de la imagen para volver a asignar la imagen detectada a una nueva imagen. Si  $f(u, v)$  es la transformación geométrica que matchea la imagen detectada con la imagen de referencia, luego la transformación inversa  $f^{-1}(x, y)$  es aplicada para mapear un píxel en la nueva imagen detectada.

Podemos clasificar los métodos de armado de imagen como:

#### **Vecino más cercano:**

Usa el valor del píxel más cercano para asignar el valor del pixel de salida 3.5.

Para determinar el píxel de salida más cercano, las coordenadas rectificadas  $(x, y)$  del pixel son retransformadas al sistema de coordenadas fuente usando la inversa de la transformación. Las coordenadas retransformadas  $(u, v)$  son usadas en la interpolación bilineal y en la convolución cúbica también. El pixel que está más cercano a las coordenadas retransformadas  $(u, v)$  es el vecino más cercano. Los niveles digitales de ese pixel se convierten en los niveles digitales de la imagen de salida.

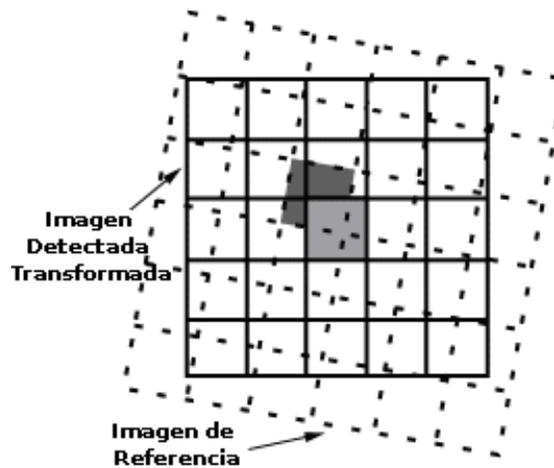


Figura 3.5: Elección de valor de pixel, mediante Vecino más Cercano

### Interpolación Bilineal:

Usa los niveles de intensidad de cuatro píxeles en una ventana de  $2 \times 2$  para calcular el valor del pixel salida mediante una función lineal.

En la interpolación bilineal, el valor del pixel se basa en las distancias entre la posición de la coordenada retransformada  $(u,v)$  y los cuatro píxeles más cercanos en la imagen de referencia 3.6.

### Convolución Cúbica:

Usa los niveles de intensidad de 16 píxeles en una ventana de  $4 \times 4$  para calcular el valor de pixel de salida con una función cúbica 3.7.

La convolución cúbica es similar a la interpolación bilineal, excepto que: Un conjunto de 16 píxeles, en un arreglo de  $4 \times 4$ , se promedia para determinar el valor del pixel de salida, y se aproxima mediante una función cúbica, en lugar de lineal, aplicando la misma a 16 valores de entrada.

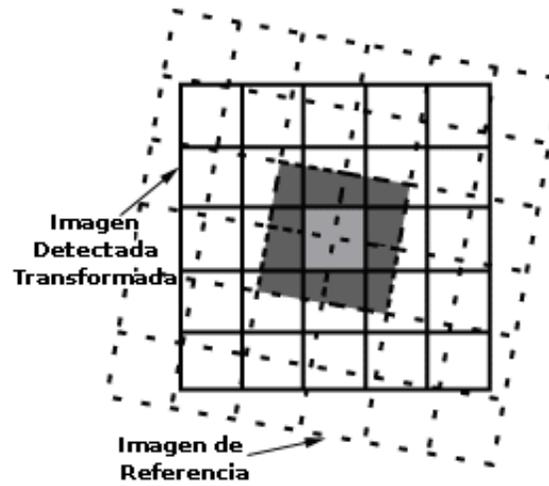


Figura 3.6: Elección de valor de pixel mediante Interpolación Bilineal

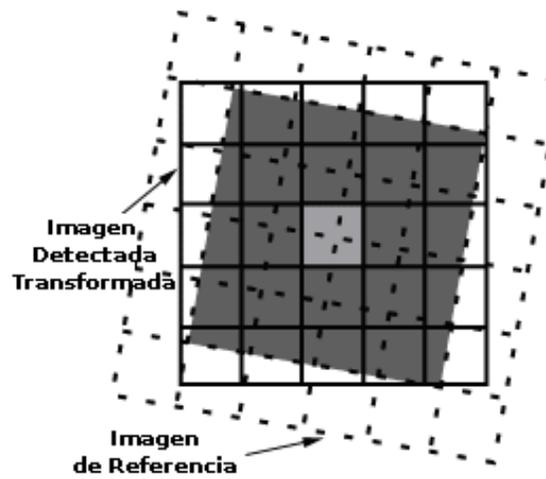


Figura 3.7: Elección de valor de pixel mediante Convolución Cúbica

---

## Capítulo 4

# Algoritmo automático de la transformada rápida de Fourier

Teniendo en cuenta lo estudiado en diferentes libros y papers, se pudo observar que los algoritmos más utilizados para el registro automático de imágenes de tele-detección son los basados en características. Por lo que se decidió ir en contra de la corriente y trabajar sobre métodos basados en intensidad. En particular se escogió un algoritmo basado en la transformada de Fourier.

Estos métodos son una de las principales alternativas rápidas para calcular coeficientes de correlación entre imágenes. Estas transformaciones de frecuencia, convierten una señal digital o imagen, en un conjunto de coeficientes que representan la fuerza de cada frecuencia en la señal original. Los métodos de transformadas de registro de imágenes se basan en la premisa de que a partir de la transformada será mucho más sencillo detectar las transformaciones geométricas necesarias para que ambas imágenes sean alineadas.

### 4.1. Fundamentos teóricos

#### 4.1.1. Transformada de Fourier y sus propiedades

Es posible implementar técnicas de procesamiento y análisis de imágenes usando la representación de datos de imágenes digitales en el “dominio de las frecuencias”,

o transformada de Fourier de la imagen.

En esta sección se enunciarán propiedades y definiciones con respecto a las transformadas de Fourier, las cuales serán de mucha importancia para el proceso de registro de imágenes.

**Definición 3.** Sea  $f(x)$  una función continua de la variable real  $x$ . La transformada de Fourier de esta función representada por  $F(f(x))$ , está definida como:

$$F(f(x)) = F(u) \quad (4.1)$$

$$= \int_{-\infty}^{\infty} f(x)e^{-i2\pi ux} dx \quad (4.2)$$

donde  $F(u)$  es denominado el componente de frecuencia  $u$  e  $i$  es la raíz cuadrada de  $(-1)$ .

