

UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA
Máster Oficial en Sistemas Inteligentes y Aplicaciones Numéricas en
Ingeniería



Trabajo Final de Master

Análisis Fotogramétrico de Imágenes Adquiridas por un Avión no Tripulado

Moisés Díaz Cabrera

Tutores: Dr. D. Jorge Cabrera Gámez
Dr. D. Ricardo Aguasca Colomo

Octubre 2011

Agradecimientos

*Los que aseguran que es imposible,
no deberían interrumpir
a los que estamos intentándolo.*
Thomas Eddison

Se hace especial mención al Club Deportivo de Radio Control y Modelismo AYA, de Santa Lucía de Gran Canaria, por la estrecha colaboración con este trabajo. Especialmente por la colaboración prestada en la preparación del UAV y montaje de la cámara, así como en el vuelo realizado.

Resumen

Los vehículos aéreos no tripulados (VANT/UAV *Unmanned Aerial Vehicle*) son empleadas en numerosas aplicaciones. A menudo son pilotados por control remoto, sin tripulación humana a bordo. Consiguen ser más útiles cuando en su espacio se alojan dispositivos sensoriales, tales como sistemas inerciales o, en particular, cámaras de vídeo. De este modo, es posible disponer de una vista aérea interesante.

El presente trabajo trata en un sentido amplio el uso de imágenes aéreas. Estas podrían emplearse para evaluar la cobertura vegetal de una zona; podrían actualizar bases de datos cartográficas realizando mosaicos de mayor resolución; podrían emplearse para medir zonas devastadas por un incendio o zonas tóxicas, de difícil acceso, etc. En el contexto planteado se estudia la posibilidad de realizar fotomosaicos, a partir de la superposición de imágenes aéreas de una determinada zona rural y con una apartamentada determinada. Es decir, el trabajo trata una parte del problema amplio del stitching.

La secuencia de vídeo es recogida en tiempo real por un receptor de vídeo conectado a un digitalizador que se conecta al PC. Se ha sintetizado un sistema de captación de imágenes integrando componentes para digitalizar un vídeo. Adicionalmente el vídeo es guardado en un dispositivo almacenador multimedia.

Mediante software se han realizado varias tareas de desarrollo o programación, las cuales se componen en digitalizar el vídeo, la corrección de distorsión de las imágenes, un filtrado para mitigar el efecto del entrelazado y ganar nitidez en las imágenes. Todo ello se realiza fuera del tiempo real en que es captado el vídeo.

En la secuencia se seleccionan imágenes tomadas a diferentes alturas, con diferentes condiciones de luminosidad, con diferentes ángulos de cabeceo y alabeo, etc. Debido al uso en este ámbito, se ha empleado la herramienta SURF para realizar un estudio de carácter exploratorio sobre los descriptores de las imágenes.

Los resultados y conclusiones obtenidas parten de unas condiciones experimentales limitadas y adversas. Con lo cual, se espera que si en estas condiciones los resultados son modestos, en mejores condiciones será posible realizar tareas

de stitching y de actualización cartográfica con mejores resultados.

Palabras Claves: detectores de puntos de interés, descriptores locales, vehículo aéreo no tripulado, trazado de mapas, fotografía aérea.

Abstract

Unmanned Aerial Vehicles (UAVs) have many applications. They are usually operated by remote control. This saves the requirement for a human pilot, saving weight and safety considerations. They house sensory devices such as inertial systems or video cameras in particular. Thus it is possible to have an aerial view, augmented by additional physical information.

This paper broadly discusses the use and processing of aerial images. These kind of images could be used to assess the vegetation of an area, to update cartographic databases by providing higher resolution mosaics or even to measure areas devastated by fire or toxic areas of difficult access. In the context under study raised the possibility of photomosaics, from overlapping aerial images of a particular rural area with a gear on the UAV and devices determined. That is, the work is only a piece of the broader problem of stitching.

The video stream is collected in real time by a video receiver, which is connected to a scanner linked to a personal computer. We have synthesized an imaging system incorporating a video camera to digitize. Additionally, the video is captured and stored on a media storage device.

We have developed software programs digitizing the video correction of image distortion, filtering to reduce the effect of interlacing and to gain clarity in the images. It was not done in real-time.

The video sequence comprises selected images taken from different heights, with different lighting conditions, with different pitch, roll and even yaw angles, etc. . . As is common in this field, we used SURF in order to explore the descriptors of the images.

The results and conclusions are based on limited and adverse experimental conditions. So that, modest results can only be achieved. We hope that better conditions will make it possible to carry out cartographic update stitching and to get better results.

Key words: keypoints detectors, local descriptors, unmanned aerial vehicle (UAV), mapping, aerial photography.

Índice

Índice	v
Índice de figuras	vii
Índice de cuadros	ix
1 Introducción	1
1.1 Introducción	1
2 Trabajos y técnicas relacionadas	7
2.1 Imágenes aéreas para la reconstrucción de mosaicos	7
2.2 Métodos de detección	10
2.2.1 Métodos de detección monoescala	11
2.2.1.1 Detector Harris	11
2.2.1.2 Detector Shi - Tomasi	11
2.2.1.3 Detector de SUSAN	13
2.2.2 Métodos de detección multiescala	13
2.2.2.1 Harris - Laplace	14
2.2.2.2 SIFT (Scale Invariant Features Transforms) . .	15
2.2.2.3 SURF	19
3 Digitalización del vídeo	27
3.1 Sistema electrónico	27
3.1.1 Sistema electrónico a bordo del UAV	28
3.1.2 Sistema electrónico en la estación de tierra	30
3.2 Programa del IC Control con OpenCV	31
4 Preprocesado de las imágenes	33
4.1 Calibración de la cámara	33
4.1.1 Modelo de cámara	33
4.1.2 Método de Zhang	34
4.1.3 Herramientas de calibración	35

ÍNDICE

4.1.4	Resultados	39
4.2	Filtrado	41
5	Estudio de las imágenes con la herramienta SURF	45
5.1	Conjunto de imágenes a estudiar	45
5.2	Eliminación de puntos no discriminantes	48
5.3	Emparejamientos manuales	49
5.3.1	Rango de emparejamiento de puntos manuales	50
5.3.2	Experimento 1	53
5.3.3	Experimento 2	56
5.3.4	Otros resultados	60
5.4	Emparejamientos automáticos	62
5.4.1	Metodología	62
5.4.2	Primer subconjunto de imágenes	63
5.4.2.1	Prueba 1	63
5.4.2.2	Prueba 2	64
5.4.2.3	Prueba 3	64
5.4.3	Segundo subconjunto de imágenes	66
5.4.3.1	Prueba 1	66
5.4.3.2	Prueba 2	68
5.4.3.3	Prueba 3	69
5.4.4	Combinación de los subconjuntos de imágenes	70
5.4.4.1	Prueba 1	70
5.4.4.2	Prueba 2	71
5.4.4.3	Prueba 3	73
6	Conclusiones y líneas futuras	77
6.1	Conclusiones	77
6.2	Líneas futuras	79
	Bibliografía	81
	Glosario	85

Índice de figuras

1.1	Escena en cartografía de GRAFCAN y en nuestras imágenes . . .	4
1.2	Cámara instalada en el avión	5
2.1	Gradientes de intensidad.	11
2.2	Magnitudes detectadas por el área USAN.	13
2.3	Imagen en varias escalas (Harris - Laplace)	14
2.4	Pirámide Gaussiana compuesta por 5 escalas y 6 octavas.	15
2.5	Pirámide Difference-of-Gaussian $D(x, y, \sigma)$	16
2.6	Comparación de un Keypoint con sus vecinos	17
2.7	Orientaciones de los vecinos de un keypoint	19
2.8	Filtro de cajas - SURF	20
2.9	Intensidad de la región \sum según las imágenes integrales	21
2.10	Izquierda: Método SIFT. Derecha: Método SURF	22
2.11	Longitud de los filtros en diferentes octavas - SURF	23
2.12	Asignación de la orientación	24
2.13	Respuestas de Haar en las sub-regiones alrededor del keypoint	25
3.1	Esquema del sistema electrónico/hardware del VANT/UAV	28
3.2	Subsistema de captación/emisión de vídeo	29
3.3	Imagen de los dispositivos de abordaje del UAV.	30
3.4	Esquema del software de grabación de vídeo	32
4.1	Formación de la imagen en una cámara pinhole	34
4.2	Plantilla de calibración pintada	37
4.3	Patrones distorsionados con puntos retroproyectados pintados	40
4.4	Patrones con y sin distorsión	42
4.5	Imagen original e imágenes resultantes tras aplicarle el filtrado	43
5.1	Conjunto de imágenes	46
5.2	Subconjunto 1	47
5.3	Subconjunto 2	47
5.4	Filtrado de keypoints singulares (1º ejemplo)	49
5.5	Filtrado de keypoints singulares (2º ejemplo)	50

ÍNDICE DE FIGURAS

5.6	Keypoints manualmente seleccionados	51
5.7	Localización del keypoint nº 8 en una serie de imágenes	52
5.8	Intervalos de distancias euclídeas	56
5.9	Aparición de keypoints según los intervalos	61
5.10	Filtrado de keypoints singulares (1º ejemplo)	65
5.11	Emparejamiento automático con SURF, subconjunto 2	67
5.12	Filtrado de keypoints singulares (2º ejemplo)	68
5.13	Keypoints manualmente seleccionados (II)	69
5.14	Emparejamiento automático con SURF, conjunto total	72

Índice de cuadros

2.1	Construcción de las octavas en el método de SURF	23
4.1	Errores de retroproyección	40
4.2	Matriz de calibración (matriz A en la ecuación 4.1).	40
4.3	Coefficientes de distorsión	41
5.1	Intervalo keypoint 1	57
5.2	Intervalo keypoint 2	57
5.3	Intervalo keypoint 3	58
5.4	Intervalo keypoint 4	58
5.5	Intervalo keypoint 5	58
5.6	Intervalo keypoint 6	59
5.7	Intervalo keypoint 7	59
5.8	Intervalo keypoint 8	59
5.9	Intervalo keypoint 9	60
5.10	Resumen de intervalos de confianza en cada punto	62
5.11	Emparejamientos automáticos con SURF (I)	64
5.12	Emparejamientos selección de puntos con SURF (I)	64
5.13	Emparejamientos con keypoints manuales con SURF (I)	66
5.14	Emparejamientos automáticos con SURF (II)	66
5.15	Emparejamientos selección de puntos con SURF (II)	68
5.16	Emparejamientos con keypoints manuales con SURF (II)	70
5.17	Emparejamientos automáticos con SURF (III)	70
5.18	Emparejamientos selección de puntos con SURF (III)	71
5.19	Emparejamientos con keypoints manuales con SURF (III)	73
5.20	Resumen de los emparejamientos en 3 experimentos	74

ÍNDICE DE CUADROS

Capítulo 1

Introducción

Resumen: Antes de abordar el desarrollo de este Trabajo Final de Máster, se pretende plasmar una idea general del problema. A continuación se expone el tema que se ha tratado en el trabajo. Adicionalmente se añaden retos planteados, no alcanzados debidos a la condiciones adversas de experimentación, es decir, debido a la mala calidad de las imágenes y a la falta de datos como GPS o datos inerciales.

1.1 Introducción

Las plataformas aéreas no tripuladas (UAV, Unmanned Aerial Vehicle) son empleadas en numerosas aplicaciones. Los vehículos aéreos no tripulados son a menudo pilotados por control remoto, son aeronaves que vuelan sin tripulación humana a bordo. Se usan mayoritariamente en aplicaciones militares.

Actualmente, los UAV militares realizan tanto misiones de reconocimiento como de ataque. A pesar de sus ataques exitosos, en ocasiones causan daños colaterales identificando falsos objetivos. Los UAV también son utilizados en aplicaciones civiles, tales como labores de lucha contra incendios o seguridad civil o como la vigilancia de los oleoductos.

Además, estos vehículos pueden ser aplicados en ambientes de alta toxicidad química y radiológicos, por ejemplo en desastres como el de la central nuclear de Fukushima, en los que sea necesario tomar muestras con alto peligro de vidas humanas y realizar tareas de control de ambiente. Incluso, pueden ser útiles para la cooperación en misiones de control de narcotráfico y contra el terrorismo.

1. Introducción

Estos vehículos consiguen ser de mayor interés cuando en su espacio se alojan dispositivos sensoriales. Es por ello que estos vehículos han de poseer un tamaño moderado para la carga útil que han de soportar. Las baterías para ofrecer una autonomía aceptable, dispositivos, sensores y, sobretodo, cámaras son los elementos principales. Ello hace que con la plataforma se puedan grabar vídeos de alta calidad y así disponer de una visión aérea amplia. En tal caso, el UAV podría llegar a ser empleado como medios de prueba en juicios, mediciones aéreas de espacios urbanos o rurales, composición de mosaicos o mapas, etc.

Los UAV son una alternativa mucho más económica y ecológica que los aviones y helicópteros pilotados. El no llevar tripulación les permite llevar a cabo misiones peligrosas sin riesgos humanos. Ciertamente, se corre el peligro de poder perderle, de ahí surge la necesidad de plataformas aéreas de bajo coste, que aumenten la rentabilidad y el atractivo de estos vehículos. También el reducido peso y tamaño de algunos UAV permiten sobrevolar personas o propiedades sin ponerlos en riesgo ni crear molestias, y pueden incluso pasar desapercibidos, lo que los hace ideales para tareas policiales o de observación de la naturaleza. Cada día se descubren nuevas aplicaciones para los UAV, se trata de una nueva tecnología que está empezando a encontrar aplicaciones prácticas.

En la fabricación de un UAV existen problemas de diseño mecánico y de diseño electrónico que hay que resolver. Se necesitan realizar sistemas en miniatura que sacrifiquen lo menos posible la calidad de los mismos. Así como, la calidad de la información percibida por ellos. La información más rica en abundancia que se puede recoger es la obtenida mediante cámaras bien individuales, estereoscópicas o cámaras infrarrojas. Por ello, la visión por computador tiene un papel fundamental en la explotación de estas plataformas.

Se desea poder realizar una superposición de las imágenes transmitidas desde el emisor de abordaje del vehículo. Mediante la visión por computador, se pretende que a medida que se reciba el flujo de imágenes en la estación de base, se logre georreferenciar y componer un mosaico de la zona sobrevolada.

Actualmente existen empresas dedicadas a este problema. Para la zona donde nos movemos existe GRAFCAN, la cual está especializada en la producción, análisis y difusión de información geográfica y territorial mediante Sistemas de Información Geográfica (SIG) e Infraestructuras de Datos Espaciales (IDE). Ellos disponen de cartografía de calidad de las islas Canarias, especialmente en zonas urbanas, no ocurriendo así en zonas rurales.

Con este trabajo final de máster se pretende realizar un estudio exploratorio de la viabilidad de construir una herramienta que sea capaz de georreferenciar imágenes en la cartografía actual de referencia, es decir, a los mapas de GRAFCAN. Especialmente en zonas para las que se carece de información de calidad,

Análisis Fotogramétrico de Imágenes Adquiridas por un UAV

como son zonas de barrancos, laderas, etc. Es decir, zonas no urbanas pero de interés.

Actualmente se dispone de un vehículo aéreo no tripulado basado en un planeador o “glider”. En él se ha instalado una cámara en posición cenital que graba un vídeo de un vuelo en cuestión. A partir de este vídeo comienza el desarrollo de este trabajo final de máster.

En un inicio ha sido necesario sintetizar un sistema hardware de captación de la secuencia de vídeo. Dicho sistema es capaz de transmitir una secuencia de vídeo, la cual es recibida y digitalizada. Debido a ello ha sido necesario diseñar un software para digitalizar el vídeo en un PC convencional, a partir de las librerías propias del digitalizador impuesto. Una vez terminada la configuración electrónica e informática, se procedió a realizar un vídeo para calibrar la cámara de abordo del vehículo aéreo. Tras este vídeo se realizó la calibración obteniendo los parámetros de calibración que se comentarán en la sección 4.1. Posteriormente se realizó el vuelo y se grabaron las secuencias de vídeo, las cuales son los datos con los que se ha trabajado. El vídeo se ha realizado en una zona de barranco, en concreto en el Barranco de Las Carboneras¹, T.M. Santa Lucia de Tirajana.

A partir de este vídeo se han seleccionado manualmente conjuntos de imágenes que capten una escena. Se ha tenido en cuenta que la escena esté a diferentes alturas, con diferentes coordenadas (X, Y, Z), incluso con diferentes orientaciones en los tres ángulos espaciales, debido al movimiento natural del vehículo aéreo.

En la figura 1.1 se representa una imagen extraída de la cartografía de GRAFCAN y una imagen tomada desde la cámara de abordo del Glider. Ambas imágenes muestran la misma escena, se trata de una parte del barranco. Se visualiza que la imagen de GRAFCAN en esa escena tiene una información pobre en comparación a la imagen tomada con nuestro UAV.

En primer lugar se ha realizado un preprocesamiento de las imágenes, las cuales han sido corregidas de distorsión (efecto barril) con la calibración previa. Adicionalmente se ha realizado un filtrado para desentrelazar las imágenes, debido a que la cámara es entrelazada. En consecuencia, se han diseñado dos programas, uno para la distorsión de las imágenes individuales y otro para el filtrado, como se explica en el capítulo 4.