Dada una transformada de Fourier  $F(u)$  podemos volver a calcular el valor de  $f(x)$  utilizando la transformada inversa de Fourier la cual definimos como:

$$F^{-1}(u) = \int_{-\infty}^{\infty} F(u)e^{-i2\pi ux} du \quad (4.3)$$

Estas dos definiciones se pueden extender fácilmente a funciones bidimensionales  $f(x, y)$ , tal que la transformada de Fourier,  $F(u, v)$ , y su inversa,  $F^{-1}u, v$ , quedan definidas como:

$$F(f(x, y)) = F(u, v) \quad (4.4)$$

$$= \int_{-\infty}^{\infty} f(x, y)e^{-i2\pi(ux+vy)} dx \quad (4.5)$$

$$F^{-1}(u, v) = \int_{-\infty}^{\infty} F(u, v)e^{-i2\pi(ux+vy)} du \quad (4.6)$$

Enunciamos algunas propiedades interesantes de la transformada de Fourier, las cuales serán necesarias para la implementación del algoritmo de registro de imágenes.

**Definición 4.** Sea  $f(x, y)$  una señal y  $F(u, v)$  su transformada de Fourier. Supongamos que trasladamos la señal por el vector  $(x_0, y_0)$ , entonces se cumple que:

$$f(x - x_0, y - y_0) \Leftrightarrow F(u, v)e^{-i2\pi(ux_0 + vy_0)} \quad (4.7)$$

Esto se denomina la **propiedad de traslación** de la transformada de Fourier.

**Definición 5.** Sea  $f(x, y)$  una señal y  $F(u, v)$  su transformada de Fourier. Si introducimos coordenadas polares tal que:

$$x = r \cos \theta \quad (4.8)$$

$$y = r \sin \theta \quad (4.9)$$

$$u = w \cos \sigma \quad (4.10)$$

$$v = w \sin \sigma \quad (4.11)$$

convertimos  $f(x, y)$  y  $F(u, v)$  en  $f(r, \theta)$  y  $F(w, \sigma)$  respectivamente, cumpliendo la siguiente propiedad:

$$f(r, \theta + \theta_0) \Leftrightarrow F(w, \sigma + \theta_0) \quad (4.12)$$

Esto se denomina la **propiedad de rotación** de la transformada de Fourier.

**Definición 6.** Sea  $f(x, y)$  una señal y  $F(u, v)$  su transformada de Fourier. Sean  $a$  y  $b$  dos escalares, entonces se cumplen las siguientes propiedades:

$$af(x, y) \Leftrightarrow aF(u, v) \quad (4.13)$$

$$f(ax, by) \Leftrightarrow \frac{1}{|ab|} F(u/a, v/b) \quad (4.14)$$

Para el registro de imágenes, las mismas son transformadas al dominio de frecuencia mediante la *transformada rápida de Fourier*. Luego estas transformadas son trabajadas eficientemente en el dominio de frecuencias, y mediante la transformada rápida inversa de Fourier, se transforma el resultado al dominio espacial, consiguiendo la traslación necesaria para alinear ambas imágenes. Este desplazamiento se encuentra en la posición donde se encuentra el valor máximo de la inversa de la transformada.

### 4.1.2. Coeficiente de correlación en el dominio de Fourier

Supongamos que dados dos vectores unidimensionales (con valores discretos), queremos encontrar el mejor matcheo posible entre ambos.

Veamos un pequeño ejemplo práctico. Consideremos  $x = [1, 1, -1, 1, -1, -1, 1]$  e  $y = [1, -1, 1]$ . Como podemos ver deberemos de comparar  $y$  en cinco posiciones distintas de  $x$ . Comparamos elemento por elemento, y contaremos la cantidad de casos en los que haya igualdad (**I**) y la cantidad de casos en la que haya desigualdad (**D**). Calculamos **I** – **D**, para definir cual es la mejor posición para la alineación de los vectores, obteniendo el siguiente vector  $[-1, 3, -3, 1, 1]$ . Como podemos observar el máximo valor se encuentra en la posición 1 (comenzando el conteo de posiciones en 0), lo cual nos indica que el mejor matcheo ocurre cuando trasladamos  $y$ , 1 posición con respecto a  $x$ .

El proceso de comparar vectores componente por componente para identificar cuando los mismos están alineados se denomina *correlación*.

Como pudimos ver en el ejemplo, para calcular la correlación entre los vectores, restábamos al número de aciertos los desaciertos. Esta medida es válida para datos binarios pero no para lo que necesitamos nosotros que son las imágenes. Por suerte, en el mundo de las estadísticas contamos con una medida que nos permite expresar la relación entre dos vectores numéricos, cuyos elementos pueden diferir a través de una transformación,  $x_i = ay_i + b$ , donde  $a$  y  $b$  son constantes. Esta función es conocida como *coeficiente de correlación*.

**Definición 7.** Sean  $\mathbf{x} = [x_0, x_1, \dots, x_{N-1}]$  e  $\mathbf{y} = [y_0, y_1, \dots, y_{M-1}]$ , vectores de longitud  $N$  y  $M$  respectivamente. Definimos el coeficiente de correlación entre  $\mathbf{x}$  y  $\mathbf{y}$  como un vector  $\text{corrcoef}(\mathbf{x}, \mathbf{y})$  de longitud  $(N - M + 1)$ , tal que cada  $j$ -ésimo componente ( $j = 0, 1, \dots, N - M + 1$ ) está dado por:

$$\text{corrcoef}(\mathbf{x}, \mathbf{y})_j = (\mathbf{x} \bullet \mathbf{y})_j \quad (4.15)$$

$$= \frac{\left( \sum_{i=0}^{M-1} x_{i+j} y_i - (1/M) \sum_{i=0}^{M-1} x_{i+j} \sum_{i=0}^{M-1} y_i \right)}{\sqrt{\left( \sum_{i=0}^{M-1} x_{i+j}^2 - (1/M) \left( \sum_{i=0}^{M-1} x_{i+j} \right)^2 \right) \left( \sum_{i=0}^{M-1} y_i^2 - (1/M) \left( \sum_{i=0}^{M-1} y_i \right)^2 \right)}} \quad (4.16)$$

La función es normalizada para que los valores resultantes esten en el rango de  $[-1, +1]$ .