Con estos datos se ha realizado un estudio de localización con marcas visuales en las imágenes. A partir del detector SURF (Speeded Up Robust Features) (Bay *et al.*, 2008), se han realizado varios estudios en aras a determinar si es viable o no poder trabajar con la cartografía de GRAFCAN y las

¹En Google Maps podemos ver la cartografía de la zona, la cual ha sido suministrada por GRAFCAN. <http://maps.google.es/?ie=UTF8&ll=27.860377,-15.490583&spn=0.00305,0.005681&t=h&z=18>

1. Introducción

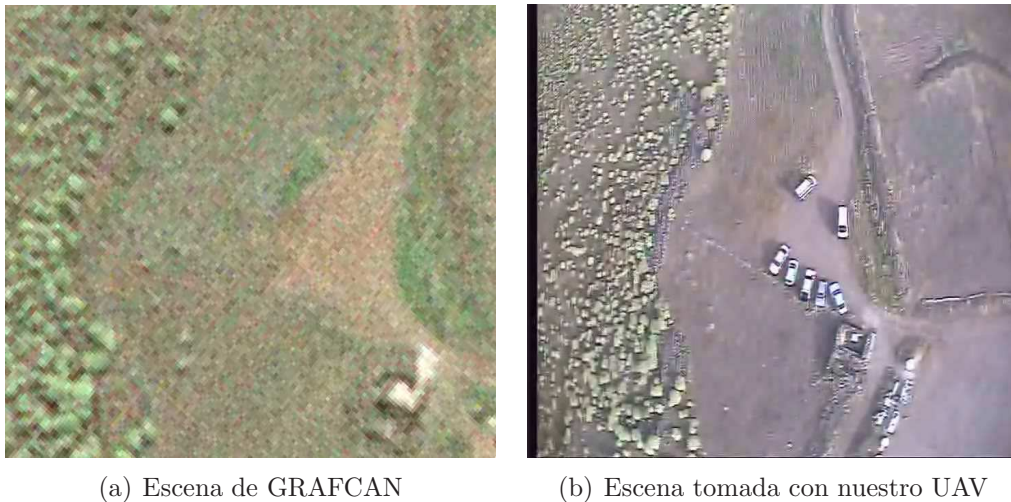


Figura 1.1: Escena tomada desde la cartografía de GRAFCAN (a) y misma escena tomada desde nuestro vehículo aéreo no tripulado (b)

imágenes tomadas.

Se ha pretendido responder a varias cuestiones preliminares que se han planteado a lo largo del trabajo. Dentro de cada una de las imágenes se ha realizado un estudio de lo singular que es o no un punto dentro de una imagen. Es decir, se pretende eliminar puntos que el detector SURF propone automáticamente que realmente no son identificables. Por ejemplo, hay imágenes que están llenas de puntos que no son representativos, como es el caso de tabaibas y/u otros arbustos. El detector SURF detecta como puntos de referencia muchos de estos arbustos.

Finalmente, se ha estudiado la determinación de un rango de error del detector SURF sobre nuestras imágenes. Para ello, se han seleccionado manualmente puntos característicos en la secuencia y se ha calculado la distancia euclídea de los descriptores de esos puntos:

$$\text{Distancia euclídea}_j = \sum_{i=0}^{n_{img}} \left(\text{Descriptor}_{img-ref_i} - \text{Descriptor}_{img-j_i} \right)^2$$

En condiciones ideales, la distancia euclídea de los descriptores de un mismo punto obtenido en dos imágenes diferentes, debe de ser nula, pero en nuestro caso es necesario determinar un rango de error.

Para finalizar el capítulo, el paso final, el cual no se aborda en este trabajo final de máster es realizar fotomosaicos a partir del stitching de un conjunto de n imágenes individuales I_i . Existen diversos métodos de realizar el mosaico con las imágenes, como se apunta en la sección 2.1. Sin embargo, mu-



Figura 1.2: Mini Cámara Sony Color, instalada en posición cenital en el avión.

chos autores proponen que la propuesta híbrida, la cual consiste combinar el tratamiento por software de las imágenes con datos de un terreno conocido, es la que mejores resultados presenta (Yahyanejad *et al.*, 2010; Pesti *et al.*, 2008; Prado-Molina *et al.*, 2009).

A continuación se resumirán los retos más importantes para resolver este problema, muchos de los cuales coinciden con (Yahyanejad *et al.*, 2010).

Baja altitud y superficie no plana

Cuando se toman imágenes a baja altitud, asumir que la superficie es plana supone asumir un riesgo en la realización del mosaico. Objetos como árboles, rugosidades o desniveles debidos a la orografía del terreno o incluso coches y edificios, pueden causar distorsión en la perspectiva de las imágenes. Sin un plano de referencia del terreno a estudiar, el solapamiento de las imágenes necesitaría de una profunda información, para la cual existen sensores o, por ejemplo, cámaras infrarrojas (Prado-Molina *et al.*, 2009). Debido a ello, los puntos a diferente nivel en las imágenes pueden causar graves errores en la superposición de las imágenes individuales I_i .

Cámara de bajo coste

Una parte delicada de este trabajo es la cámara empleada en el avión no tripulado. Dicha cámara es de bajo coste y, por tanto, las características técnicas no son sofisticadas. La consecuencia es directa en cuanto a la calidad de las imágenes seleccionadas. Se trata de una mini cámara Sony Color CCD, como vemos en la imagen 1.1. Dicha cámara captura imágenes entrelazadas y de baja resolución.

1. Introducción

Información inexacta de la posición y la orientación

El sensor que indica la posición del avión no tripulado es un GPS, el cual nos proporciona datos de la altitud de vuelo. Para este trabajo no se han empleado estos datos. Por otro lado, es muy deseable conocer los ángulos de cabeceo (pitch), balanceo (roll), guiñada (yaw). Sin embargo, la plataforma aérea dispone de un sensor que nos indica el horizonte artificial. En resumen, no se dispone de datos en todas las direcciones del movimiento de la aeronave pues falta el dato de guiñada. Sin embargo, estos datos en pequeños vehículos aéreos tienen una exactitud limitada, la cual no puede ser comparada con grandes aviones o aircraft. Debido a ello, no puede ser enorme la dependencia de estos datos, por tanto, es necesario hacer frente a esta inexactitud en todo el proceso cuyo final es el mosaico de las imágenes.

Limitación de recursos

Mientras el vehículo aéreo no tripulado está en vuelo, el vídeo es grabado en un dispositivo de almacenamiento, alojado en dicho vehículo. El vídeo es transmitido en analógico y grabado en formato digital en la estación de base. Dicha estación consistirá en ordenadores portátiles. Por otro lado, la responsabilidad de la información de referencia del vehículo durante el vuelo recae en el GPS a bordo. El vuelo realizado para este trabajo no contaba con el desarrollo de la telemetría, debido a ello no ha sido posible usar la información que se debe enviar desde el UAV hacia la estación de base.

Capítulo 2

Trabajos y técnicas relacionadas

Resumen: Este capítulo presenta una revisión bibliográfica de trabajos relacionados con el presente trabajo. Una parte importante se dedica a explicar algunos métodos de detección de puntos “singulares” o de interés en imágenes, en concreto se explica con más detenimiento los detectores y descriptores de características SIFT y SURF.

2.1 Imágenes aéreas para la reconstrucción de mosaicos

La fotogrametría es una técnica que utiliza imágenes aéreas para reconstruir mapas aéreos sumando imágenes más pequeñas, es decir, es el arreglo de fotografías aéreas que cubren una zona con el fin de conseguir un mapa continuo de un área determinada.

En (Prado-Molina *et al.*, 2009) se hace un estudio acerca de los métodos que se deben emplear para reconstruir mapas aéreos. Estos métodos son: generación de mosaicos no-controlados, semi-controlados, controlados u ortorectificados. Además, para cualquiera de esos métodos, se apuesta por cámaras de bajo costo montadas en aeronaves aerofotográficas¹. Ellos proponen formar mosaicos aéreos con dos cámaras: una cámara para adquirir imágenes a color RGB y otra cámara de infrarrojos. La razón de la cámara de infrarrojos es debida a que de esta manera es posible llevar a cabo una cuantificación de

¹Vehículo aéreo con una parte externa dedicada al montaje giroestabilizado de la cámara.

2. Trabajos y técnicas relacionadas

áreas afectadas, pues con ella se aprecia fácilmente la vegetación y su estado de actividad fotosintética². En cuanto a la generación de mosaicos ortorectificados, en (López, 2006) se proponen unos pasos a seguir para generar ortomágenes, estos consisten en disponer de elementos externos (GPS y ángulos de orientación de cada toma), la corrección de la distorsión previa calibración de la cámara, por lo menos tres puntos de control en el terreno y un modelo digital de elevación.

Por otro lado, en (Pesti *et al.*, 2008) se desarrolla un software llamado MapSticher, el cual es un sistema que a partir de imágenes obtenidas con capturas de bajo coste, es capaz de ortorectificar la imagen y crear mapas aéreos. MapSticher permite de manera manual-automática crear los mapas: se introducen manualmente las imágenes y algunos puntos característicos del terreno y mediante el stitching se van formando el mapa aéreo. Se empleó esa técnica para crear una porción de mapa del valle de Skagit River en las cercanías de la ciudad de Concrete, Washington (población 790). Además se hace un análisis de la resolución que debe emplearse según la extensión del terreno. Por ejemplo, si se desea hacer el ortomosaico de un terreno, los autores apuestan porque la resolución que se debe emplear sea de 0.15 m/píxel. Por otro lado, la técnica que se aplica es más precisa que la técnica clásica del stitching seguida del “rubber-sheeting³”. Así se consigue menos esfuerzo y se produce una mejor composición de las imágenes.

En (Ladd *et al.*, 2006) se presenta un proyecto dónde se explica con detalle la adquisición de imágenes, rectificación, georreferenciación y mosaicos de imágenes para generar una imagen compuesta de todo el campo. Los datos recogidos por un sistema de monitoreo se comparan con la información obtenida de las imágenes. Dicho de otro modo, este documento se centra en la utilización de técnicas fotogramétricas para la minería de datos de las imágenes capturadas por pequeñas plataformas aéreas con control remoto. El documento explica las diferentes técnicas de fotogrametría, las cuales se explican en el glosario, y narran cómo fueron utilizadas en imágenes tomadas desde diferentes plataformas aéreas y con diferentes cámaras. El documento concluye con que el programa desarrollado para la construcción de mapas aéreos ha cumplido las expectativas. Por otro lado, se deja abierto la mejora del código en Matlab con métodos de normalización para el color de las imágenes. Además de las mejoras en las cámaras empleadas y en la plataforma aérea. Un interesante punto que propone el documento (Ladd *et al.*, 2006) de cara al presente trabajo, es que existen varias técnicas para la construcción de mapas

²La actividad fotosintética se refiere a plantas u organismos capaces de realizar la fotosíntesis.

³Deformación de las imágenes que han sido unidas por coordenadas globales con la técnica del stitching (ver glosario).

Análisis Fotogramétrico de Imágenes Adquiridas por un UAV

aéreos. Sin embargo, el método más adecuado es el de georreferenciar cada imagen por separado y luego realizar el mosaico. De esta manera se reducen los errores, permitiendo aumentar la resolución, precisión y exactitud en todo el mosaico⁴.

Por otro lado, las aplicaciones de las imágenes tomadas desde un vehículo aéreo no tripulado son de gran utilidad en un buen número de aplicaciones. En (Quaritsch *et al.*, 2008) se apunta que las vistas aéreas de ojo de pájaro que proporcionan las cámaras instaladas en estos vehículos, tienen tres dominios de aplicación: supervisión del medio que captan las imágenes, vigilancia y posterior aplicación de la ley, y la gestión de desastres en el medio. En nuestro caso, la aplicación queda como una segunda capa de trabajo. Es decir, el cometido de este trabajo es construir un mapa georreferenciado con la unión de muchas imágenes tomadas desde un vehículo aéreo.

En (Yahyanejad *et al.*, 2010) se apuesta por un método híbrido para lograr el mosaico general de una región de interés. En este trabajo, los autores capturan imágenes a partir de un conjunto de UAVs (vehículos aéreos no tripulados) que vuelan a baja altitud. Tras una comparación con tres métodos de solapamiento de imágenes aéreas, se propone un método, llamado método híbrido, que combina el tratamiento por software de las imágenes con un método a partir de los datos de un terreno conocido. Entre los logros conseguidos, los autores destacan que los errores cometidos entre los datos del GPS y de la distancia estimada son menores de 30 cm. Por otro lado, el tiempo de computación para un conjunto de 37 imágenes se reduce al 70 %, en comparación con uno de los tres métodos propuestos en el artículo.

En el seno del equipo AeroDreams, en el trabajo (Consentino, Julio 2008), se ha realizado un trabajo de reconstrucción de mapas aéreos, cuya extensión supera las 250000 hectáreas, en la parte norte de Argentina. La compañía desarrolló un sistema completo que incluye una plataforma en el avión para estabilizar la cámara, un PC integrado y un software de control. Se empleó una cámara reflex, la cual captura los datos de imágenes para su posterior procesamiento. Con el fin de gestionar un gran volumen de imágenes aéreas para la posterior generación de grandes mosaicos, llevaron a cabo un exitoso método automático. Desde el PC de abordaje en el vehículo aéreo, se envían los datos a la estación de control de tierra vía UHF o a través de datos de satélite. El sistema de abordaje consiste en un PC del tipo SBC, 1 GHz, 256 MB RAM, 8 GB de disco flash con un sistema operativo Linux. Las principales entradas de datos al PC son el GPS + IMU⁵ + datos magnéticos. Además hay otros datos interesantes como el nivel de combustible, las revoluciones por minuto

⁴Ver en el glosario los siguientes términos: rectificación, georeferenciar, imagen ortográfica y mosaico

⁵IMU: Unidad de medida inercial

2. Trabajos y técnicas relacionadas

(rpm) del motor, temperatura de los gases de escape, así como información atmosférica (temperatura del aire, humedad y presión), etc. En cada unidad de tiempo, el software captura y guarda los datos del GPS/IMU en un fichero con la imagen para futuros procesados. Es decir, en cada imagen tenemos los datos siguientes:

$$Imagen_i = \phi(x, y, z, \alpha, \beta, \gamma, t) \quad (2.1)$$

Estos datos se envían a la estación de control de tierra, donde a tiempo real muestra el rumbo de la aeronave y los datos de la imagen tomada. Lo más interesante del artículo de (Consentino, Julio 2008) son las dos posibilidades propuestas de realizar el ortomosaico:

- La primera es usar un conjunto de puntos de control de tierra, normalmente 3 por imagen, los cuales son tomados de imágenes proporcionadas por satélites o por aeronaves tripuladas por humanos, es decir, imágenes de cartografía de baja o media resolución.
- La segunda es usar la información de la localización con el GPS y la de los ángulos de giro con la IMU, guardada en cada imagen. En este caso, podríamos emplear puntos de control de tierra sólo para comprobar los resultados.

Sin embargo, el objetivo de este trabajo es evitar emplear el uso de los puntos de control de tierra para la reconstrucción de mosaicos. Una interesante conclusión de este trabajo ha sido el logro de acumular errores de menos de 8 metros, con lo cual consiguieron un mosaico de resolución 20 cm/px.

2.2 Métodos de detección

A continuación se presenta una variedad de herramientas para detectar puntos de interés en imágenes. Estos métodos principalmente tratan de buscar puntos esquinas en imágenes. Se dividen en dos tipos: métodos monoescala y métodos multiescala. El método monoescala realiza el tratamiento de detección de puntos de la imagen, en su escala original. El método multiescala realiza variaciones en la escala de la imagen con el fin de estudiar la estabilidad de puntos detectados a diferentes escalas.

Muchos trabajos analizan varios de estos métodos, por ejemplo para tareas de navegación visual mediante SLAM, con el fin de seleccionar el detector más adecuado (Ballesta *et al.*, 2010) (Gil *et al.*, 2010). En el presente trabajo no se analizan varios métodos, sino que se ha decidido emplear el detector SURF para detectar puntos de interés, debido a que muchos autores usan dicho método sobre imágenes aéreas.

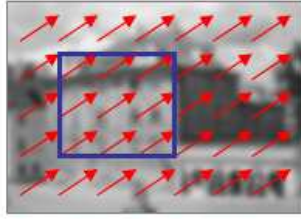


Figura 2.1: Gradientes de intensidad.

2.2.1 Métodos de detección monoescala

A continuación se listarán y se explicarán algunos métodos monoescala de detección de puntos característicos. Estos métodos realizan la detección de puntos sin escalar la imagen, es decir, sin modificar la escala original de la imagen. Los métodos que se explicarán son los más usados por diversos autores.

2.2.1.1 Detector Harris

El detector de esquinas de Harris (Harris and Stephens, 1988) busca las variaciones de intensidad en todas las direcciones: en esquinas o en puntos aislados. Se combina la detección de esquinas y de bordes con una función de autocorrelación local para cada píxel y sus vecinos. Por otro lado, este detector mejoró el detector de esquinas propuesto por Moravec (Moravec, 1980), el cual fue una de las primeras implementaciones conocidas. Este detector, en comparación con el de Moravec, reduce el ruido y, además, la dirección de la región evaluada es independiente de la imagen total. A modo de resumen, el detector de Harris realiza lo que se expone en el algoritmo 1.

Este detector se ha empleado en trabajos de procesamiento de imágenes aéreas. En (Merino Cabañas and Ollero Baturone, 2002) se emplea este detector para monitorizar y detectar. También se usa en técnicas para la estabilización de imágenes y geolocalización de características basadas en algoritmos de seguimiento de regiones y modelos del terreno.

2.2.1.2 Detector Shi - Tomasi

El método de detección de Shi - Tomasi (Shi and Tomasi, 1994) calcula el gradiente de intensidad por cada píxel. Al igual que el detector de Harris, el gradiente proporciona información de la no uniformidad de los niveles de grises en la imagen. Los puntos que se eligen son aquellos puntos que presentan un cambio abrupto en el gradiente vertical y en el horizontal al mismo tiempo.

La elección automática de puntos se determina a partir de la matriz A . En ella se alojan todos los píxeles de la imagen.