A pesar de ser de gran ayuda para el cálculo de matcheo entre dos vectores, en el caso de tener que matchear grandes imágenes, el proceso será muy lento, a menos que trabajemos en algún dominio matemático que permita aprovechar propiedades del mismo para aumentar velocidad en este proceso. En particular, nos referimos al dominio de Fourier.

Para ello debemos de poder calcular el coeficiente de correlación en este dominio, lo cual es posible gracias a propiedades de la transformada de Fourier.

**Definición 8.** Sean  $\mathbf{x}$  e  $\mathbf{y}$  vectores de longitud  $N$ . Sea  $(\mathbf{x} \otimes \mathbf{y})$  la **convolución circular** entre  $\mathbf{x}$  e  $\mathbf{y}$  es definida por:

$$\text{circonv}(\mathbf{x}, \mathbf{y})_j = (\mathbf{x} \otimes \mathbf{y})_j \quad (4.17)$$

$$= \sum_{i=0}^{N-1} x_{(j-i) \bmod N} y_i. \quad (4.18)$$

A partir de (8), podemos enunciar el **teorema de convolución**:

**Teorema 1.** Sean  $\mathbf{x}$  e  $\mathbf{y}$  vectores de longitud  $N$ . Sean  $F(\mathbf{x})$ ,  $F(\mathbf{y})$  las transformadas de Fourier de  $\mathbf{x}$  e  $\mathbf{y}$ . Sean  $F(\mathbf{x})(u)$  y  $F(\mathbf{y})(u)$  los componentes  $u$  de la transformadas fourier correspondientes entonces:

$$F(\mathbf{x} \otimes \mathbf{y})(w) = F(\mathbf{x})(u) \cdot F(\mathbf{y})(u) \quad (4.19)$$

Como podemos ver a partir del teorema de convolución se facilitaría el cálculo. El problema es que el registro de imágenes depende de correlaciones y no de convoluciones. Lo positivo es que tanto las correlaciones como las convoluciones tienen la misma estructura, salvo pequeñas diferencias. A partir de esto definimos la **correlación circular**:

**Definición 9.** Sean  $\mathbf{x}$  e  $\mathbf{y}$  vectores de longitud  $N$ . Sea  $(\mathbf{x} \circ \mathbf{y})$  la **correlación circular** entre  $\mathbf{x}$  e  $\mathbf{y}$ , la cual es definida por:

$$\mathbf{circcorr}(\mathbf{x}, \mathbf{y})_j = (\mathbf{x} \circ \mathbf{y})_j \quad (4.20)$$

$$= \sum_{i=0}^{N-1} x_{(j+i) \bmod N} y_i. \quad (4.21)$$

Luego podemos enunciar el **teorema de correlación**. Tengamos en cuenta que la diferencia entre la correlación y convolución circular está dado por el valor del índice con el cual obtendremos el elemento de  $\mathbf{x}$  con el cual haremos el cálculo (en particular es un cambio de signo). Esta diferencia genera que precisemos de la conjugada compleja de la transformada de Fourier.

**Teorema 2.** Sean  $\mathbf{x}$  e  $\mathbf{y}$  vectores de longitud  $N$ . Sea  $F(\mathbf{x})$  la transformada de Fourier de  $\mathbf{x}$  y  $F^*(\mathbf{y})$  la conjugada compleja de la transformada de Fourier de  $\mathbf{y}$ . Sean  $F(\mathbf{x})(u)$  y  $F^*(\mathbf{y})(u)$  los componentes  $u$  de la transformadas, entonces:

$$F(\mathbf{x} \otimes \mathbf{y})(w) = F(\mathbf{x})(u) \cdot F^*(\mathbf{y})(u) \quad (4.22)$$

El teorema de correlación es muy importante para el registro de imágenes, ya que que la correlación entre imágenes es uno de los pasos centrales de este proceso. Para correlacionar dos imágenes, cada una con  $N$  píxeles, necesitamos de  $N^2$  operaciones. Gracias al teorema de correlación, podemos llevar a cabo el proceso de correlación en el dominio de Fourier, donde precisamos de  $N$  operaciones para encontrar la correlación circular entre las dos imágenes. En general los métodos basados en las transformadas de Fourier tienen un orden de  $N \log N$ .

Para poder terminar de calcular el coeficiente de correlación en el dominio de Fourier debemos de poder calcular las operaciones del mismo mediante correlaciones circulares, respetando que los vectores a correlacionar deben de tener la misma longitud.

Basándonos en la definición 7, considerando los vectores  $\mathbf{x}$  y  $\mathbf{y}$  de la misma, tenemos un primer problema para poder pasar el problema al dominio de Fourier: los vectores  $\mathbf{x}$  e  $\mathbf{y}$  pueden tener longitudes diferentes. Por suerte se puede resolver con facilidad. Antes que nada supongamos que se cumple que  $N > M$ . Para que las longitudes sean iguales *rellenamos* el vector  $\mathbf{y}$  con ceros hasta que el mismo tenga

longitud  $N$ .

$$\mathbf{y} = [y_0, y_1, \dots, y_{M-1}, 0, 0, \dots, 0] \quad (4.23)$$

También creamos un nuevo vector  $\mathbf{m}$  de longitud  $N$ , en el cual representaremos con 1 los valores válidos de  $\mathbf{y}$  y con 0 los agregados para que tenga longitud  $N$ . También podemos utilizar la máscara  $\mathbf{m}$  para representar los valores que existían en el vector  $\mathbf{y}$ , pero cuyo valor es inválido por diferentes motivos.

De la misma manera podemos crear un vector máscara  $\mathbf{h}$  para el vector  $\mathbf{x}$  el cual represente los valores válidos e inválidos del vector  $\mathbf{x}$ .

Una vez definidos estos nuevos vectores, nos enfocaremos en calcular en forma de correlaciones circulares la fórmula (7).