2. Trabajos y técnicas relacionadas

Algoritmo 1 Algoritmo de detección de Harris

- 1: Filtrado o emborronamiento Gaussiano de la imagen
- 2: Cálculo de los gradientes de intensidad (figura 2.1) en cada píxel $\nabla I(x, y)$
- 3: **para** $i = 1 \rightarrow n_{px}$ **hacer**
- 4: Cálculo de la matriz de autocorrelación

$$A = \sum \begin{pmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{pmatrix}$$

- 5: Evaluación de la función de respuesta:

$$R(A) = Det(A) - k \cdot Tr(A)^2$$

dónde:

$$R(A) \gg 0 \text{ para } Rank(A) = 2$$

$$R(A) \Rightarrow 0 \text{ para } Rank(A) < 2$$

- 6: **fin para**

- 7: Escoger los mejores candidatos a partir de un umbral (thresholding)
-

$$A = \sum_{u=1}^{u_{max}} \sum_{v=1}^{v_{max}} \begin{pmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{pmatrix}$$

En dónde:

$$I_x = \frac{I(x+1, y) - I(x-1, y)}{2}$$

$$I_y = \frac{I(x, y+1) - I(x, y-1)}{2}$$

El siguiente paso es calcular los autovalores λ_x y λ_y de la matriz A. Un punto será seleccionado si sus autovalores λ_x y λ_y tienen una magnitud elevada. Por ejemplo, si sucede que $\lambda_x \gg \lambda_y$, se interpreta que el píxel en cuestión es interesante en una dirección (en la dirección horizontal) y no en la vertical. Sin embargo, si los autovalores son elevados en ambas direcciones, se interpreta que estamos ante un punto esquina, el cual es interesante detectarlo.

Otros autores, como los del trabajo (Yao *et al.*, 2010), concluyen que en imágenes aéreas el método Shi - Tomasi detecta menos "ground truth" que el método SUSAN a pesar de que el tiempo de procesamiento sea muy parecido. En dicho trabajo realizan experimentos con imágenes aéreas y en el análisis resulta que el método Shi - Tomasi detecta 265 puntos esquinas, frente a los

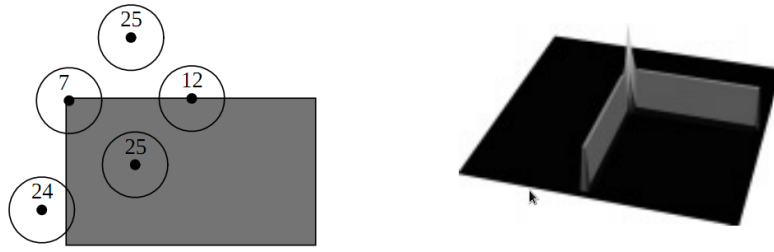


Figura 2.2: Magnitudes detectadas por el área USAN.

598 detectados con SUSAN. Sin embargo, en cuanto a la robustez, los autores prefieren el detector de Shi - Tomasi debido al número de falsas esquinas del detector de SUSAN.

2.2.1.3 Detector de SUSAN

El detector de SUSAN (*Smallest Univalve Segment Assimilating Nucleus*) desarrollado por Brandy y Smith (Smith and Brady, 1997) intenta detectar esquinas, aristas o uniones de manera rápida. Este método crea, para cada punto de la imagen, una pequeña máscara o filtro circular con la misma intensidad que el píxel central (núcleo de la máscara). Esta región se conoce como región USAN. Dentro de la máscara circular se cuenta el número de píxeles con un nivel de gris similar al nivel de gris del núcleo de la máscara. A modo de ejemplo, podemos observar la imagen 2.2. En la imagen de la izquierda podemos ver varias máscaras circulares aplicadas sobre una región rectangular. En la imagen de la derecha, se observa el resultado del detector de SUSAN sobre la imagen. Es un resultado invertido, por tanto el pico, que intenta describir la esquina, se trata del punto que posee menos píxeles similares al píxel central. Es decir, se resuelve un problema de búsqueda de mínimos. Debido a la sencillez, el método es preciso y rápido aunque poco robusto.

En el trabajo (Singh and Rahman, 2008) se aplica el detector de SUSAN en un problema muy concreto: determinar la pista de aterrizaje en condiciones de alta nubosidad. El objetivo es ayudar al piloto a aterrizar cuando las condiciones atmosféricas sean desfavorables. En dicho trabajo se emplean imágenes claras e imágenes cubiertas de nubes con el fin de detectar, de manera robusta, los bordes de la pista de aterrizaje.

2.2.2 Métodos de detección multiescala

Muchos de los detectores descritos hasta ahora extraen las características en una única escala, determinado por parámetros internos de cada detector. Para tratar con cambios de escalas, un enfoque directo consiste en extraer puntos

2. Trabajos y técnicas relacionadas



Figura 2.3: Imagen en varias escalas (Harris - Laplace)

sobre un rango de escalas de la imagen y usarlos todos estos puntos juntos para representar la imagen. Así trabajan los métodos de detección multiescala o la aproximación multiresolución (Crowley and Sanderson, 1987).

2.2.2.1 Harris - Laplace

El método Harris - Laplace (Mikolajczyk and Schmid, 2001) es uno de los primeros métodos invariantes a cambios de escala y a cambios de posición. Surge como una extensión del detector de esquinas de Harris (Harris and Stephens, 1988).

En dicho método se crea un espacio de escalas con la imagen en cuestión, clásicamente Gaussiano, como vemos en la imagen 2.3. Posteriormente, en cada escala se detecta puntos de interés aplicando Harris.

El método tiene un parecido metodológico con el método SIFT, el cual se describe con más profundidad a continuación. Por un lado, Harris - Laplace localiza el extremo máximo sobre la pirámide de escalas aplicando el detector de esquinas de Harris en espacio y la Laplaciana en escala. El método de SIFT busca el extremo máximo sobre la pirámide de escalas aplicando diferencia de Gaussianas en espacio y en escala.

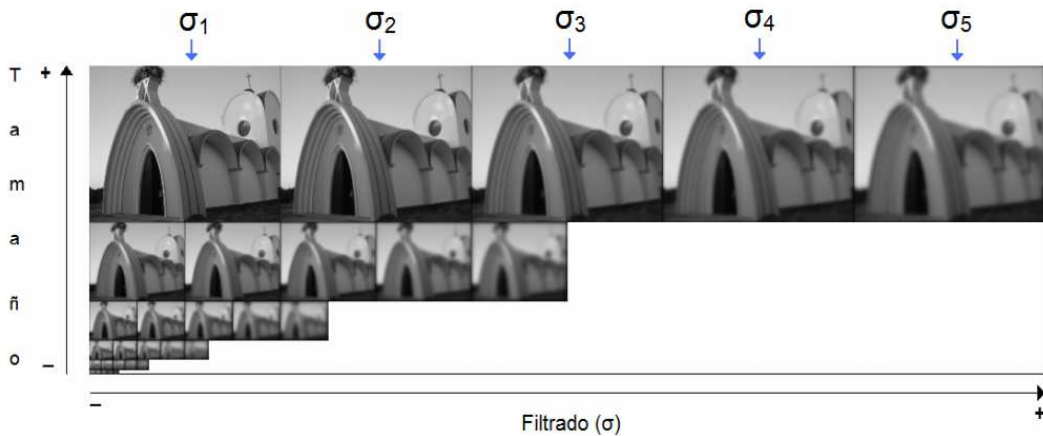


Figura 2.4: Pirámide Gaussiana compuesta por 5 escalas y 6 octavas.

2.2.2.2 SIFT (Scale Invariant Features Transforms)

El algoritmo SIFT (Lowe, 2004) realiza una transformación de la imagen a unas coordenadas que permanecen constantes, principalmente a los cambios de escala, rotación e iluminación. Cada punto invariante de la imagen lleva consigo una serie de valores que describen el punto. Estos valores se denominan descriptores del punto y alojan datos referentes a las características invariantes: posición, escala, orientación y otros valores descriptivos. Típicamente, el tamaño del descriptor es de 128 características. SIFT es un método complejo, con un alto coste computacional. A pesar de ello, algunos autores lo emplean para tareas en tiempo real, debido a una previa y adecuada implementación sobre imágenes de dimensiones reducidas.

El algoritmo se divide en 4 partes bien diferenciadas, éstas son:

1. Detección de puntos extremos en escala - espacio

Se realiza una búsqueda de puntos por todas las localizaciones que forman la pirámide Gaussiana. A modo de ejemplo, ver la imagen 2.4. La pirámide Gaussiana se realiza variando los tamaños de la imagen original y variando el parámetro σ del filtro de Gauss, obteniendo para cada parámetro σ una octava.

Para determinar las localizaciones se realiza la convolución de la Gaussiana G con la imagen $I(x, y)$, obteniendo una función continua conocida como scale-space $L(x, y, \sigma)$:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (2.2)$$

Dónde la Gaussiana sigue la expresión siguiente:

2. Trabajos y técnicas relacionadas

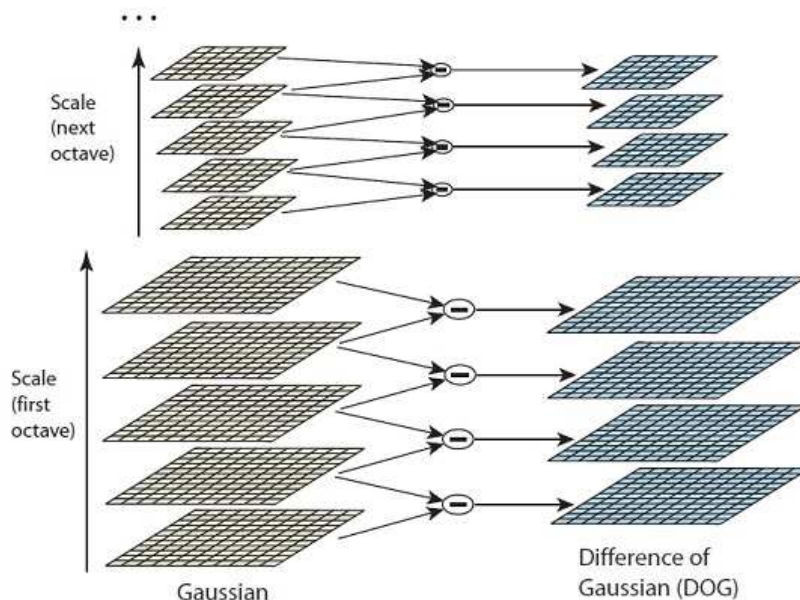


Figura 2.5: Pirámide Difference-of-Gaussian $D(x, y, \sigma)$.

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \cdot e^{-\frac{(x^2 + y^2)}{2\sigma^2}} \quad (2.3)$$

El siguiente paso es calcular la diferencia de Gaussianas (DoG) por escalas, ecuación 2.4. Se trata simplemente de una resta de dos imágenes vecinas dentro de una octava, figura 2.5. Es decir, si observamos la figura 2.4, la diferencia de Gaussianas se calcula con la resta de las imágenes por columnas. De esta manera se pretende detectar rápidamente (la DoG es una simple resta) los keypoints más estables. Nótese que usando la diferencia de Gaussianas pasamos, por cada escala, a tener una imagen menos, debido a la resta. Es decir, en la figura 2.4 se pasaría de 5 a 4 imágenes por escala. A pesar de que no se realiza su demostración, cabe indicar que la diferencia de Gaussianas se justifica debido a su aproximación con la diferencia de Laplacianas, cuyo cálculo computacional es más pesado.

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (2.4)$$

Posteriormente se realiza el cálculo de máximos y mínimos locales de la función $D(x, y, \sigma)$. Como muestra la figura 2.6, cada píxel es comparado

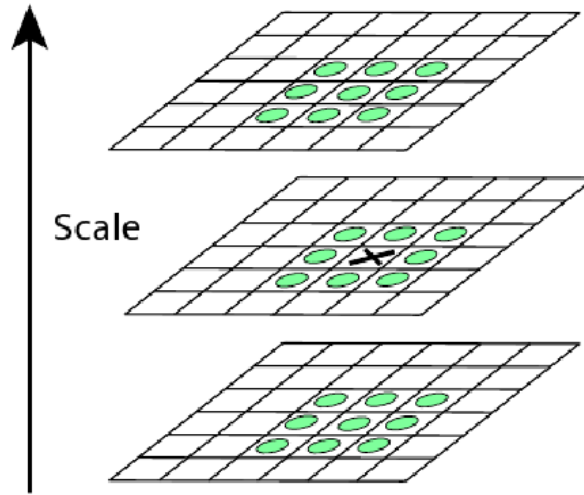


Figura 2.6: La cruz representa el keypoint a estudiar. Los 8 círculos verdes alrededor de la cruz son los vecinos en la escala actual. El resto de círculos rojos son los 9 puntos vecinos en la escala posterior y los 9 en la anterior.

con sus 8 píxeles vecinos y con sus 9 píxeles de la imagen posterior y anterior en la misma escala. Definiendo así un píxel como keypoint cuando éste tenga una magnitud mayor en la función $D(x, y, \sigma)$ que sus 26 píxeles⁶ vecinos.

2. Localización de puntos característicos (Keypoints)

En este punto el algoritmo guarda la información de cada keypoint. La información que se guarda es la escala y la octava a la que pertenece dentro de la pirámide y la posición del píxel en términos de filas y columnas. Para evitar almacenar keypoints sensibles al ruido, aquellos píxeles cuyo contraste sea bajo y/o se localicen en un borde, serán descartados.

La eliminación de puntos debido a un bajo contraste se realiza aplicando un desarrollo en serie de Taylor de la expresión diferencias de Gaussianas $D(x, y, \sigma)$, explicado en la sección 4 del artículo de D. Lowe (Lowe, 2004).

Por otra parte, es de interés suprimir puntos candidatos que se encuentran en el borde, debido a que tendrán una respuesta pobre en la dirección del borde pero una respuesta rica en la dirección perpendicular al borde. Para ello se empleará el Hessiano, al igual que en el detector de esquinas

⁶Ver figura 2.6. La suma de los círculos de la figura resulta 26 (9 + 8 + 9).

2. Trabajos y técnicas relacionadas

de Harris (Harris and Stephens, 1988).

$$H = \begin{vmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{vmatrix} \quad (2.5)$$

De esta matriz podemos deducir la traza y el determinante.

$$Tr(H) = D_{xx} + D_{yy} = \alpha + \beta \quad (2.6)$$

$$Det(H) = D_{xx} \cdot D_{yy} - D_{xy}^2 = \alpha\beta \quad (2.7)$$

Siendo α y β los vectores propios de H . Seguidamente se define, en base al coeficiente $r = \frac{\alpha}{\beta}$ y la relación siguiente:

$$\frac{Tr(H)^2}{Det(H)} < \frac{(r+1)^2}{r} \quad (2.8)$$

Esta relación es eficiente computacionalmente, ya que eliminamos muchos puntos en las fases siguientes. De manera experimental se propone escoger un valor menor que 20 para eliminar los puntos de los bordes.

3. Asignación de la orientación

En este punto se calcula la orientación de cada keypoint seleccionado en el apartado anterior. Para ello se establece una región de 16x16 píxeles alrededor del keypoint. A cada píxel se le calcula su gradiente mediante la diferencia entre píxeles. El gradiente queda definido con la magnitud m y la orientación Θ .

$$m(x, y) = \quad (2.9)$$

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

$$\Theta(x, y) = \tan^{-1} \frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \quad (2.10)$$

Una vez acabado este apartado se poseerá de información de la localización, octava, escala y orientaciones principales de cada región importante de la imagen.

4. Construcción del descriptor de puntos característicos

Para cada keypoint se toman 256 puntos alrededor de él y se calculan las orientaciones, según la ecuación 2.10. Estos 256 puntos surgen de una

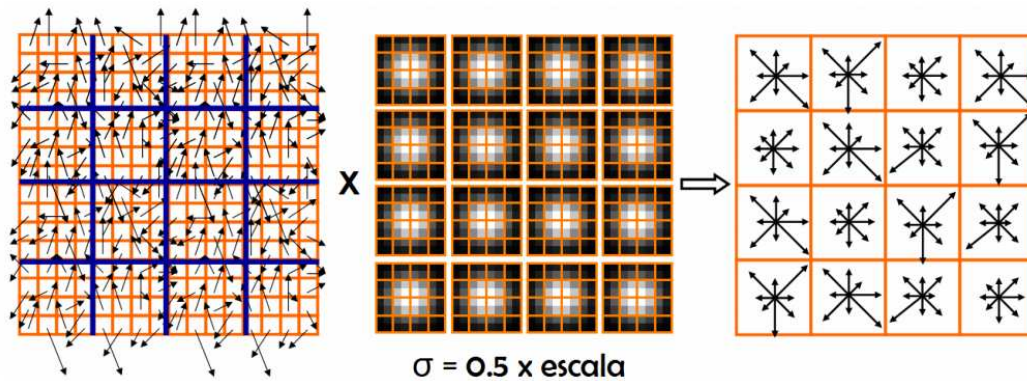


Figura 2.7: Izquierda: Subdivisiones 4x4. Centro: Ventanas circulares gaussianas. Derecha: Descriptor compuesto por 16 histogramas de 8 bins.

gran matriz de 16x16, dónde cada celda se trata de un píxel. Seguidamente se crea una máscara divida en 4x4 partes, la cual se trata de una ventana circular gaussiana con un factor $\sigma = 0.5 \cdot \text{escala}$. El resultado lo vemos en la figura 2.7, dónde obtendremos 16 histogramas de 8 bins cada uno de ellos.

Hay dos parámetros que fijan el tamaño del descriptor, según la expresión 2.11. El tamaño de descriptor más empleado por los autores es de 128 elementos.

$$\text{Tamaño} = r \cdot n^2 = 8 \cdot 4^2 = 128 \quad (2.11)$$

Este último paso pretende robustecer los keypoints ante cambios de luminosidad, contraste y a las no linealidades en la luz, producida bien por la cámara o bien por los cambios de luminosidad de las superficies 3D.

Tras completar estos pasos, el algoritmo logra determinar un número de keypoints con descriptores altamente distintivos. Con lo cual se consigue que una características sea correctamente emparejada con una probabilidad elevada, realzando el adjetivo robusto en este método.