El término  $\sum_{i=0}^{M-1} x_{i+j}y_i$ , es tal cual la definición de correlación circular dada en la ecuación (7) por lo que:

$$(x \circ y)_j = \sum_{i=0}^{M-1} x_{i+j}y_i \quad (4.24)$$

Ahora consideremos el término

$$\sum_{i=0}^{M-1} y_i \quad (4.25)$$

, que calcula la suma de los elementos de  $\mathbf{y}$ . Teniendo en cuenta que queremos calcular la correlación con el vector  $\mathbf{x}$ , entonces tendría mas importancia sumar los valores de  $\mathbf{y}$  válidos en  $\mathbf{x}$ . Por lo tanto podemos transformar la ecuación (4.25) como:

$$\sum_{i=0}^{M-1} y_i = \sum_{i=0}^{M-1} h_{j+i}y_i \quad (4.26)$$

$$= (h \circ y)_j \quad (4.27)$$

También podemos realizar algo parecido con el término

$$\sum_{i=0}^{M-1} x_{j+i} \quad (4.28)$$

, donde sumaremos únicamente los valores donde  $\mathbf{y}$  sea válido. Por lo que podemos escribir el término 4.28:

$$\sum_{i=0}^{M-1} x_{j+i} = \sum_{i=0}^{M-1} x_{j+i} m_i \quad (4.29)$$

$$= (x \circ m)_j \quad (4.30)$$

Otro elemento que cambiará su valor es el divisor  $M$ . Como únicamente relacionaremos los valores válidos de ambos vectores, no será  $M$  la cantidad de elementos por la que deberemos promediar. Utilizando las máscaras  $\mathbf{h}$  y  $\mathbf{m}$  podemos calcular con facilidad la cantidad de elementos que utilizaremos:

$$\sum_{i=0}^{M-1} h_{i+j} m_i = (h \circ m)_j \quad (4.31)$$

Por lo tanto podemos definir el **coeficiente de correlación** en términos de correlaciones circulares:

$$(x \bullet y)_j = \frac{((x \circ y)_j - (1/(h \circ m)_j)(x \circ m)_j(h \circ y)_j)}{\sqrt{((x^2 \circ m)_j - (1/(h \circ m)_j)((x \circ m)_j)^2)((h \circ y^2)_j - (1/(h \circ m)_j)((h \circ y)_j)^2)}} \quad (4.32)$$

Entonces es posible calcular el coeficiente de correlación en el dominio de Fourier. Todas estas propiedades y definiciones aplicadas para dominios espaciales unidimensionales, se aplican de la misma manera para bidimensionales.

Además podemos utilizar la propiedad de **correlación de fase**, para el cálculo de los valores de traslación.

**Definición 10.** Sean  $f(x, y)$  y  $g(x, y)$ , imágenes con valores discretos y reales. Supongamos que las mismas difieren por una traslación  $(x_0, y_0)$ . Sean  $F(u, v)$  y  $G(u, v)$  las transformadas discretas de Fourier de  $f(x, y)$  y  $g(x, y)$  respectivamente. Por propiedad de traslación de Fourier 4,  $F(u, v)$  y  $G(u, v)$  difieren idealmente por el factor de fase  $e^{-i2\pi(ux_0+vy_0)}$ . Sea  $G^*(u, v)$  la conjugada compuesta de  $G(u, v)$ . Definimos la **correlación de fase** como la transformada inversa de Fourier de:

$$\frac{F(u, v) \cdot G^*(u, v)}{\|F(u, v)\| \cdot \|G(u, v)\|}. \quad (4.33)$$

Debido a que  $f(x, y)$  y  $g(x, y)$  difieren por  $(x_0, y_0)$ , sabemos que por la propiedad de traslación 4, sus transformadas de Fourier satisfacen que  $G(u, v) = F(u, v) \cdot e^{-i2\pi(ux_0+vy_0)}$ . Por este motivo es que la correlación de fase es definida como la transformada de Fourier inversa de:

$$\frac{F(u, v) \cdot G^*(u, v)}{\|F(u, v)\| \cdot \|G(u, v)\|} = e^{i2\pi(ux_0+vy_0)}. \quad (4.34)$$

La transformada inversa de (4.34), nos devuelve la función de impulso  $\delta(x - x_0, y - y_0)$ , la cual es aproximadamente 0 en todos los valores menos en el punto  $(x_0, y_0)$  donde idealmente el valor debería ser aproximadamente 1.

Como observamos en el sección 4.4.2, teníamos diferentes modelos de transformaciones para poder lograr el registro de imágenes. Para el algoritmo que desarrollaremos, trabajaremos con imágenes que difieren en escala, ángulo de rotación y traslaciones.

Para estudiar como cada una de estas transformaciones, puede ser calculada en el dominio de Fourier, dividiremos en dos casos generales. [8] Consideremos  $f(x, y)$  y  $g(x, y)$  las imágenes a registrar, y  $F(u, v)$  y  $G(u, v)$  sus transformadas de Fourier.

### 1. Sin cambio de escala:

Consideremos  $g(x, y)$  una replica de  $f(x, y)$  con una traslación de  $(x_0, y_0)$  y rotación de  $\theta_0$ . Las propiedades de Fourier de rotación y traslación, son muy importantes en este punto. Esto se debe ya que nos aseguran, que estas transformaciones no interactúan entre sí en el dominio de Fourier. La importancia de esto lo podemos reflejar de la siguiente manera:

*Cuando aplicamos transformaciones de rotación y traslación en una imagen, los píxeles cambian dependiendo de ambas operaciones. Algo totalmente distinto pasa en el dominio de Fourier. La traslación de una imagen no altera la magnitud del componente de frecuencia, por lo que podemos encontrar la rotación examinando las magnitudes.*

Aprovechando este beneficio podemos calcular de manera secuencial la rotación y traslación. Para calcular el ángulo de rotación utilizaremos la propiedad de rotación de la transformada de Fourier (5). Consideremos  $f(r, \theta)$  y  $g(r, \theta)$  su representación polar. Luego:

$$f(r, \theta) = g(r, \theta - \theta_0). \quad (4.35)$$

Por la propiedad de rotación de Fourier y tomando  $M_1$  y  $M_2$  como las magnitudes de  $F$  y  $G$ :

$$M_1(r, \theta) = M_2(r, \theta - \theta_0). \quad (4.36)$$

Utilizando la correlación de fase podemos calcular fácilmente el valor de  $\theta_0$ .