2.2.2.3 SURF

Este método ha sido presentado por (Bay *et al.*, 2008), sus siglas se entienden por *SURF: Speeded Up Robust Features*. Según experimentos realizados por diferentes autores, este método añade mejoras frente al método de SIFT, en cuanto a la robustez, la repetibilidad (constancia de puntos de interés), así como la distinción de los descriptores. El método emplea la matriz Hessiana, igual que SIFT, y añade el concepto de imágenes integrales, las cuales reducen

2. Trabajos y técnicas relacionadas

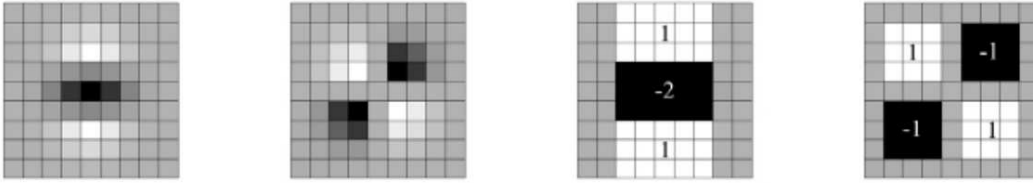


Figura 2.8: De izquierda a derecha: Las derivadas parciales de segundo orden de la Gaussiana en la dirección de y y en la dirección xy y las aproximaciones realizadas usando filtros de caja. Las regiones grises son igual a 0.

considerablemente la carga computacional. Otra mejora es el uso del determinante de la matriz Hessiana para calcular tanto la posición como la escala de los puntos de interés. Se procede a explicar a modo resumen el método.

1. Puntos de interés (keypoints)

Esta parte es muy similar al método de SIFT, aunque SURF introduce unas novedades interesantes desde el punto de vista del tiempo de cálculo. SURF aprovecha en todos los casos el determinante de la matriz Hessiana para la localización y la escala de los puntos. Se justifica este uso en la velocidad de cálculo y en la precisión obtenida. Este punto es interesante debido a que no se usa una métrica para la posición y otra para la escala de los puntos de interés, sino que se emplea la misma medida en todos los casos.

Por ello, supongamos un punto $p = (x, y)$ de la imagen I , la matriz Hessiana $H(p, \sigma)$ del punto $p(x, y)$, que pertenece a la escala σ se define como:

$$H(p, \sigma) = \begin{bmatrix} L_{xx}(p, \sigma) & L_{xy}(p, \sigma) \\ L_{xy}(p, \sigma) & L_{yy}(p, \sigma) \end{bmatrix} \quad (2.12)$$

Donde $L_{xx}(p, \sigma)$ representa la convolución de la derivada parcial de segundo orden de la Gaussiana $\frac{\partial^2}{\partial x^2}g(\sigma)$ con la imagen I en el punto p . Idem para los términos $L_{xy}(p, \sigma)$ y $L_{yy}(p, \sigma)$.

Con el objetivo de conseguir eficiencia, se suprimen las segundas derivadas por una aproximación de filtros en forma de caja (Simard *et al.*, 1999), tal como se explica en la figura 2.8. Esto hace que las derivadas parciales de segundo orden L_{ij} se denoten con D_{ij} .

A continuación se realizan las convoluciones con estos filtros a partir de imágenes integrales, debido a la rapidez que ello conlleva. El cálculo de una imagen integral se lleva a cabo mediante esta fórmula:

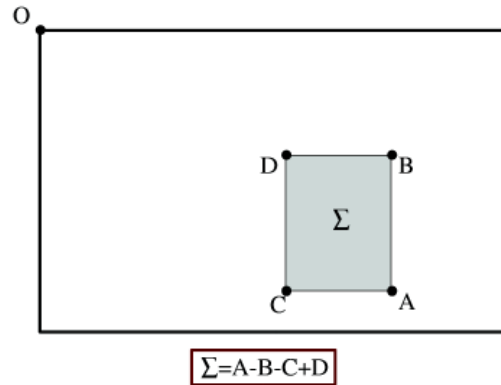


Figura 2.9: Intensidad de la región Σ según las imágenes integrales

$$I_{\Sigma}(x, y) = \sum_{i=1}^{i \leq x} \sum_{j=1}^{j \leq y} I(i, j) \quad (2.13)$$

Dónde (x, y) es el punto en cuestión de la imagen y I es la intensidad de la imagen en el punto. Con ello ya podemos calcular la suma de las intensidades de una región determinada. Supóngase la región Σ de la figura 2.9. El paso siguiente es calcular las imágenes integrales de los puntos A, B, C y D , para finalmente calcular la suma de las imágenes según la ecuación 2.14.

$$\Sigma = A - B - C + D = I_{\Sigma_A}(x, y) - I_{\Sigma_B}(x, y) - I_{\Sigma_C}(x, y) + I_{\Sigma_D}(x, y) \quad (2.14)$$

Como se explicaba en el detector SIFT, la imagen original se suavizaba paulatinamente mediante filtros gaussianos dentro de una escala. Este proceso se iba repitiendo a diferentes escalas. En el caso del detector SURF, al inicio se transforma la imagen original en imagen integral. Aquí no es necesario aplicar un filtro gaussiano a la salida de una imagen previamente filtrada, sino que es posible aplicar los filtros tipo caja, de cualquier tamaño, sobre la imagen original. Por ello, este método no irá disminuyendo de tamaño la imagen como el detector SIFT, sino que irá elevando el tamaño del filtro de cajas. Esta explicación se ilustra en la figura 2.10.

En consecuencia, el determinante del Hessiano se calcula aproximando

2. Trabajos y técnicas relacionadas

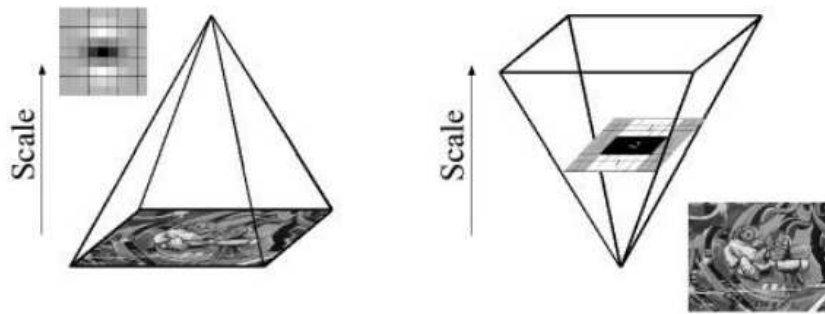


Figura 2.10: Izquierda: Método SIFT. Derecha: Método SURF

las derivadas parciales L_{ij} a D_{ij} :

$$\det(H_{approx}) = D_{xx}D_{yy} - (0.9 \cdot D_{xy})^2 \quad (2.15)$$

El valor 0.9 surge de una relación que parte de la conservación de la energía entre el núcleo gaussiano y una aproximación del núcleo de Gauss. En dicha relación se utiliza la norma de Frobenius. En la práctica se usa 0.9 como valor constante. La aproximación de las derivadas parciales tiene una representación visual que se ilustra en la figura 2.8.

Por otro lado, en el espacio de escala, al igual que sucedía en SIFT, las imágenes se suavizan por octavas, variando la σ de Gauss. Sin embargo, en SURF las octavas se componen de un número fijo de imágenes que resultan de la convolución de la imagen original con una serie de filtros de cajas de diferentes tamaños o resoluciones. La construcción del espacio de escala empieza con un filtro de 9×9 , va aumentando el tamaño según el paso o incremento. El paso de los filtros dentro de una octava es el doble respecto del paso de la octava anterior: en la primera octava el tamaño del filtro es de 6; en la segunda octava el paso (o incremento) del filtro es de 12; en el tercero es de 24; etc. Además el primer filtro de cada octava es el segundo filtro de la octava predecesora. En la tabla 2.1 se pretende mostrar la construcción de las octavas y en la figura 2.11 muestra una vista de conjunto de los tamaños de los filtros para las primeras tres octavas.

Para localizar un punto de interés en todas las escalas, paulatinamente se van eliminando los puntos que no sean máximo con respecto a sus vecinos en un espacio de $3 \times 3 \times 3$. Por ello, el máximo determinante de la matriz Hessiana se interpola en la escala y en la posición de la imagen. En ese punto se concluye la etapa de detección.

Análisis Fotogramétrico de Imágenes Adquiridas por un UAV

1° octava	9 x 9	$\xrightarrow{6}$	15 x 15	$\xrightarrow{6}$	21 x 21	$\xrightarrow{6}$	27 x 27
2° octava	15 x 15	$\xrightarrow{12}$	27 x 27	$\xrightarrow{12}$	39 x 39	$\xrightarrow{12}$	51 x 51
3° octava	27 x 27	$\xrightarrow{24}$	51 x 51	$\xrightarrow{24}$	75 x 75	$\xrightarrow{24}$	99 x 99
⋮							

Cuadro 2.1: Construcción de las octavas en el método de SURF

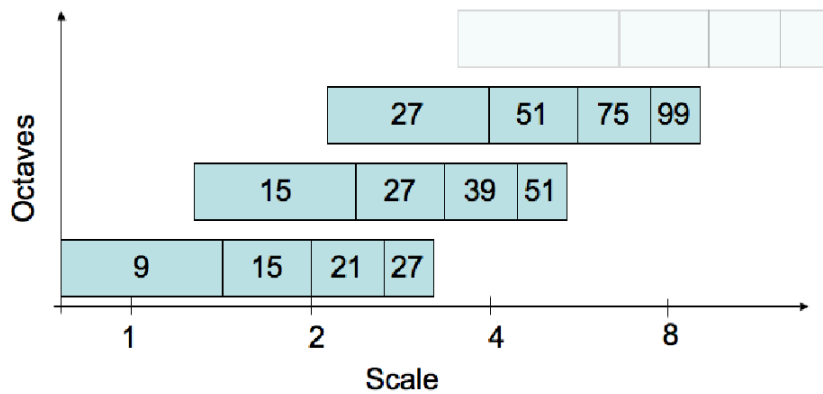


Figura 2.11: Representación gráfica de la longitud de los filtros para tres octavas diferentes. El eje logarítmico horizontal representa las escalas. Las octavas se superponen para cubrir todas las posibles escalas.

2. Trabajos y técnicas relacionadas

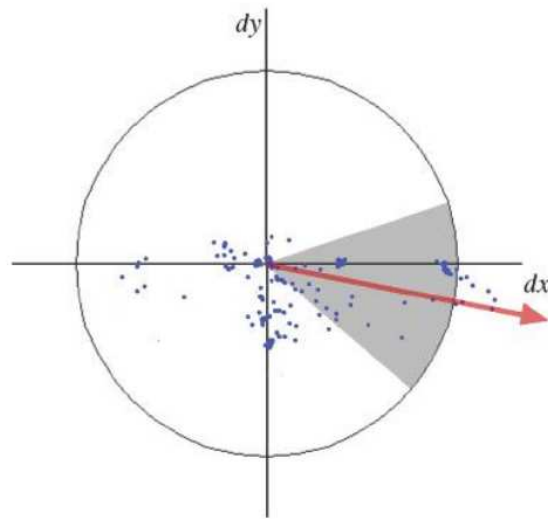


Figura 2.12: Vector de orientación representado en rojo. Los puntos azules son las respuestas obtenidas a raíz de las funciones de Haar. En gris el sector circular de $\frac{\pi}{3}$

2. Asignación de la orientación

El siguiente paso es asignar la orientación a cada uno de los keypoints (*puntos de interés*) encontrados posteriormente. Este paso es importante debido a que se crea invarianza a la orientación a los descriptores de cada punto de interés encontrado. Primero se calcula la respuesta del wavelet de Haar: se trata de una cierta secuencia de funciones. Para el cálculo se toma un área circular centrada en el keypoint y cuyo radio es $6s$, donde s es la escala en la que el keypoint en cuestión fue encontrado. Con lo cual, las funciones de Haar son funciones onduladas, proporcionales a la escala encontrada. Las respuestas de las funciones de Haar son representadas como vectores en el espacio, donde la respuesta en x es la proyección del vector en abscisas y la respuesta en y la proyección en ordenadas. Finalmente, la orientación dominante es la suma de todas las respuestas dentro de una ventana de orientación móvil que abarca un ángulo de $\frac{\pi}{3}$. La respuesta vertical y horizontal son sumadas dentro de la ventana, originando así, un nuevo vector con las sumas resultantes en x y en y . El vector de mayor longitud es el que da la orientación al punto de interés. Ver figura 2.12.

3. Extracción del descriptor

Se comienza construyendo una región cuadrada de tamaño $20s$ alrededor

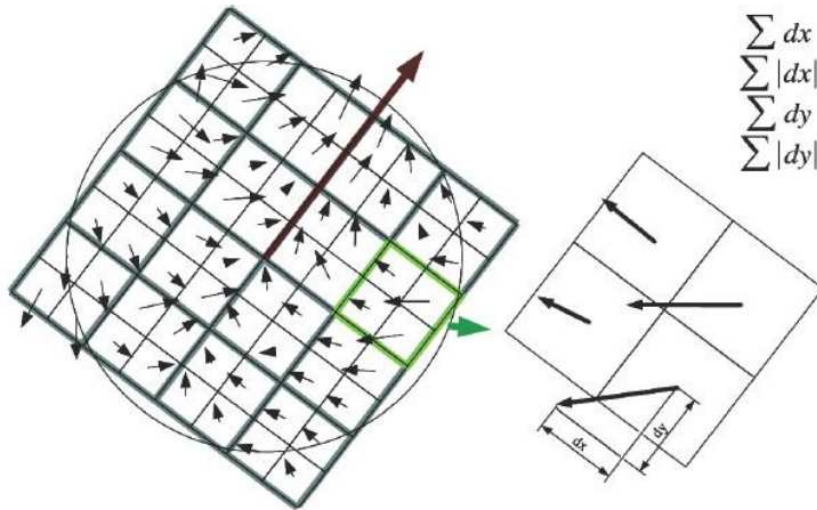


Figura 2.13: Respuestas de Haar en las sub-regiones alrededor del keypoint

del keypoint y orientada según la orientación asignada en el paso anterior. Esta región se subdivide en regiones pequeñas de 4×4 . Cada región pequeña se subdivide de nuevo en regiones 4×4 . En cada pequeña región se calcula las respuestas de Haar y, en consecuencia, una orientación nueva, según ilustra la figura 2.13.

La respuesta de Haar en la dirección horizontal d_x y en la dirección vertical d_y es ponderada con una Gaussiana de $\sigma = 3.3s$ centrada en el punto de interés, con el fin de obtener mayor robustez ante deformaciones geométricas y errores de posicionamiento.

Seguidamente se procede a sumar en cada subregión las respuestas d_x y d_y . Además se suman los valores absolutos de las respuestas $|d_x|$ y $|d_y|$ en cada una de las subregiones, para almacenar información de la polaridad sobre los cambios de intensidad. Con lo cual, cada subregión tiene como descriptor un vector v de longitud 4 que describe su estructura de intensidad. El descriptor total vendrá dado por la expresión 2.16.

$$v = \left(\sum d_x, \sum d_y, \sum |d_x|, \sum |d_y| \right) \quad (2.16)$$

4. Matching *Emparejamiento entre keypoints*

Esta etapa está presente en los demás métodos. No forma parte del método SURF, puede entenderse como un apartado siguiente tras realizar el método SURF (u otro método de detección de puntos de interés).

2. Trabajos y técnicas relacionadas

Este apartado consiste en emparejar los puntos de interés detectados que se corresponden en dos imágenes. Entonces, se emplea en el cálculo la distancia euclídea de sus descriptores. De este modo, un emparejamiento es válido cuando la distancia relativa de los descriptores de ambos puntos sea menor de X veces la distancia respecto al segundo vecino más cercano. Siendo X un valor umbral que ronda el valor de 0.7 (según algunos autores). Se recomienda además añadir restricciones geométricas, con el fin de aminorar falsos emparejamientos.

Capítulo 3

Digitalización del vídeo

Resumen: Este trabajo requiere de unos componentes hardware que limitan la tecnología empleada. En este capítulo se explicarán los componentes empleados, así como la conexión entre ellos. Por un lado aparecen los dispositivos que se alojan en el Glider, y por otro lado los dispositivos que reciben de manera analógica los datos en la estación de tierra. Se hablará además del software realizado para la digitalización del vídeo, a partir de las librerías de OpenCV y de las librerías propias de los dispositivos empleados para digitalizar la secuencia de vídeo.

3.1 Sistema electrónico

El sistema electrónico de vídeo y audio se compone en dos subsistemas, los cuales se conectan por una señal analógica de radiofrecuencia a 2.4 GHz. Por un lado existe un transmisor abordo del Glider UAV. Este se encarga de dos funciones, las cuales son: envío de señal de vídeo y audio, dónde se aloja en formato analógico la información de vídeo y envío de la telemetría a tierra. Cabe mencionar que para este trabajo, el canal de la telemetría no se desarrolló, por lo tanto no se pudo disponer de otros datos diferentes a las imágenes. Por otro lado, en la estación de tierra existe un receptor que recibe la secuencia de vídeo y audio en formato analógico. Además, desde tierra se envían comandos por radio control para controlar la navegación del UAV. La figura 3.1 muestra un esquema de los diferentes subsistemas mencionados.

3. Digitalización del vídeo

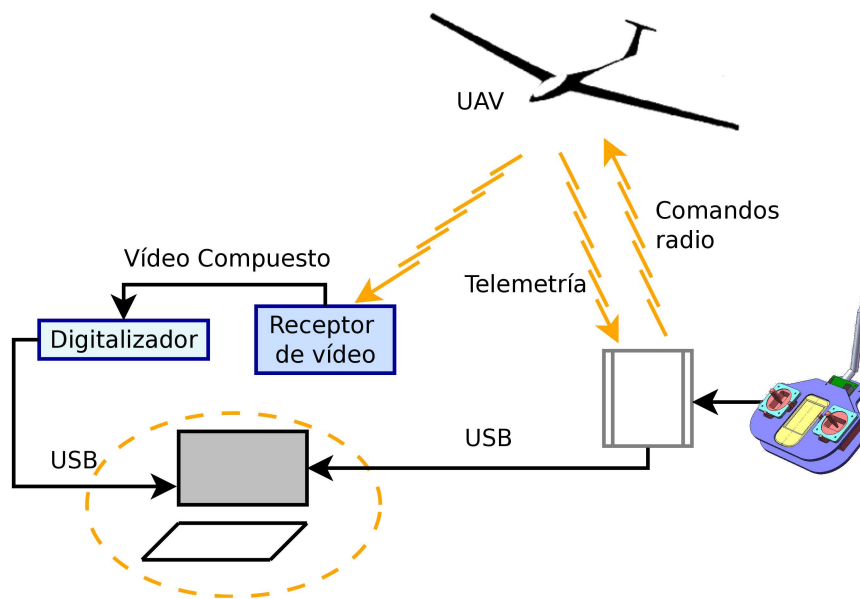


Figura 3.1: Esquema del sistema electrónico/hardware del VANT/UAV

3.1.1 Sistema electrónico a bordo del avión no tripulado

En el UAV se aloja un sistema electrónico que se compone de un receptor, un transmisor de radio, sensores (IMU, GPS, RPM, una cámara), además del OSD y del emisor de vídeo. El circuito electrónico ha sido suministrado por la empresa DMD¹, los cuales desde 2007 suministran productos profesionales de radio control (RC) de largo alcance (LRS) para UAVs, FPV² en aviones, helicópteros, drones, cuadracópteros, vehículos especiales y robots entre otros.

La figura 3.2 muestra un conexionado de radio con un sistema de vídeo de ejemplo para FPV, omitiendo receptor, GPS y los sistemas que no tienen relación directa con el sistema de vídeo. Sobre el modelo de cámara empleado se hablará en el próximo capítulo.

En el esquema mencionado se aprecia el OSD, el cual es el responsable de mezclar la información de vídeo junto con la de GPS para mostrarlas en el cockpit. Aunque no se ha trabajado con él, el GPS es conectado en el OSD, según vemos en la figura 3.2. Del OSD la información se deriva al transmisor de vídeo a 2.4GHz el cual es conectado a una antena. Cabe mencionar que el OSD sobrepone la señal información de muchos sistemas y subsistemas, no sólo del GPS.

Por otro lado, en la figura 3.3 vemos un conexionado idéntico al visto anteriormente al que se le ha añadido los cables del voltímetro de la batería de

¹Digital Micro Devices <http://www.dmd.es>

²FPV: First-person view, en castellano vuelo en primera persona (ver glosario).

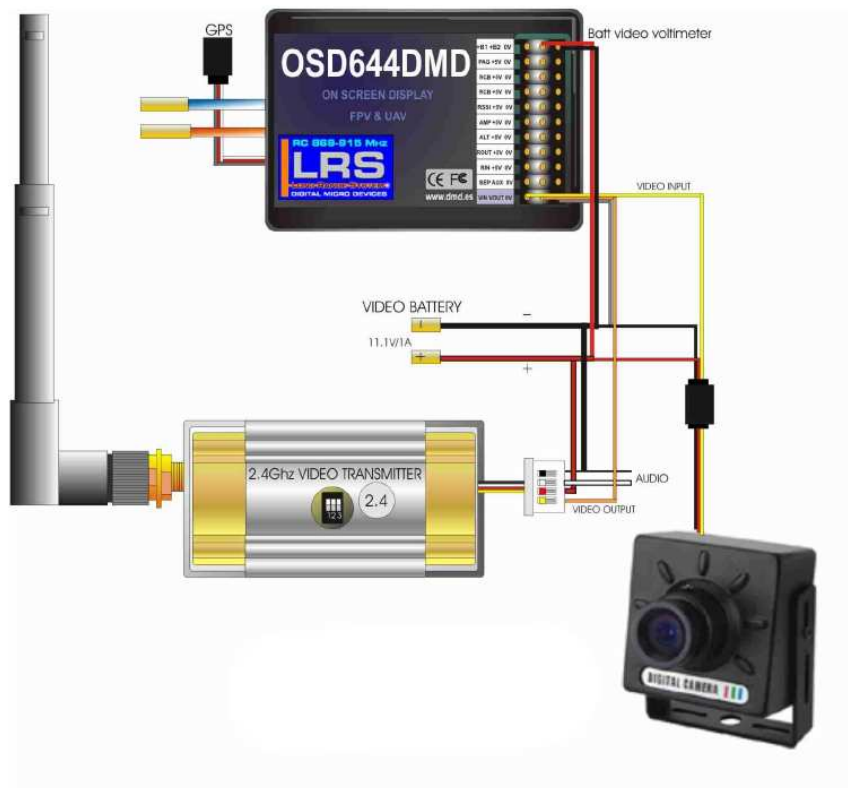


Figura 3.2: Conexionado del subsistema de captación/emisión de vídeo a bordo del UAV.

3. Digitalización del vídeo

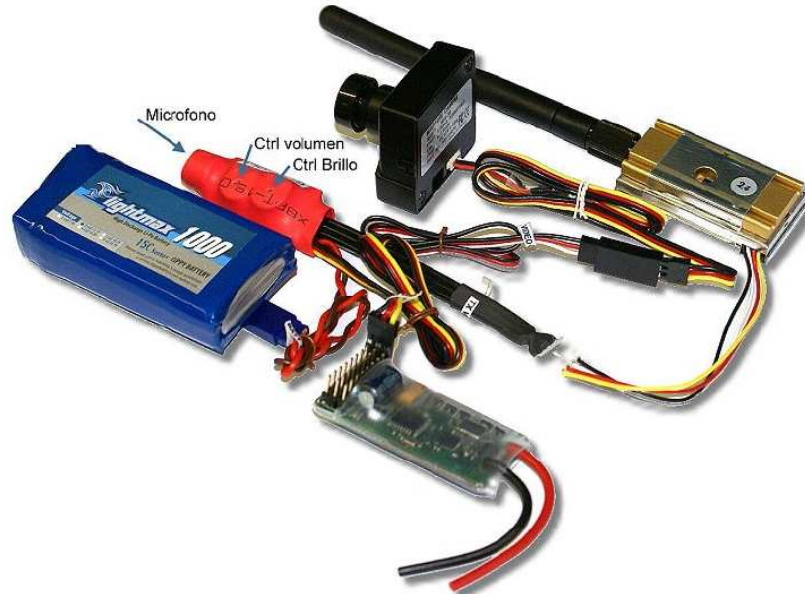


Figura 3.3: Imagen de los dispositivos de abordo del UAV.

vídeo. En este caso también se ha omitido mostrar el receptor de radio, GPS y los sistemas que no tienen relación directa con el sistema de vídeo.

3.1.2 Sistema electrónico en la estación de tierra

El flujo de vídeo es guardado en un PC convencional mediante un digitalizador de imágenes, el cual se encuentra en la estación de tierra. El digitalizador elegido ha sido impuesto debido a los recursos del laboratorio de robótica del Instituto Universitario de Sistemas Inteligentes y Aplicaciones Numéricas en Ingeniería (SIANI).

IC Imaging Control, el digitalizador empleado, es compatible con una entrada de vídeo FireWire y USB. Debido a las restricciones del PC empleado, se ha utilizado el puerto USB para digitalizar el vídeo.

Una característica útil de IC Imaging Control es la posibilidad de que a partir de una transmisión de vídeo analógico, el digitalizador es capaz de convertir las secuencias a su equivalente digital. De esta manera se captura el vídeo digital para su posterior procesado.

En paralelo, en tierra se dispone de una antena conectada a un receptor de 2,4GHz y una salida de vídeo compuesto. Esta es conectada al monitor. Así pues, el vídeo es guardado en tiempo real en un disco duro multimedia o grabadora digital. De esta manera conseguimos una dualidad de la información

digital. Adicionalmente, en tierra el cockpit puede visualizarse con unas gafas especializadas en tiempo real.

3.2 Programa del IC Control con OpenCV

IC Imaging Control dispone de software propio con entorno de ventanas para la captura de vídeo y su posterior almacenamiento en formato AVI. El software además permite cambiar parámetros de configuración del vídeo con el fin de configurar por hardware la secuencia. Además, IC Imaging Control dispone de ejemplos en su página web³ de códigos en lenguajes C++. Para poder ejecutar estos código ha sido necesario solicitarles la última versión, IC Imaging Control 3.2.

La idea que se ha llevado a cabo es poder digitalizar fotograma a fotograma con el software de IC Imaging Control y guardarlos como vídeo en formato AVI, con unos codecs determinados. Por otro lado se trataba de disponer de los fotogramas como imágenes en un formato compatible con la librería OpenCV para poder procesar a tiempo real. La versión de OpenCV empleada en este trabajo es la 2.1. Para realizar el desarrollo de este software se ha hecho uso de la ayuda del foro⁴ de la propia página página de IC Imaging Control, dónde se ha compartido el software final, el cual ha sido comprobado y empleado en este trabajo final de máster. A modo de resumen, se presenta el pseudocódigo empleado en el algoritmo 2

El programa diseñado incluye la posibilidad de procesar en paralelo con dos hebras de trabajo. En una hebra se guardarían los fotogramas previamente digitalizados. En la otra hebra, se da la posibilidad de ir tomando un fotograma y realizar un procesado. Por ejemplo, se da la posibilidad para que de manera oportunista se tome un fotograma y se envíe a una función para realizar un proceso de stitching entre imágenes. La figura 3.4 muestra un esquema que pretende aclarar el software diseñado. Podemos añadir que para cada fotograma capturado y recibido por el transmisor, se repite dicho esquema.

³Ver la web: <http://www.imagingcontrol.com>

⁴Ver la web: <http://www.theimagingsourceforums.com/>

3. Digitalización del vídeo

Algoritmo 2 Grabación de vídeo con IC Imaging Control + OpenCV

- 1: Abrir el digitalizador
 - 2: Configurar parámetros de vídeo con IC Imaging Control
 - 3: Configurar las variables de vídeo en OpenCV 2.1
 - 4: Guardamos memoria de Buffer con IC Imaging Control
 - 5: **mientras** VIDEO = OK **hacer**
 - 6: Adquisición de un fotograma con IC Imaging Control
 - 7: **Hebra 1:** Grabar la imagen con la función de OpenCV
 - 8: **Hebra 2:** (De manera oportunista) Procesado de la última imagen disponible
 - 9: **si** tecla = ESC **entonces**
 - 10: Break
 - 11: **fin si**
 - 12: **fin mientras**
 - 13: Liberar memoria
 - 14: Destruir y cerrar todas las ventanas
-

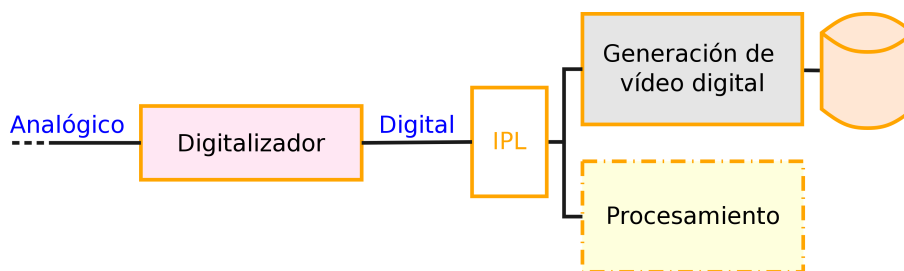


Figura 3.4: Esquema del software de grabación de vídeo con IC Imaging Control + OpenCV

Capítulo 4

Preprocesado de las imágenes

Resumen: El flujo de fotogramas que se obtiene desde el avión es procesado dos veces. Por un lado se ha realizado un filtro que procesa la imagen entrelazada proporcionada por la cámara. Por otro lado se ha realizado una calibración a la cámara de abordo. Los parámetros obtenidos en la calibración se aplican a las imágenes para corregir de distorsión.

4.1 Calibración de la cámara

4.1.1 Modelo de cámara

En este apartado del capítulo se hablará del modelo de cámara, la cual proyecta un punto 3D del espacio en un plano de imagen 2D. Existen diversos trabajos, tales como (Hartley and Zisserman, 2000) (Isern-González, 2003), donde se explican los diferentes modelos de cámaras. Los modelos de cámaras que se emplean ahorran coste y tiempo con respecto a calibrar modelos de cámara 3D. Su justificación aparece en los resultados que se obtienen al simplificar el modelo 3D en un plano de imagen; a pesar de que éstos no sean exactos. Existen varios modelos de cámaras tales como el modelo de Gauss (Hecht and Zajac, 1979) o el modelo pin-hole (Pinhole lens), que es el usado en este trabajo y el más usado en general (Popescu *et al.*).

El modelo de cámara de agujero o “pinhole” son muy simples debido a que no poseen lentes y a que tienen una pequeña apertura. Esta cámara puede ser explicada como una caja a prueba de luz con un orificio diminuto que permite entrar la luz dentro de la caja. Cuando la luz de una imagen pasa por este agujero, entonces se forma una imagen invertida en el lado opuesto de la caja, tal como vemos en la figura 4.1.1.

4. Preprocesado de las imágenes

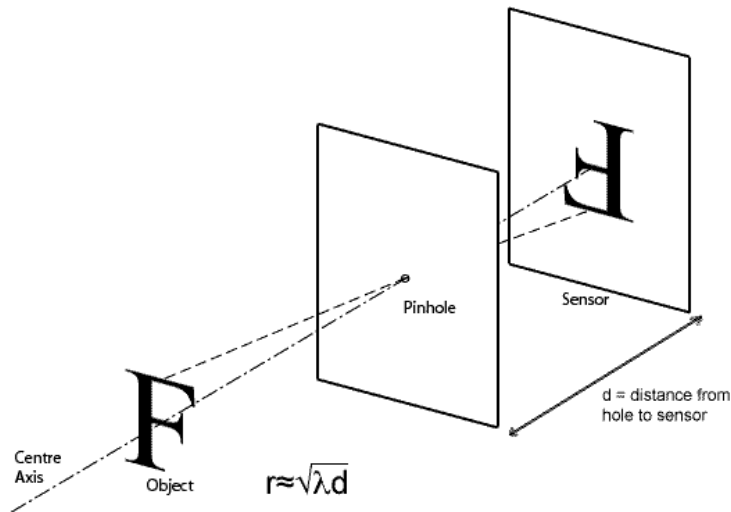


Figura 4.1: Formación de la imagen en una cámara pinhole

En estas cámaras, mientras más pequeño sea el agujero por el que atraviesa la luz, más nítida y definida se consigue la imagen que se forma en el interior de la misma. Por el contrario, la imagen se hace más oscura cuanto más pequeño sea el orificio. Es decir una imagen bien definida será al mismo tiempo más opaca. Por ello, se recomienda que el orificio sea del 1 % o menos de la distancia entre el orificio (pinhole) y el sensor. En la figura 4.1.1, dicha distancia es denominada como d .

4.1.2 Método de Zhang

El método de calibración empleado es el nombrado método de Zhang (Zhang, 2000), (Zhang, 1999). El autor propone una técnica de calibración de cámaras basada en la observación de un patrón de calibración observado desde varias posiciones. Las principales ventajas del método radican en la posibilidad de obtener los parámetros extrínsecos e intrínsecos de la cámara a partir de un patrón. Donde no es necesario conocer la distancia a la que son tomadas las imágenes del patrón; ni métricas sobre puntos de interés. En este trabajo se ha ido presentando un patrón a la cámara en diversas posiciones y orientaciones. Sin embargo, de la misma manera, hubiera sido posible mover la cámara sobre el patrón con la mano. Estas ventajas introducen el adjetivo de flexible en el método: la escena de calibración no necesita de más preparación.

El modelo de calibración empleada sigue la ecuación 4.1, o más en detalle la ecuación 4.2.

$$s \cdot m' = A \cdot [R|t] \cdot M' \quad (4.1)$$

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (4.2)$$

Donde (X, Y, Z) son las coordenadas de un punto 3D en el espacio de coordenadas, (u, v) son las coordenadas del punto de proyección en píxeles. A se llama matriz de cámara, o una matriz de parámetros intrínsecos. (c_x, c_y) es un punto principal (que suele estar en el centro de la imagen). Por otro lado, f_x, f_y es la distancia focal expresada en unidades de píxeles relacionados. Por lo tanto, si una imagen de la cámara es escalado por algún factor, todos estos parámetros deben ser escalados (multiplicados o divididos, respectivamente) por el mismo factor. La matriz de parámetros intrínsecos no depende de la escena vista y, una vez estimados, se puede volver a utilizar (siempre y cuando la longitud focal sea fija (en el caso de zoom)). La matriz $[R|t]$ es conocida como la matriz de parámetros extrínsecos. Se emplea para describir el movimiento de la cámara alrededor de una escena estática, o viceversa, movimiento rígido de un objeto delante de la cámara fija. Es decir, $[R|t]$ traduce las coordenadas de un punto (X, Y, Z) a otro sistema de coordenadas, el cual es fijo con respecto a la cámara.

4.1.3 Herramientas de calibración

Para la calibración se ha hecho uso de las librerías de OpenCV. En ellas se implementa el método de Zangh para cámara pinhole, según se explica en la documentación de OpenCV¹.

Se ha modificado el ejemplo *calibracion.cpp* que suministra la versión 2.1 de OpenCV, para realizar el proceso de calibración de la cámara.