Finalmente sobre la nueva imagen  $g$  con la rotación exacta, aplicamos la correlación de fase para calcular los valores de  $(x_0, y_0)$

## 2. Con cambio de escala:

Sea  $g(x, y)$  una replica de  $f(x, y)$ , la cual difiere en escala por los factores  $a$  y  $b$ . Por la propiedad de escala de la transformada de Fourier sabemos que las mismas están relacionadas por:

$$G(u, v) = \frac{1}{|ab|} F(u/a, v/b). \quad (4.37)$$

Luego transformando el dominio espacial hacia el dominio logarítmico podemos reducir el problema de localizar la escala a una simple traslación (ignoremos el factor  $\frac{1}{|ab|}$ ):

$$G(\log u, \log v) = F(\log u - \log a, \log v - \log b) \quad (4.38)$$

$$G(w, z) = F(w - c, z - d). \quad (4.39)$$

tomando  $w = \log u, z = \log v, c = \log a$  y  $d = \log b$ .

De esta manera podemos aplicar la definición de correlación de fase, para obtener el vector de traslación  $(c, d)$ , computar los valores de escala  $a$  y  $b$ , cómo:  $a = e^c$  y  $b = e^d$ .

En el caso en que  $g(x, y)$  sea una replica trasladada, rotada y con cambio de escala de  $f(x, y)$ , las magnitudes espectrales de sus transformadas de Fourier en representación **log polar** están relacionadas por:

$$M_1(r, \theta) = M_2(r/a, \theta - \theta_0) \quad (4.40)$$

$$M_1(\log r, \theta) = M_2(\log r - \log a, \theta - \theta_0) \quad (4.41)$$

$$M_1(q, \theta) = M_2(q - c, \theta - \theta_0) \quad (4.42)$$

Donde,  $q = \log r$  y  $c = \log a$ .

Luego podemos aplicar en (4.42) la correlación de fase y obtener los valores de  $a$  y  $r$ .

Finalmente, aplicamos las transformaciones de escala y rotación a la imagen  $g$  y calculamos mediante correlación de fase nuevamente, los valores de  $(x_0, y_0)$

### 4.1.3. Algoritmo y su implementación

Utilizando las propiedades y conceptos anteriores se implementó un algoritmo basado en la transformada de Fourier para el registro de imágenes, las cuales podían diferir en escala, traslación y rotación.

Para la implementación del mismo se decidió utilizar el lenguaje de programación **Python**. Se eligió este lenguaje debido a que contiene librerías Numerical Python (NumPy) y Scientific Python (SciPy), las cuales son muy eficientes para las operaciones entre vectores y permiten manipular con facilidad los mismos. Para la manipulación de las imágenes, ya sea como entrada o salida, se utilizó la librería (PIL) ya que permite un manejo sencillo de estas actividades. Además el objetivo de este trabajo es poder contribuir con la comunidad del software libre, agregando nuevas herramientas en el campo del procesamiento de imágenes.

### Algoritmo de la transformada de Fourier

A continuación se explicitará el algoritmo utilizado. El mismo toma como entrada la imagen de referencia ( $I_1$ ) y la imagen detectada ( $I_2$ ), cuyos tamaños pueden ser distintos. Devuelve, una nueva imagen, la cual es el resultado del registro, además los valores del vector de traslación  $(x_0, y_0)$ , el ángulo de rotación  $(\theta_0)$  y el valor de la escala  $(s)$ .

1. Comparamos el tamaño de las imágenes. En el caso en que no sean iguales, *rellenamos* las imágenes con valores 0 y creamos máscaras que nos indiquen las posiciones de los valores válidos e inválidos.
2. Aplicamos las transformadas rápidas de Fourier a  $I_1$  e  $I_2$ , obteniendo  $F_1(u, v)$  y  $F_2(u, v)$ .
3. Aplicamos filtro pasa-alta (*high pass*) a  $F_1(u, v)$  y  $F_2(u, v)$  para remover ruidos de baja frecuencia.
4. Transformamos las coordenadas de  $F_1(u, v)$  y  $F_2(u, v)$  rectangulares a log-polares.
5. Calculamos los valores absolutos de  $F_1(u, v)$  y  $F_2(u, v)$ , obteniendo  $|F_1(u, v)|$  y  $|F_2(u, v)|$ .
6. Aplicamos la transformada rápida de Fourier a las imágenes en coordenadas log-polares, obteniendo  $Flp_1(u, v)$  y  $Flp_2(u, v)$ .
7. Calculamos la correlación de fase entre  $Flp_1(u, v)$  y  $Flp_2(u, v)$ . Es decir la transformada inversa de Fourier del resultado obtenido por la ecuación (4.33). Denominemos la misma  $IR_1$ .
8. Localizamos el punto  $(\log(s), \theta_0)$ , el cuál debe ser la posición donde se encuentra el valor máximo de  $|IR_1|$ . Conseguimos los valores de la escala y el ángulo de rotación.
9. Construimos una nueva imagen  $I_3$ , aplicando los valores de rotación y de escala obtenidos a  $I_2$ .

10. Aplicamos transformada rápida de Fourier a  $I_3$ , obteniendo  $F_3(u, v)$ .
11. Calculamos la correlación de fase entre  $F_1(u, v)$  y  $F_3(u, v)$ (4.33). Representamos la misma con  $IR_2$ .
12. Obtenemos el vector de traslación  $(x_0, y_0)$ , buscando la posición del valor máximo de  $|IR_2|$
13. Aplicamos traslación a la imagen  $I_3$ , y construimos imagen final resultante del registro.

---

## Capítulo 5

# Recolección de datos y Evaluación del Algoritmo

Luego de implementarse el algoritmo de la transformada de Fourier para el registro automático de imágenes, se evaluó el rendimiento del mismo. Para llevar a cabo el mismo se decidió adoptar la siguiente metodología: Se toma una imagen y la misma se le realiza algún tipo de transformación, ya sea traslación, rotación o modificación de escala. Luego se aplica el algoritmo implementado para poder verificar si los valores que se obtienen son los mismos que se han utilizado en las transformaciones. [8]

### 5.1. Imágenes utilizadas para la evaluación

Para poder verificar el rendimiento y precisión del algoritmo se decidió trabajar con imágenes satelitales obtenidas por diferentes sensores del satélite Landsat. Las imágenes utilizadas son de la ciudad de Córdoba, Argentina.

El algoritmo fue evaluado con dos imágenes. La primera fue sensada por el sensor TM del satélite Landsat 5. Las imágenes captadas por el mismo, están compuesto por 7 bandas espectrales y tienen una resolución terrestre de  $30m$  de tamaño de celda y banda térmica de  $120m$  de resolución terrestre. Proporciona una resolución global de  $30m$ , lo que permite la elaboración de mapas temáticos a una escala de  $1 : 100,000$ .