El software que se ha empleado para la calibración es descrito en pseudocódigo en los algoritmos 3 y 4. Las modificaciones que se realizaron a partir del ejemplo de OpenCV son:

- La primera modificación se ha realizado sobre los flags que usa la función de calibración (*cvCalibrateCamera*) de OpenCV. Se han realizado 5 iteraciones sobre la función de calibración, de manera que los parámetros de la calibración son calculados 5 veces de manera diferente. Originalmente, el ejemplo no utiliza el recurso de los flags. Cada uno de estos flags hace un efecto diferente sobre la función de calibración. En el algoritmo 3 y 4 podemos ver el orden en el que han sido aplicados. Se va a comentar cada uno de esos flags a continuación:

¹Confrontar con http://opencv.willowgarage.com/documentation/camera_calibration_and_3d_reconstruction.html

4. Preprocesado de las imágenes

- CV_CALIB_USE_INTRINSIC_GUESS: Parte la calibración con unos valores iniciales de la matriz de calibración, es decir, de la matriz de calibración A , según la ecuación 4.1.
 - CV_CALIB_FIX_ASPECT_RATIO: La relación f_x/f_y es constante en el cálculo de la calibración. Sólo itera el valor de f_y de la matriz de calibración A , según la ecuación 4.1.
 - CV_CALIB_FIX_PRINCIPAL_POINT: Se fija el punto principal de la imagen c_x y c_y justo en el centro de la misma.
 - CV_CALIB_ZERO_TANGENT_DIST: Los parámetros de la distorsión tangencial de la ecuación 4.3, permanecen a 0 en el cálculo de los parámetros de la calibración.
 - CV_CALIB_FIX_K?: Los coeficientes de distorsión se establecen todos a 0 inicialmente a menos que empleemos este flag. En este caso, el parámetro en cuestión $K_?$, tomará un valor inicial.
- Otra modificación ha sido la extracción del 5º coeficiente de distorsión. En total obtenemos estos coeficientes:

$$k_1, k_2, p_1, p_2, [k_3] \quad (4.3)$$

El ejemplo original proporcionaba 4 coeficientes: dos coeficientes de distorsión radial k_1, k_2 y los dos de la distorsión tangencial p_1, p_2 .

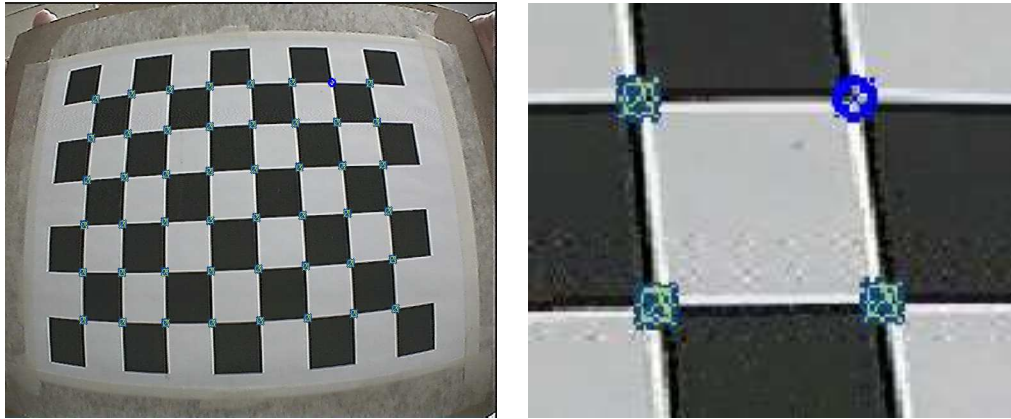
Modificando el programa se ha obtenido el 3º coeficiente de la distorsión radial k_3 .

- La última modificación consiste en determinar de manera visual la retro-proyección de los puntos esquinas de las imágenes del patrón de calibración. Esta modificación consiste en pintar sobre las imágenes de entrada al programa de calibración los puntos esquinas en diferente color, como muestra la figura 4.1.3.

En azul oscuro se han pintado un círculo y una cruz de los puntos esquinas detectados originalmente.

En amarillo verdoso se ha pintado un círculo y una cruz de los puntos esquinas corregidos de distorsión a los que se han aplicado la corrección de los parámetros de distorsión (ecuación 4.3). Es decir, sobre las imágenes corregidas de distorsión se han calculado los puntos esquinas. Estos puntos esquinas han sufrido la deformación aplicando los parámetros de distorsión.

En azul se han pintado aquellos puntos ideales. Es decir, cuando las coordenadas de un punto amarillo verdoso coincide con un punto azul



(a) Retroproyección de los puntos esquinas del patrón de calibración (b) Detalle de los puntos dibujados sobre la plantilla

Figura 4.2: Plantilla de calibración pintada

oscuro, se pinta un círculo y una cruz en azul. Es un caso ideal, pues en esos puntos los parámetros de la distorsión son muy acertados.

4. Preprocesado de las imágenes

Algoritmo 3 Calibración de la cámara con las herramientas de OpenCV
(Parte 1)

```
1: para  $i = 1 \rightarrow n_{imgpatron}$  hacer
2:   Cargo  $img_i$ 
3:   Guardo puntos esquinas de cada patrón
4:   Guardo tamaño de imágenes, tamaño del tablero, número de cuadros,
   número de puntos esquinas y número de imágenes
5: fin para
6: para  $i = 1 \rightarrow 5$  hacer
7:   si  $i = 1$  entonces
8:     flags = CV_CALIB_FIX_PRINCIPAL_POINT OR
     CV_CALIB_FIX_ASPECT_RATIO OR
     CV_CALIB_ZERO_TANGENT_DIST OR CV_CALIB_FIX_K3;
9:   fin si
10:  si  $i = 2$  entonces
11:    flags = CV_CALIB_USE_INTRINSIC_GUESS OR
    CV_CALIB_FIX_PRINCIPAL_POINT OR
    CV_CALIB_FIX_ASPECT_RATIO OR
    CV_CALIB_ZERO_TANGENT_DIST OR CV_CALIB_FIX_K3;
12:  fin si
13:  si  $i = 3$  entonces
14:    flags = CV_CALIB_USE_INTRINSIC_GUESS OR
    CV_CALIB_FIX_ASPECT_RATIO OR
    CV_CALIB_ZERO_TANGENT_DIST OR CV_CALIB_FIX_K3;
15:  fin si
16:  si  $i = 4$  entonces
17:    flags = CV_CALIB_USE_INTRINSIC_GUESS OR
    CV_CALIB_ZERO_TANGENT_DIST OR CV_CALIB_FIX_K3;
18:  fin si
19:  si  $i = 5$  entonces
20:    flags = CV_CALIB_USE_INTRINSIC_GUESS;
21:  fin si
22:  Calibro con cvCalibrateCamera2
23: fin para
24: Guardo resultados en Fichero
```

Algoritmo 4 Calibración de la cámara con las herramientas de OpenCV (Parte 2)

```

1: para  $i = 1 \rightarrow n_{imgpatron}$  hacer
2:   Guardo ptos. esquinas retroproyectados en imágenes corregidas de distorsión
3:   Retroproyecto ptos. iniciales con parámetros obtenidos en la calibración
4:   para  $j = 1 \rightarrow n_{ptos\_esquinas}$  hacer
5:     si  $ptos\_esquinas_i = ptos\_retroproyectados_i$  entonces
6:       Pinto círculo con cruz en azul en la esquina j de la imagen i
7:     si no
8:       Pinto círculo con cruz en azul oscuro en la esquina j de la imagen i
9:       Pinto círculo con cruz en amarillo verdoso en la esquina j de la imagen i
10:    fin si
11:  fin para
12:  Corrijo la imagen i de distorsión
13:  Muestro y guardo la imagen i corregida
14: fin para

```

4.1.4 Resultados experimentales de la calibración de la cámara

El programa proporciona un fichero con diferentes parámetros de salida. Para determinar un criterio de control de la calidad de la calibración, se ha intentado minimizar los errores de retroproyección de los punto de calibración. Es decir, los puntos esquinas. Para ello se ha empleado una manera analítica (tabla 4.1) y otra visual (figura 4.1.4) con el fin de fijar un criterio de parada. El error calculado se ha realizado en base a la norma que se muestra en la ecuación 4.4.

$$\varepsilon = \frac{\|arr_1 - arr_2\|_{L1}}{\sum Ptos. esquina} = \frac{\sum arr_1(I) - arr_2(I)}{\sum Ptos. esquina} \quad (4.4)$$

Donde arr_1 son los puntos esquina retroproyectados en las imagen original y arr_2 los puntos esquinas de la imagen original. Originalmente se disponían de 7 imágenes de la plantilla de calibración. Tras las pruebas realizadas, se determinaron que las 5 imágenes empleadas eran las que proporcionaban mejores resultados.

Adicionalmente, las coordenadas del centro de la imagen (c_x, c_y según la ecuación 4.2), han sido de ayuda para concluir con una calibración aceptable. Estos valores los podemos ver en los resultados de la matriz de calibración A (ecuación 4.1), en la tabla 4.2. Nótese que el tamaño de las imágenes empleadas son: $360x288 px$. Así mismo, en la tabla 4.3 se incluyen los parámetros de los

4. Preprocesado de las imágenes

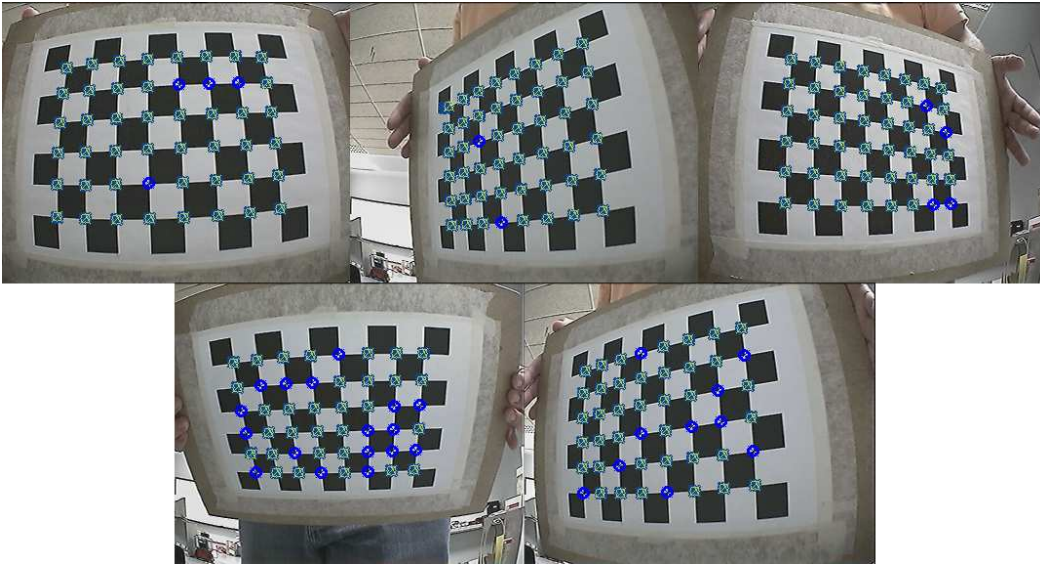


Figura 4.3: Imágenes de las plantillas de calibración distorsionadas con puntos pintados: en azul oscuro los puntos esquinas; en amarillo verdoso los puntos esquinas corregidos y retroproyectados; en azul los puntos cuyas retroproyecciones coinciden con los puntos esquinas originales.

Imagen 1	Imagen 2	Imagen 3	Imagen 4	Imagen 5	media
$1.49 \cdot 10^{-1}$	$3.75 \cdot 10^{-1}$	$1.86 \cdot 10^{-1}$	$1.18 \cdot 10^{-1}$	$1.25 \cdot 10^{-1}$	$1.91 \cdot 10^{-1}$

Cuadro 4.1: Errores de retroproyección de los puntos esquinas de cada imagen patrón

coeficientes de distorsión tangencial p_2 y radial k_2 , vistos en la ecuación 4.3. La distorsión tangencial es casi nula, por ello, estos parámetros resultan más bajos que los de la distorsión radial.

No usamos los parámetros extrínsecos sino los intrínsecos. Debido a que en nuestro caso de aplicación la cámara no es fija sino su punto de referencia varía con el tiempo. Los parámetros extrínsecos nos dan información de la rotación y traslación geométrica de la imagen con respecto a la cámara. Con lo cual, son parámetros interesantes para cámaras fijas, por ejemplo cámaras

338.18	0	176.34
0	358.35	143.69
0	0	1

Cuadro 4.2: Matriz de calibración (matriz A en la ecuación 4.1).

k_1	k_2	p_1	p_2	$[k_3]$
$-5.13 \cdot 10^{-1}$	$-9.73 \cdot 10^{-2}$	$-9.79 \cdot 10^{-4}$	$-6.51 \cdot 10^{-4}$	$9.74 \cdot 10^{-1}$

Cuadro 4.3: Coeficientes de la distorsión radial $k_?$ y tangencial $p_?$ (ecuación 4.3).

de vigilancia, no cámaras móviles como en nuestro caso.

Finalmente, en la figura 4.1.4 se muestran las imágenes empleadas y el efecto de cómo han quedado tras aplicarle los parámetros de calibración. Estas imágenes han sido obtenidas de un vídeo realizado con la misma cámara.

4.2 Filtrado

Existen dos modos de visualizar vídeo según las cámaras: el “barrido” progresivo y el “escaneo” entrelazado, según la cámara o monitor empleado.

El escaneo progresivo se usa en monitores de ordenador, proyectores, muchas cámaras y televisión digital y visualiza todas las líneas horizontales de una sola vez como si fuesen un único fotograma.

El escaneo entrelazado se usa en los formatos estándar de televisión NTSC, PAL y SECAM y visualiza sólo la mitad de las líneas horizontales en cada pasada (cada fotograma se divide en dos campos, el primero contiene todas las líneas de número impar y el segundo las de número par). Esto se ha utilizado tradicionalmente para obtener altas velocidades de refresco (50 Hz en PAL, 60 Hz en NTSC) con únicamente la mitad del flujo de datos. A cambio, la resolución horizontal queda afectada, ya que cada fotograma se compone de dos “medias imágenes”, que al mezclarse pueden dar lugar a parpadeo, doble imagen (“ghosting”), etc.

La cámara a bordo del UAV es entrelazada. Debido a ello la señal de vídeo es entrelazada, la cual, como se ha comentado, está compuesta por dos campos horizontales, el campo impar (líneas de vídeo horizontales impares) y el campo par (líneas de vídeo horizontales pares). El término entrelazado se refiere en la forma en que los campos son barridos secuencialmente, primero el campo impar y después del campo par, y entrelazados posteriormente para formar la imagen. Este efecto causa problemas a la hora de registrar objetos en movimiento, como es nuestro caso.

Debido a ello, se ha realizado un pequeño programa, cuyo pseudocódigo lo podemos ver en el algoritmo 5. El programa consiste en un simple submuestreo con interpolación bilineal. De esta manera se “limpian” las imágenes desentrelazándolas y submuestreándolas en horizontal. En la imagen 4.2 podemos ver una imagen original (imagen 4.5(a)) obtenida desde la cámara. Dicha imagen

4. Preprocesado de las imágenes

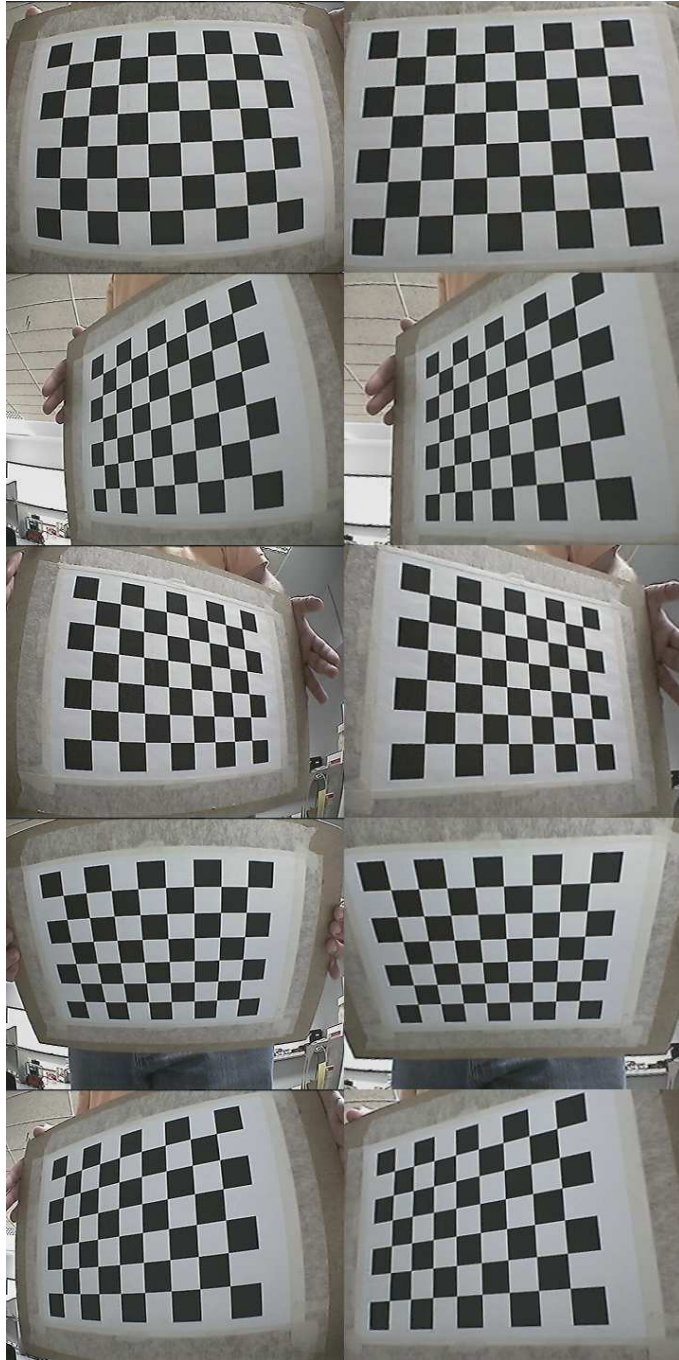


Figura 4.4: Imágenes de las plantillas de calibración con distorsión y corregidas de distorsión.