La segunda de ellas fue obtenida por el sensor **ETM+** de la serie de satélites Landsat 7. Una imagen Landsat 7 ETM+ está compuesta por 8 bandas espectrales. Las bandas del espectro visible y del infrarrojo mantienen la resolución espacial de  $30m$ , lo que permite generación de imágenes con aplicaciones directas hasta una escala de  $1 : 25,000$ . Además cuenta con resolución espacial de la banda térmica de  $60m$ , y la banda pancromática con una resolución espacial de  $15m$ .

Ambas imágenes, representan la ciudad de Córdoba. A continuación se darán detalles de las mismas:

- Imagen de la ciudad de Córdoba detectada por Landsat TM:

Representa la ciudad de Córdoba, capital de la Provincia de Córdoba. La misma fué detectada el día 17 de enero de 1992(5.1).

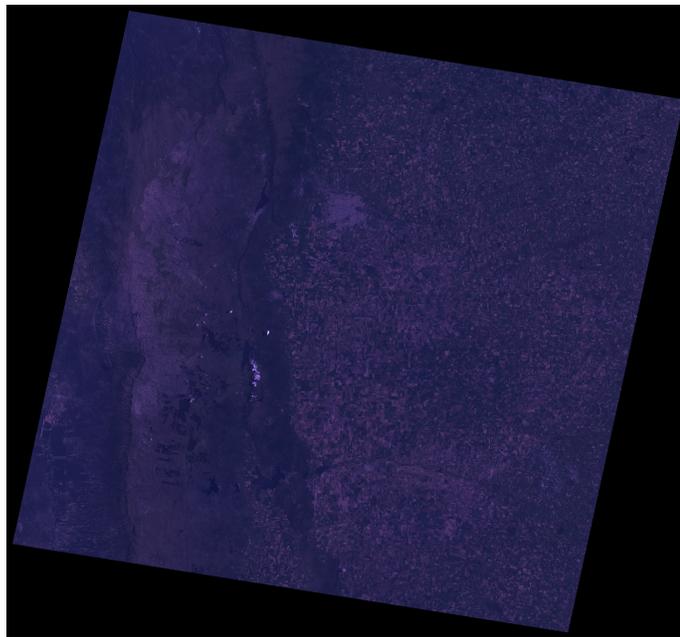


Figura 5.1: Imagen completa de la ciudad de Córdoba (RGB) tomada por Landsat TM

- Imagen de la ciudad de Córdoba detectada por Landsat ETM+: También representa a la ciudad de Córdoba. Fué detectada el día 01 de enero de 2001. (5.2)

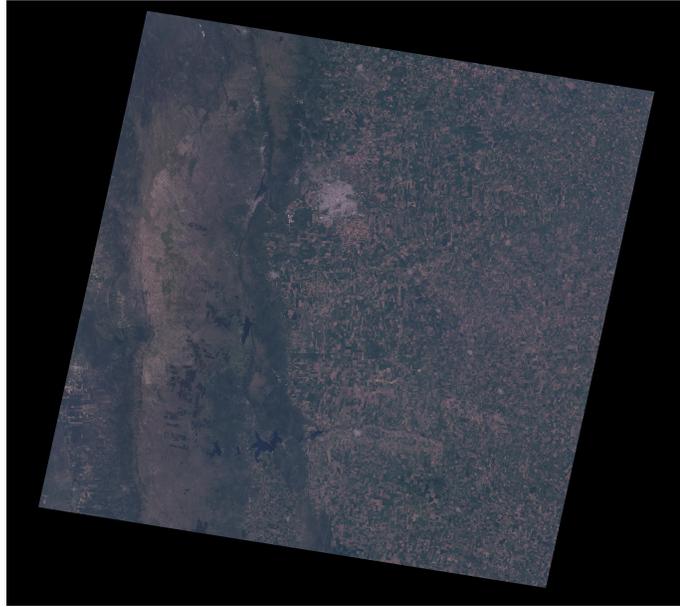


Figura 5.2: Imagen completa de la ciudad de Córdoba (RGB) tomada por Landsat ETM+

Para llevar a cabo la recolección de datos, utilizaremos únicamente, una banda del espectro visible para aplicar el algoritmo, y trabajaremos con recortes de las imágenes nombradas anteriormente.

### 5.1.1. Evaluación del algoritmo de con transformaciones individuales

Se realizaron varios tests, para poder conseguir datos relevantes del algoritmo implementado. En el primero de ellos tomamos un recorte de  $512 \times 512$  de la imagen detectada por Landsat TM y le aplicamos un serie de transformaciones, para poder verificar el algoritmo de la transformada rápida de Fourier. De esta manera, simplificamos el trabajo de verificación de los valores.

Primero, generamos nuevas imágenes a partir de la original, a través de traslaciones aplicadas sobre la misma. El valor  $(\mathbf{x}, \mathbf{y})$ , es el valor original de la traslación,  $(\mathbf{x}_0, \mathbf{y}_0)$  la traslación estimada por el algoritmo, **tpeak** es el máximo valor de la transformada inversa de la correlación de fase y **tiempo** es el tiempo de ejecución. Podemos ver los resultados en la tabla siguiente:

$(x, y)$	$(x_0, y_0)$	<b>tpeak</b>	<b>tiempo</b> (segs)
(50, 50)	(50, 50)	0,84	0,101
(100, 100)	(100, 100)	0,72	0,094
(150, 150)	(150, 150)	0,606	0,099
(200, 200)	(200, 200)	0,49	0,102
(250, 250)	(250, 250)	0,4	0,976
(257, 257)	(255, 255)	0,38	0,104

Cuadro 5.1: Resultados registro de imágenes afectadas por traslaciones

En la tabla se representan algunas de las transformaciones realizadas, ya que eran demasiadas para explicitarlas todas. El último valor de la misma nos indica el primer valor donde el cálculo de traslación no es correcto. Esto sucede cuando el valor máximo es menor a 0,3, considerado como el valor mínimo confiable definido por el coeficiente de correlación. Podemos ver que cuando su valor es menor a 0,3, pierde exactitud el método. Podemos notar también que cuando el porcentaje de coincidencia entre las dos imágenes es menor al 25 por ciento, disminuye la precisión del algoritmo.