(a) Imagen Original



(b) Imagen líneas horizontales pares

(c) Imagen líneas horizontales impares

Figura 4.5: Imagen original e imágenes resultantes tras aplicarle el filtrado

es de muy mala calidad, con mucho ruido debido al problema de la cámara entrelazada. Este programa divide la imagen en dos, una con las líneas horizontales pares y otra con las líneas horizontales impares, como vemos en las imágenes 4.5(b) e 4.5(c). El resultado es bastante aceptable y nos permite reutilizar las imágenes tomadas con el UAV. El inconveniente que presenta es que se reduce aún más la resolución de las imágenes, consiguiendo imágenes útiles de $360 \times 288 \text{ px}$. Es decir, se obtienen imágenes con un cuarto de las imágenes originales.

Finalmente comentar que las cámaras que se deberían utilizar a bordo de los aviones de RC deberían ser progresivas y no entrelazadas, como en nuestro caso.

4. Preprocesado de las imágenes

Algoritmo 5 Filtrado de las imágenes obtenidas en el vuelo del UAV.

```
1: para  $i = 1 \rightarrow n_{img}$  hacer
2:   Cargo imagen  $i$ 
3:   Obtengo el ancho y alto de la imagen  $i$ 
4:   para  $j = 1 \rightarrow n_{img}$  hacer
5:     si  $fila == par$  entonces
6:       Copio Fila  $j$  de Imagen original en imagen par
7:     si no
8:       Copio Fila  $j$  de Imagen original en imagen impar
9:     fin si
10:  fin para
11:  Ancho imagen par e impar = Ancho imagen original / 2
12:  Alto imagen par e impar = Alto imagen original / 2
13:  Guardo imagen par e impar
14: fin para
```

Capítulo 5

Estudio de las imágenes con la herramienta SURF

Resumen: Se mostrará en este capítulo el estudio realizado con las imágenes aéreas tras realizar el preprocesado, comentado en el capítulo anterior. Para ello se parte de un conjunto de imágenes concreto, extraído de un vídeo que sobrevoló una misma escena en varias ocasiones y con diferentes condiciones de vuelo. El estudio se basa en analizar el uso de SURF, en la versión implementada en OpenCV. Se han realizando filtros y se ha estudiado la posibilidad de conseguir el emparejamiento más adecuado entre las imágenes seleccionadas. SURF ha sido descrito con anterioridad en el capítulo 2.

5.1 Conjunto de imágenes a estudiar

Para la realización de este estudio se ha usado un vídeo que pertenece al vuelo sobre una zona de barrancos, tomado con el vehículo aéreo no tripulado (VANT). La cámara, en posición cenital, ha tomado un vídeo de la zona pasando varias veces por ciertas escenas a diferente altura, orientación o condiciones de luminosidad. No existen datos externos procedentes de dispositivos tales como GPS, giróscopos, etc.

Debido a ello, se ha observado el vídeo y se han extraído fotogramas de una zona por donde el vehículo aéreo ha pasado a diferentes alturas y en diferentes condiciones.

En la figura 5.1 se observa el conjunto de 7 imágenes extraídas del vídeo. Estos serán los datos con los que se hará el estudio. Por tanto, las condiciones experimentales llevan consigo pocas imágenes. Además aparece el efecto de

5. Estudio de las imágenes con la herramienta SURF



Figura 5.1: Conjunto de imágenes en el orden en que serán mencionadas en el capítulo. Empezando a numerarlas por la imagen superior derecha y terminando por la imagen inferior izquierda.

aliasing perceptual, donde muchas regiones de una imagen se confunde con otras regiones de la misma imagen.

En el conjunto aparecen imágenes muy diferentes unas de otras y algunas que son muy parecidas con alguna otra. Con el fin de realizar un estudio con el menor sesgo posible, se ha separado el conjunto en dos subconjuntos, los cuales se estudiarán por separado. El criterio de separación es la altura a la que han sido tomadas las imágenes. Por ello, agrupamos las imágenes según dos tipos de alturas. En la figura 5.2 tenemos 4 imágenes que serán estudiadas por separadas. Con un recuadro rojo se remarca la imagen que será usada como referencia en el estudio. La justificación se debe a que una línea futura es georreferenciar estas imágenes en mapas de menor resolución. Es decir, poder encajar en una imagen aérea muy lejana una imagen de mayor resolución, por ello la imagen de referencia es la que ha sido tomada a más altura.

Por otro lado, en la figura 5.3 señalamos el subconjunto de imágenes tomadas a menor altura. De nuevo se señala en rojo la imagen que será empleada como referencia. En este caso, se ha elegido esa imagen debido a que era central, es decir, contenía una escena central con respecto a las otras dos.

Por último, apuntar que debido al digitalizador empleado, en los bordes de las imágenes aparece un recuadro negro. Debido a que este borde tiene las mismas dimensiones en todo el conjunto, en el software diseñado elegiremos una región de interés (ROI) donde desaparecerá dicho marco negro.



Figura 5.2: Subconjunto 1



Figura 5.3: Subconjunto 2

5. Estudio de las imágenes con la herramienta SURF

5.2 Eliminación de puntos no discriminantes

En este apartado se pretende tomar una imagen y discriminar puntos que sean parecidos dentro de la misma imagen. Por ejemplo, muchas de las escenas de las imágenes muestran zonas muy parecidas, que corresponden a zonas verdes como son las tabaibas. Como posibilidad a la hora de mejorar el emparejamiento de puntos en las imágenes se ha decidido eliminar ciertos puntos de interés que pueden llegar a ser falsos positivos o distractores. La técnica empleada se realizará a partir de la distancia euclídea.

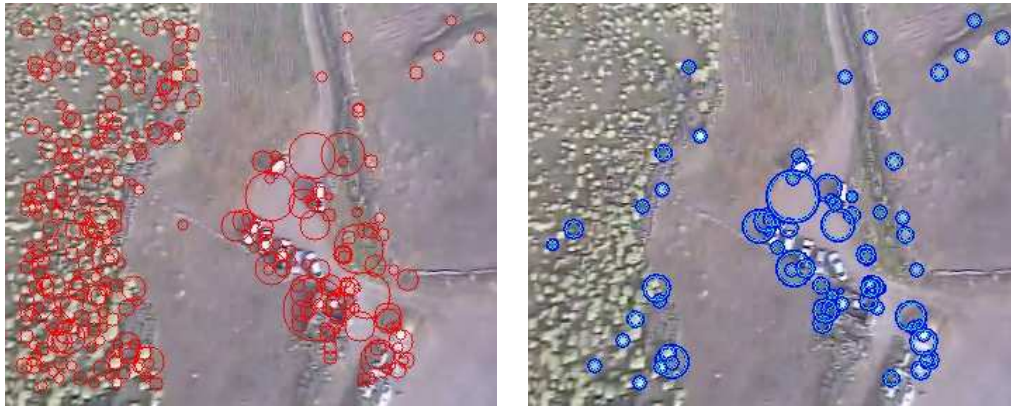
Debido a ello, se ha diseñado una función la cual acoge dos entradas y devuelve una salida. Una entrada son los keypoints detectados en una imagen al aplicarle SURF y otra entrada son los descriptores de cada keypoint. La salida consiste en un vector que posee los índices por orden de keypoint singular. Es decir, el elemento “0” aloja el índice i del keypoint que menos se parece al resto de keypoints detectados. Así sucesivamente. Este método parte del cálculo de la matriz de distancias euclídeas de los descriptores que pertenecen a los keypoints encontrados. Nótese que esta matriz se considerará triangular superior o inferior a efectos de cálculo.

Posteriormente realizamos un cálculo de interdistancias de la matriz distancia euclídea. Para ello se barre la matriz de distancias euclídeas sumando cada elemento que pertenece a la diagonal superior y finalmente se calcula la media. Este valor se usará como umbral en pasos posteriores.

En el siguiente paso se trata de ver los descriptores de cada uno de los keypoints que sean menores que la media. Es decir, se toma la fila i de la matriz distancia euclídea y se cuenta el número de descriptores j que sean menores que el valor medio de la matriz de distancias euclídeas. El número de keypoints que sean menor que el valor medio para la fila i , es decir, para el keypoint i , se guarda en el elemento i de un nuevo vector v . Con lo cual, obtenemos un vector cuya posición nos indica el keypoint en el orden que ha sido encontrado según SURF y cuyo valor nos indica el número de descriptores por debajo de la media.

Para un keypoint i , cuanto mayor sea el número de descriptores o cuantos más impactos tenga por encima de la media, más singular será ese keypoint, es decir, menos se parecerán sus descriptores a los descriptores de otros keypoints. Por ejemplo, si dos keypoints son iguales, su valor de distancia euclídea en la matriz de distancias euclídeas será nulo. Un valor muy cercano a 0 en la matriz de distancias euclídeas significa que el keypoint i y el keypoint j son muy parecidos. A alto nivel, se dice que dos pequeña zonas en una escena, una i y otra j , son muy parecidas, por ejemplo, dos arbustos verdes.

Posteriormente se ordena el vector v de modo que el mayor valor del vector será el primer valor y así sucesivamente, generando un vector ordenado ψ .



(a) Todos los keypoints detectados con SURF (b) Keypoints más singulares detectados

Figura 5.4: En la figura (a) aparecen todos los keypoints detectados con la herramienta SURF implementada en OpenCV. En la figura (b) observamos los puntos singulares tras el filtrado realizado

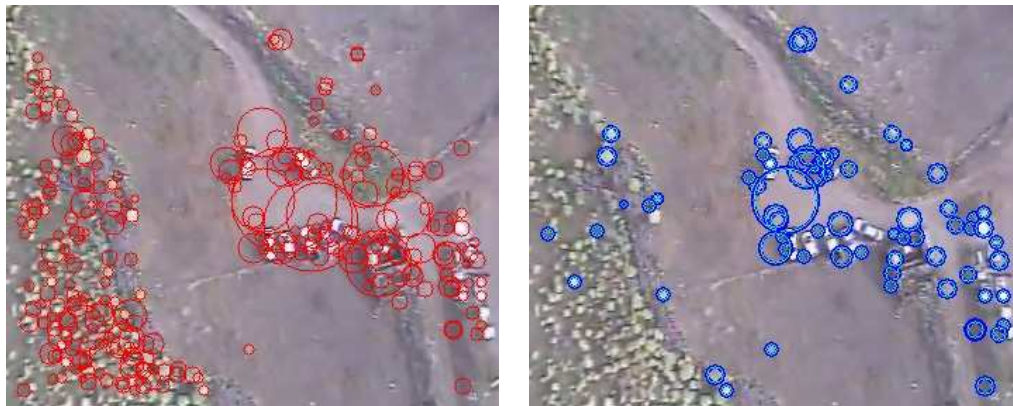
Comparando ambos vectores v y ψ , finalmente generamos un nuevo vector ν donde se aloja el índice de los keypoints de más singular (o del keypoint cuyos descriptores son menos parecidos con otros) hasta el keypoint cuyos descriptores se parecen mucho con otro keypoint. Este vector ν será el parámetro de salida en la función diseñada.

Una explicación gráfica se ilustra en la figura 5.4. En ella podemos observar dos imágenes de la misma escena. La figura (a) representa todos los puntos detectados automáticamente con la herramienta SURF implementada en OpenCV. Tras realizar el procedimiento explicado, se redibujan sólo el 25% de los puntos detectados, ver figura (b). Como vemos, los keypoints detectados en la parte de arbustos ahora se han suprimido en su gran mayoría. Esto coincide con lo esperado tras observar a alto nivel el comportamiento de SURF. Pues muchos arbustos son parecidos, y en consecuencias, sus keypoints también lo son. Estos puntos detectados podemos decir que actúan como “distractores” en la tarea siguiente de emparejamiento. De manera análoga sucede en el segundo ejemplo ilustrado en la figura 5.5. Se pretende nuevamente realzar con otra imagen de la misma escena, tomada con una orientación diferente, el resultado de este filtrado.

5.3 Emparejamientos manuales

En este apartado se ha tratado de realizar otro estudio de los descriptores de las imágenes aéreas. Se pretende estudiar los descriptores de las imágenes con

5. Estudio de las imágenes con la herramienta SURF



(a) Todos los keypoints detectados con SURF (b) Keypoints más singulares detectados

Figura 5.5: En la figura (a) aparecen todos los keypoints detectados con la herramienta SURF implementada en OpenCV. En la figura (b) observamos los puntos singulares tras el filtrado realizado

el fin de emparejar imágenes con escenas parecidas. En este apartado se han seleccionado manualmente 9 puntos que previamente han sido detectados en el conjunto de imágenes que se estudia. Estos 9 puntos han sido seleccionados después de hacer una inspección visual de puntos que aparecían siempre y/o eran más singulares a la vista. Para explicar la metodología se hará hincapié en cómo se ha estudiado un keypoint. La explicación se extiende a tantos puntos de interés como se quiera estudiar. En la figura 5.6 visualizamos los 9 puntos seleccionados en la imagen de referencia. A partir de ahora, se nombrarán los puntos según este esquema.

Por otro lado, cabe destacar que para este estudio se han empleado 128 descriptores por cada keypoint. Los descriptores tienen valores muy pequeños, del orden de centésimas o milésimas. Por ello, en el cálculo de la distancia euclídea, como se comentará, se puede generar errores de redondeo.

5.3.1 Rango de emparejamiento de puntos manuales según sus descriptores

Se ha empleado la herramienta SURF, implementada en OpenCV para determinar keypoints sobre las imágenes. Se ha seleccionado un punto sobre la primera imagen y se han guardado sus 128 descriptores. El criterio de selección ha sido visual, es decir, el punto seleccionado ha sido un punto que visualmente tiene características diferentes con otros keypoints. A modo de ejemplo se señalan en la figura 5.7 el keypoint número 8. En ella se puede observar el mismo keypoint coloreado en rojo, detectado en 4 de 7 imágenes.



Figura 5.6: Keypoints manualmente seleccionados en la imagen de referencia

En las 3 restantes no se detectaba o no aparecía dicho punto de interés.

Posteriormente se ha seleccionado el mismo punto en todas las imágenes de test. La herramienta SURF en ocasiones no encontraba automáticamente el punto en cuestión en alguna imagen. En ese caso, para esa imagen, no existían datos de sus descriptores.

Seguidamente, una vez se almacenan todos los descriptores, se procede a realizar el cálculo de la distancia euclídea de los descriptores, ecuación 5.1. Ello da lugar a una matriz triangular de 7×7 , debido a que el conjunto de imágenes comprende 7 fotografías aéreas. Nótese que en la ecuación 5.1, m y n hace referencia a dos imágenes distintas del conjunto de 7 imágenes.

$$\text{dist. euclídea}_{mn} = \sqrt{\sum_{i=1}^{128} (\text{Descr}_{m_i} - \text{Descr}_{n_i})^2} \quad (5.1)$$

Luego, se procede a calcular la media de la triangular superior de la matriz de distancias euclídeas de los descriptores del keypoint encontrado en algunas de las imágenes del conjunto. Este valor es importante en este paso, pues indica un valor umbral. Este valor es un dato que indicará la distancia euclídea promedio de los descriptores del mismo punto encontrado en n imágenes. Nótese, que en el caso ideal, los descriptores de un mismo punto son iguales y por tanto, este valor sería 0.

5. Estudio de las imágenes con la herramienta SURF



Figura 5.7: Localización del keypoint n° 8 en una serie de imágenes

5.3.2 Experimento 1

A continuación se muestran los resultados encontrados en 9 puntos de interés seleccionados a mano, siguiendo la metodología explicada. Se muestran las matrices de distancias euclídeas calculadas. Nótese que en algunos elementos aparece *NaN* (Not a Number). Esto ocurre para la fila i y columna j , las cuales corresponden a dos fotogramas determinados (fotograma i y fotograma j), cuando no comparten el keypoint que se seleccionó en la imagen de referencia. En este experimento se han empleado 128 descriptores en el algoritmo SURF.