Luego, evaluamos como afectaría al algoritmo de la transformada de Fourier, el trabajar con imágenes que difirieran por alguna rotación. El valor  $\theta$ , es el valor original de la rotación,  $\theta_0$  la rotación estimada por el algoritmo, **speak** es el máximo valor de la transformada inversa de la correlación de fase y **tiempo** es el tiempo de ejecución. A continuación la tabla con los resultados:

$\theta$	$\theta_0$	<b>rpeak</b>	<b>tiempo</b> (segs)
10°	10°	0,3831	7,26
20°	20°	0,399	1,87
30°	30°	0,243	4,47
50°	50°	0,3429	3,27
70°	70°	0,3947	1,89
90°	90°	0,7715	0,87

Cuadro 5.2: Resultados registro de imágenes afectadas por rotaciones

Finalmente evaluamos como se comporta el algoritmo frente a los cambios de escala. El valor **s**, es el valor original de la escala, **s<sub>e</sub>** la escala estimada por el algoritmo, **speak** es el máximo valor de la transformada inversa de la correlación de fase y **tiempo** es el tiempo de ejecución.

<b>s</b>	<b>s<sub>e</sub></b>	<b>speak</b>	<b>tiempo(segs)</b>
0,4	0,3992	0,046	5,35
0,6	0,602	0,141	3,05
0,8	0,803	0,128	1,13
1,2	1,1948	0,264	4,15
1,4	1,3916	0,198	3,18
1,6	1,5938	0,297	2,089
1,8	1,7931	0,324	7,98

Cuadro 5.3: Resultados registro de imágenes afectadas por cambio de escala

### 5.1.2. Evaluación con múltiples transformaciones

A diferencia de los tests realizados anteriormente, observaremos el desempeño del algoritmo, al tratar de registrar una dos imágenes que difieren tanto en escala, traslación y rotación. Para ello utilizaremos el mismo método que en los tests anteriores: a partir de una imagen, le aplicamos las transformadas necesarias, para poder después comparar correctamente los resultados del algoritmo. En este caso utilizaremos un recorte de la imagen de la ciudad de Córdoba tomada el año 2001 (5.3). El tamaño de la misma es de  $1024 \times 1024$ . Para registrar la imagen inicial, con la imagen transformada, se utilizó la banda 1 de esta imagen, la cual es equivalente a la banda azul del espectro electromagnético.

Se realizaron varias pruebas, para poder evaluar el rendimiento y precisión del algoritmo. Para ello realizamos diferentes tipos de transformaciones sobre el recorte, para luego aplicar el algoritmo de registro. A continuación mostramos uno de los resultados:

- Aplicamos una rotación de  $45^\circ$ , una traslación de (252, 235) y un cambio de escala de 1,45 a la imagen (5.3) y obtenemos:



Figura 5.3: Recorte de la imagen (RGB) de la ciudad de Córdoba tomada por Landsat ETM+

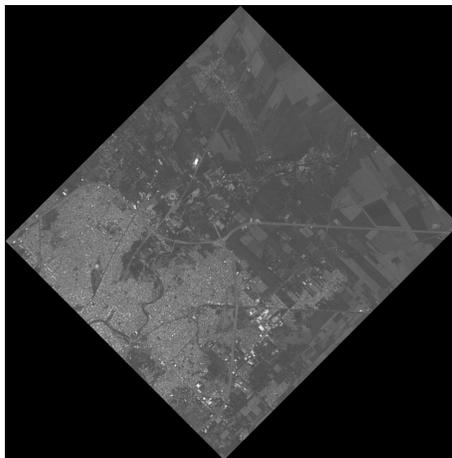


Figura 5.4: Aplicación de transformaciones a 5.3

Luego de aplicar el algoritmo de la transformada rápida de Fourier, obtenemos los siguientes valores de transformación:

- $s = 1,456$
- $(x_0, y_0) = (252, 235)$
- $\theta_0 = 45^\circ$
- Tiempo de ejecución =  $6,03\text{seg}$
- $t_{\text{peak}} = 0,3869$

- $s_{peak}$  y  $r_{peak} = 0,2354$

Luego de aplicar las transformaciones geométricas, estimadas por el algoritmo, a la imagen transformada, procedemos a generar la imagen resultante del registro (5.5).



Figura 5.5: Imagen generada por el registro entre 5.3 y 5.4

### 5.1.3. Obtención de una Imagen a partir de varios recortes

Como última evaluación, se decidió verificar el método para la construcción de imágenes, una vez obtenidos los parámetros necesarios para que se pueda realizar el registro. Para ello utilizamos la interpolación bilineal. Para llevar a cabo esta serie de tests, se realizaron recortes sobre la imagen tomada de la ciudad de Córdoba por el satélite Landsat ETM+. Luego fuimos registrando las imágenes de a pares, construyendo nuevas imágenes, las cuales nuevamente serán registradas hasta poder generar la imagen final. Este proceso se denomina generación de mosaicos a partir de varias imágenes.

A continuación vemos, en imágenes como se realizó el proceso de construcción de imágenes. Utilizamos tres recortes para llevar a cabo este proceso. Las imágenes utilizadas fueron:

Luego aplicamos el algoritmo de registro a los imágenes (5.6a) y (5.6b), obteniendo de esta manera la siguiente imagen:

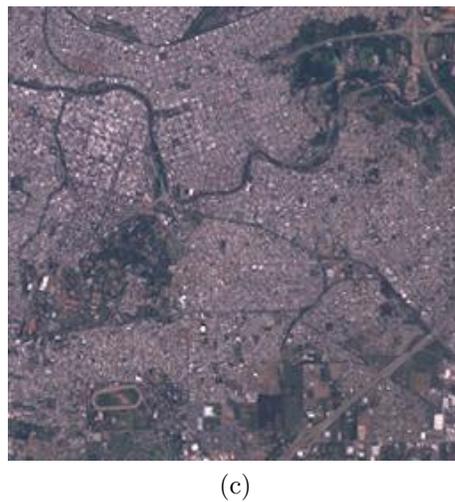
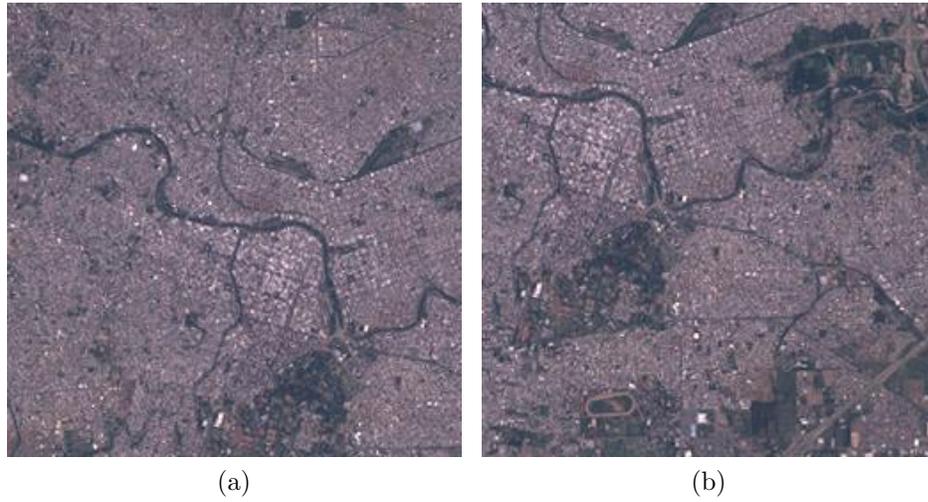


Figura 5.6: Recortes de la imagen (RGB) de la ciudad de Córdoba tomada por Landsat ETM+.



Figura 5.7: Registro entre imágenes (5.6a) y (5.6b)

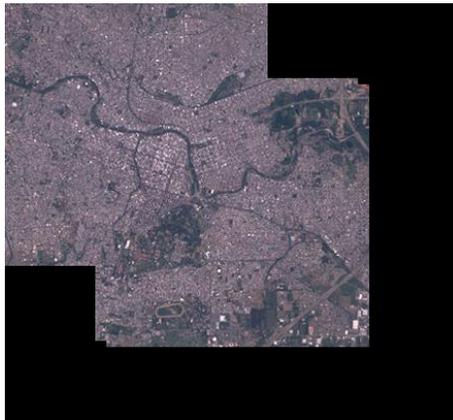


Figura 5.8: Registro entre imágenes (5.7) y (5.6c)

Finalmente aplicamos el algoritmo de registro entre (5.6c) y 5.7, para así obtener el mosaico entre las tres imágenes.

---

## Capítulo 6

# Conclusión y Trabajo Futuro

El registro de imágenes se ha transformado en un área fundamental de la tele-detección espacial, que como describimos anteriormente forma parte de resolución de problemas de vital importancia para el ser humano. Por ende lo necesario de poder definir e implementar métodos automáticos para que lleven a cabo esta tarea. Como dijimos al principio, es muy común que hoy en día, se sigan utilizando métodos manuales para el registro, lo que generará soluciones no muy eficientes y con sus problemáticas.

Tratando de abarcar este problema, de una manera pequeña comparado con lo que es el problema real, hemos podido implementar un algoritmo totalmente automático para el registro de imágenes: el algoritmo basado en la transformada rápida de Fourier.

Como pudimos ver en los resultados obtenidos en el capítulo anterior, el algoritmo cumple muy bien con sus requisitos. Es más, podemos considerar que el mismo resuelve, tres problemas típicos en el mundo real de las imágenes satelitales:

1. Registro de imágenes cercanas
2. Generar mosaicos de imágenes.
3. Registro automático de imágenes satelitales, procedentes del mismo satélite.

En particular, los resultados han sido muy positivos. Se ha logrado poder registrar imágenes con diferencias de rotación de importancia. A lo mejor una de sus limitaciones, es no se pueden registrar imágenes que tengan una coincidencia menor al 25

por ciento, lo cual pasa a ser un problema menor, en el caso en que podamos recortar las imágenes, y de esta manera generamos una coincidencia mayor al 25 por ciento, lo que generará resultados precisos.

La escala permitida para el registro puede ser una limitación grande, para poder generar registro de imágenes entre imágenes tomadas por diferentes satélites. Por lo que se transforma en una herramienta central para registro entre imágenes captadas por un mismo satélite. Aunque puede haber casos en donde se puede registrar imágenes de diferentes satélites: si comparamos imágenes capturadas por el Landsat TM, cuya resolución es de  $30m$  y por el SPOT multiespectral de resolución  $20m$ , funcionará ya que la escala será de  $30/20 = 1,5$ .

En cuanto a la eficiencia, es una muy buena noticia, contar con un algoritmo que tarde menos de de 20 segundos para realizar el registro entre dos imágenes de tamaño  $1024 \times 1024$ . Es una excelente forma de reemplazar el registro manual, utilizado en la gran mayoría de softwares de procesamiento de imágenes.

Como trabajo futuro, proponemos extender el algoritmo de la transformada de Fourier, para poder trabajar con diferencias de escala mayores. Se propone agregar el trabajo con imágenes multi temporales. Además, consideramos de gran idea, utilizar las múltiples bandas para poder generar un registro mucho más preciso.

---

# Bibliografía

- [1] Lisa Gottesfeld Brown. A survey of image registration techniques. 1991.
- [2] Emilio Chuvieco. *Fundamentos de Teledetección Espacial*. 2010.
- [3] Python Software Foundation. <http://www.python.org/>. 2013.
- [4] A. Ardeshir Goshtasby. *Image Registration*. Springer, 2012.
- [5] Gang Hong y Yun Zhang. The image registration technique for high resolution remote sensing image in hilly area.
- [6] Nathan S. Netanyahu Jacqueline Le Moigne y Roger D. Eastman. *Image Registration for Remote Sensing*. Cambridge, 2011.
- [7] Mario A. Lamfri Oscar H. Bustos, Alejandro C. Frery y Carlos M. Scavuzzo. *Técnicas Estadísticas en Teledetección Espacial*. 2004.
- [8] B. Srinivasa Reddy y B. N. Chatterji. An fft-based technique for translation, rotation and scale-invariant image registration. *Transactions on Image Processing*, 5:977–1000, 1996.
- [9] J.A. Richards y X. Jia. *Remote Sensing Digital Image Analysis*. Springer, 1999.
- [10] José A. Sobrino. *Teledetección*. 2001.
- [11] Hongjie Xie y Nigel Hicks. An idl/envi implementation of the fft based algorithm for automatic image registration. *Computers and Geosciences*, 29:1045–1055, 2003.

- [12] Barbara Zitova y Jan Flusser. Image registration methods: a survey. *Image and Vision Computing*, 21:977–1000, 2003.