1. Punto 1:

$$Pto_1 = \begin{pmatrix} 0 & 0.4439 & 0.5724 & 0.1937 & 0.3279 & NaN & NaN \\ - & 0 & 0.3195 & 0.4532 & 0.4062 & NaN & NaN \\ - & - & 0 & 0.5179 & 0.5271 & NaN & NaN \\ - & - & - & 0 & 0.3313 & NaN & NaN \\ - & - & - & - & 0 & NaN & NaN \\ - & - & - & - & - & 0 & NaN \\ - & - & - & - & - & - & 0 \end{pmatrix} \quad (5.2)$$

2. Punto 2:

$$Pto_2 = \begin{pmatrix} 0 & 0.4302 & 0.4214 & 0.5055 & NaN & 0.4722 & NaN \\ - & 0 & 0.2608 & 0.5139 & NaN & 0.2669 & NaN \\ - & - & 0 & 0.3639 & NaN & 0.4107 & NaN \\ - & - & - & 0 & NaN & 0.6169 & NaN \\ - & - & - & - & 0 & NaN & NaN \\ - & - & - & - & - & 0 & NaN \\ - & - & - & - & - & - & 0 \end{pmatrix} \quad (5.3)$$

3. Punto 3:

$$Pto_3 = \begin{pmatrix} 0 & 0.4616 & 0.5973 & 0.3876 & NaN & 0.5698 & 0.4388 \\ - & 0 & 0.2865 & 0.2053 & NaN & 0.4625 & 0.3997 \\ - & - & 0 & 0.3688 & NaN & 0.4433 & 0.5401 \\ - & - & - & 0 & NaN & 0.5199 & 0.3915 \\ - & - & - & - & 0 & NaN & NaN \\ - & - & - & - & - & 0 & 0.5405 \\ - & - & - & - & - & - & 0 \end{pmatrix} \quad (5.4)$$

5. Estudio de las imágenes con la herramienta SURF

4. Punto 4:

$$Pto_4 = \begin{pmatrix} 0 & 0.5846 & NaN & 0.3594 & NaN & NaN & NaN \\ - & 0 & NaN & 0.5045 & NaN & NaN & NaN \\ - & - & 0 & NaN & NaN & NaN & NaN \\ - & - & - & 0 & NaN & NaN & NaN \\ - & - & - & - & 0 & NaN & NaN \\ - & - & - & - & - & 0 & NaN \\ - & - & - & - & - & - & 0 \end{pmatrix} \quad (5.5)$$

5. Punto 5:

$$Pto_5 = \begin{pmatrix} 0 & 0.4651 & 0.4700 & 0.4570 & NaN & 0.3627 & NaN \\ - & 0 & 0.3616 & 0.2466 & NaN & 0.3775 & NaN \\ - & - & 0 & 0.3863 & NaN & 0.3736 & NaN \\ - & - & - & 0 & NaN & 0.4213 & NaN \\ - & - & - & - & 0 & NaN & NaN \\ - & - & - & - & - & 0 & NaN \\ - & - & - & - & - & - & 0 \end{pmatrix} \quad (5.6)$$

6. Punto 6:

$$Pto_6 = \begin{pmatrix} 0 & 0.3540 & 0.5061 & 0.3320 & NaN & NaN & 0.5429 \\ - & 0 & 0.2963 & 0.2919 & NaN & NaN & 0.4108 \\ - & - & 0 & 0.4787 & NaN & NaN & 0.3544 \\ - & - & - & 0 & NaN & NaN & 0.5864 \\ - & - & - & - & 0 & NaN & NaN \\ - & - & - & - & - & 0 & NaN \\ - & - & - & - & - & - & 0 \end{pmatrix} \quad (5.7)$$

7. Punto 7:

$$Pto_7 = \begin{pmatrix} 0 & 0.4313 & 0.3925 & 0.5148 & 0.4944 & NaN & NaN \\ - & 0 & 0.1340 & 0.2779 & 0.5700 & NaN & NaN \\ - & - & 0 & 0.3359 & 0.5903 & NaN & NaN \\ - & - & - & 0 & 0.4755 & NaN & NaN \\ - & - & - & - & 0 & NaN & NaN \\ - & - & - & - & - & 0 & NaN \\ - & - & - & - & - & - & 0 \end{pmatrix} \quad (5.8)$$

8. Punto 8:

$$Pto_8 = \begin{pmatrix} 0 & 0.4684 & 0.4636 & 0.3736 & NaN & NaN & NaN \\ - & 0 & 0.1307 & 0.3090 & NaN & NaN & NaN \\ - & - & 0 & 0.2903 & NaN & NaN & NaN \\ - & - & - & 0 & NaN & NaN & NaN \\ - & - & - & - & 0 & NaN & NaN \\ - & - & - & - & - & 0 & NaN \\ - & - & - & - & - & - & 0 \end{pmatrix} \quad (5.9)$$

9. Punto 9:

$$Pto_9 = \begin{pmatrix} 0 & 0.3490 & 0.3139 & 0.3315 & NaN & 0.2555 & NaN \\ - & 0 & 0.1403 & 0.1523 & NaN & 0.3750 & NaN \\ - & - & 0 & 0.1716 & NaN & 0.3797 & NaN \\ - & - & - & 0 & NaN & 0.3839 & NaN \\ - & - & - & - & 0 & NaN & NaN \\ - & - & - & - & - & 0 & NaN \\ - & - & - & - & - & - & 0 \end{pmatrix} \quad (5.10)$$

En la figura 5.8 se pueden observar los intervalos calculados para cada punto seleccionado a mano y su media. Se muestra en el eje horizontal los 9 keypoints seleccionados y en el eje vertical el valor de las distancias euclídeas. En color rosa se señala el punto medio. La amplitud del margen donde se encuentran los valores de distancias euclídeas, queda señalado en dicho gráfico. Adicionalmente, se observa en el gráfico varios keypoints cuyas cruces (llamadas “Distancias Euclídeas” en la leyenda del gráfico) están prácticamente superpuestas unas con otras. Una posible interpretación es que pueden existir keypoints muy cercanos, incluso superpuestos, al keypoint real. Al ser señalado, podría haberse señalado erróneamente un keypoint prácticamente

5. Estudio de las imágenes con la herramienta SURF

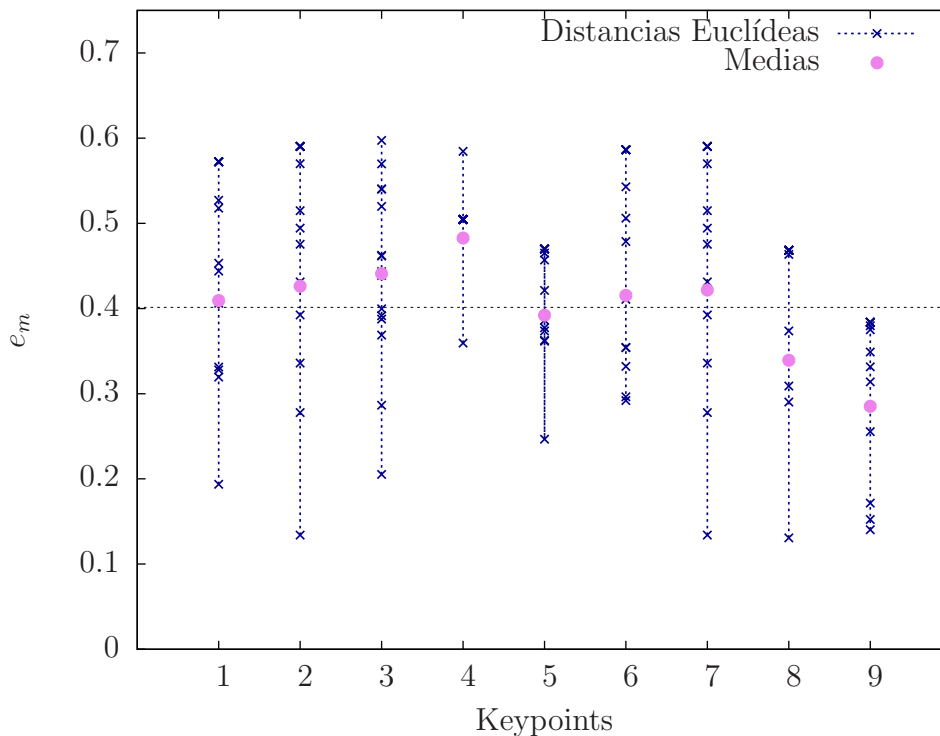


Figura 5.8: Intervalos de las distancias euclídeas para keypoints idénticos en distintas imágenes

superpuesto, debido a que en las imágenes aparecen muchos puntos cercanos detectados. Sin embargo, este sesgo podría ser mejorado definiendo un radio de localización alrededor del punto seleccionado y buscando el keypoint más similar dentro de ese radio de localización/imprecisión.

En resumen, en esta gráfica se observa la medida de distancia euclídea que indica cuán similares son dos detecciones del mismo punto de interés sobre una serie de imágenes. Evidentemente lo ideal sería encontrar que esta distancia fuese lo menor posible o nula, pues estaríamos ante dos keypoints, de dos imágenes distintas, que son idénticos. Sin embargo, el comportamiento es muy dispar y poco homogéneo con relación a la distancia euclídea de los keypoints seleccionados.

5.3.3 Experimento 2

En este caso se han comparado con la imagen de referencia las imágenes de test. En ellas se compara un keypoint seleccionado a mano en la imagen de referencia con el mismo keypoint en una imagen i de test, según la distancia

Análisis Fotogramétrico de Imágenes Adquiridas por un UAV

Img.	Kp. Conf.	Lim. In.	Lim. Sup.	Rango
1	3	0.444	0.449	0.005
2	17	0.574	0.581	0.007
3	0	0.197	0.387	0.190
4	0	0.336	0.416	0.080
5	-	-	-	-
6	-	-	-	-
Media	5	0.387	0.458	0.071

Cuadro 5.1: Intervalo keypoint 1. Img = Imagen. Kp Conf. = Keypoints confundidos. Lim. In. = límite inferior. Lim. Sup. = límite superior

Img.	Kp. Conf.	Lim. In.	Lim. Sup.	Rango
1	1	0.435	0.456	0.021
2	2	0.435	0.440	0.005
3	5	0.522	0.531	0.009
4	-	-	-	-
5	4	0.481	0.485	0.004
6	-	-	-	-
Media	3	0.468	0.487	0.019

Cuadro 5.2: Intervalo keypoint 2. Img = Imagen. Kp Conf. = Keypoints confundidos. Lim. In. = límite inferior. Lim. Sup. = límite superior

euclídea de los descriptores. Se pretende buscar intervalos de valores donde se asegure que no hay otro keypoint que pueda confundirse con el keypoint de referencia. Sin embargo, en algunos casos, se comprueba que antes de detectarse el mismo keypoint en la imagen de test, otros keypoints son detectados primeros. Por ello, a continuación se presentarán los intervalos donde no se confunde el emparejamiento y los keypoints mal detectados en algunos casos. El número de descriptores empleados han sido de 128. Nótese que hay imágenes donde no se detecta el keypoint señalado en la imagen de referencia. Los datos se muestran en los cuadros 5.1 - 5.9.

Además, como vemos en las columnas Lim. In. y Lim. Sup., los intervalos son muy estrechos. Esto disminuye la confianza de emparejar puntos de interés, con un margen amplio. Por otro lado, en la mayoría de los casos no existen puntos en las imágenes de test con descriptores más parecidos a los descriptores del keypoint seleccionado en la imagen de referencia. Lo ideal es que esto se cumpliera en todos los casos. Sin embargo, hay falsos positivos en las imágenes de test cuyos descriptores son más parecidos a los descriptores del keypoint seleccionado, que el keypoint correcto. Debido a ello se puede pensar

5. Estudio de las imágenes con la herramienta SURF

Img.	Kp. Conf.	Lim. In.	Lim. Sup.	Rango
1	5	0.462	0.464	0.002
2	21	0.599	0.604	0.005
3	0	0.396	0.460	0.064
4	-	-	-	-
5	13	0.570	0.575	0.005
6	4	0.439	0.448	0.009
Media	9	0.439	0.510	0.017

Cuadro 5.3: Intervalo keypoint 3. Img = Imagen. Kp Conf. = Keypoints confundidos. Lim. In. = límite inferior. Lim. Sup. = límite superior

Img.	Kp. Conf.	Lim. In.	Lim. Sup.	Rango
1	20	0.585	0.588	0.003
2	-	-	-	-
3	12	0.359	0.422	0.063
4	-	-	-	-
5	-	-	-	-
6	-	-	-	-
Media	16	0.472	0.505	0.033

Cuadro 5.4: Intervalo keypoint 4. Img = Imagen. Kp Conf. = Keypoints confundidos. Lim. In. = límite inferior. Lim. Sup. = límite superior

Img.	Kp. Conf.	Lim. In.	Lim. Sup.	Rango
1	0	0.421	0.478	0.057
2	1	0.470	0.512	0.042
3	0	0.457	0.472	0.015
4	-	-	-	-
5	0	0.363	0.512	0.149
6	-	-	-	-
Media	0	0.428	0.494	0.066

Cuadro 5.5: Intervalo keypoint 5. Img = Imagen. Kp Conf. = Keypoints confundidos. Lim. In. = límite inferior. Lim. Sup. = límite superior

Análisis Fotogramétrico de Imágenes Adquiridas por un UAV

Img.	Kp. Conf.	Lim. In.	Lim. Sup.	Rango
1	0	0.354	0.377	0.023
2	0	0.506	0.513	0.007
3	0	0.332	0.545	0.213
4	-	-	-	-
5	-	-	-	-
6	0	0.446	0.448	0.002
Media	0	0.409	0.471	0.062

Cuadro 5.6: Intervalo keypoint 6. Img = Imagen. Kp Conf. = Keypoints confundidos. Lim. In. = límite inferior. Lim. Sup. = límite superior

Img.	Kp. Conf.	Lim. In.	Lim. Sup.	Rango
1	0	0.431	0.524	0.093
2	0	0.393	0.503	0.110
3	0	0.515	0.547	0.032
4	0	0.494	0.590	0.096
5	-	-	-	-
6	-	-	-	-
Media	0	0.458	0.541	0.083

Cuadro 5.7: Intervalo keypoint 7. Img = Imagen. Kp Conf. = Keypoints confundidos. Lim. In. = límite inferior. Lim. Sup. = límite superior

Img.	Kp. Conf.	Lim. In.	Lim. Sup.	Rango
1	3	0.468	0.492	0.024
2	0	0.464	0.487	0.023
3	0	0.374	0.483	0.109
4	-	-	-	-
5	-	-	-	-
6	-	-	-	-
Media	1	0.435	0.487	0.052

Cuadro 5.8: Intervalo keypoint 8. Img = Imagen. Kp Conf. = Keypoints confundidos. Lim. In. = límite inferior. Lim. Sup. = límite superior

5. Estudio de las imágenes con la herramienta SURF

Img.	Kp. Conf.	Lim. In.	Lim. Sup.	Rango
1	0	0.349	0.544	0.195
2	0	0.314	0.563	0.249
3	0	0.332	0.494	0.162
4	-	-	-	-
5	0	0.256	0.556	0.300
6	-	-	-	-
Media	0	0.312	0.539	0.227

Cuadro 5.9: Intervalo keypoint 9. Img = Imagen. Kp Conf. = Keypoints confundidos. Lim. In. = límite inferior. Lim. Sup. = límite superior

en dos hipótesis: los keypoints no son tan singulares debido a la calidad de las imágenes o el método visual para señalar los keypoints no es correcto.

En el cuadro 5.10 se muestra a modo de resumen los valores medios obtenidos por los conjuntos de imágenes en cada punto. En términos medios se ha determinado un valor de confianza representativo: (0.423, 0.499). Además, en términos medios vemos que hay 4 keypoints cuyos descriptores se van a confundir con los descriptores de un keypoint seleccionado en la imagen de referencia. Además, en un margen de distancias euclídeas de descriptores de keypoints de 0.076 (diferencia del intervalo) podemos encontrar el keypoint esperado, confundido previamente con otros 4. También se ha añadido una columna final con una representación *s/t*. “s” indica el número de imágenes en las que ha sido detectado el punto, “t” indica el número de imágenes en las que no se confunde el keypoint con otros keypoints.

5.3.4 Otros resultados

Finalmente, en la primera subfigura perteneciente a la figura 5.9, se observa una imagen de test. En ella se selecciona sólo un keypoint, el cual corresponde con el keypoint seleccionado en la figura de referencia, la cual no se ha representado. El valor de la distancia euclídea es de 0.146. En la segunda subfigura aparece el segundo keypoint importante para un valor de distancia euclídea de 0.535. Finalmente, en la última subfigura vemos que para el valor de distancia euclídea 0.659 aparecen muchos keypoints detectados. De este ejemplo se puede deducir que a medida que el valor de distancia euclídea aumenta, de manera no gradual aumentan los keypoints: en el inicio aparecen pocos keypoints para un incremento rango de distancia euclídea, después de un valor, un incremento ocasiona una notable aparición de keypoints.

Análisis Fotogramétrico de Imágenes Adquiridas por un UAV

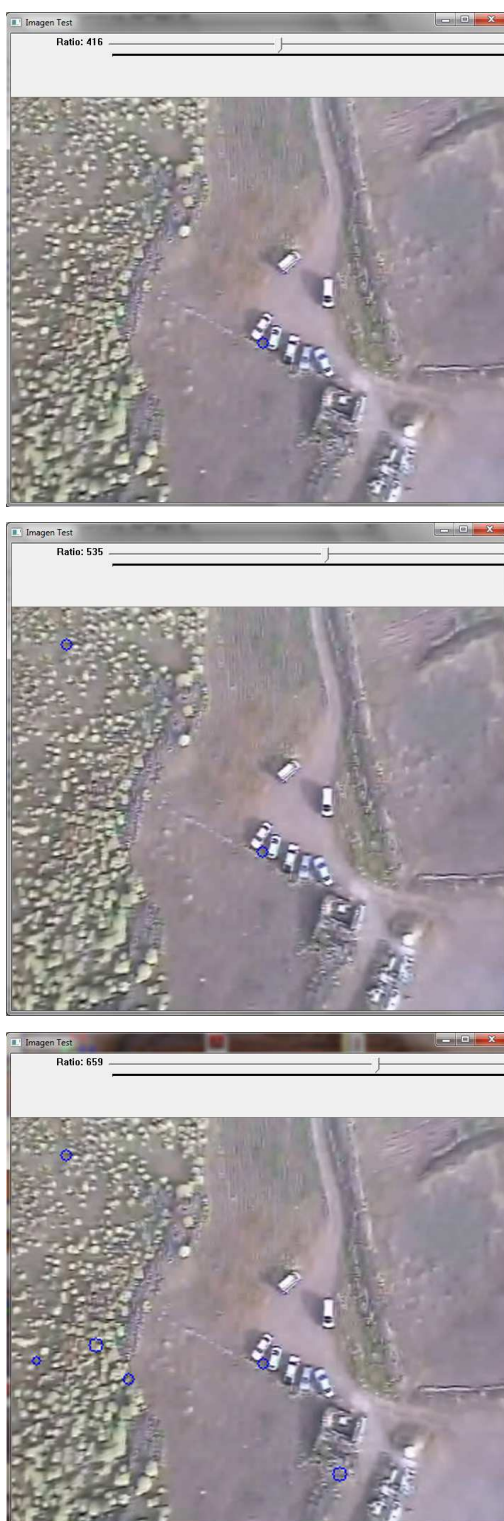


Figura 5.9: Aparición de keypoints según diferentes intervalos.

5. Estudio de las imágenes con la herramienta SURF

Punto	Kp. Conf.	Lim. In.	Lim. Sup.	Rango	Detecciones
1	5	0.387	0.458	0.071	4/2
2	3	0.468	0.487	0.019	4/0
3	9	0.439	0.510	0.071	5/1
4	16	0.472	0.505	0.033	2/0
5	0	0.428	0.494	0.066	4/3
6	0	0.409	0.471	0.062	4/4
7	0	0.458	0.541	0.083	4/4
8	1	0.435	0.487	0.052	3/2
9	0	0.312	0.539	0.227	4/4
Media				0.076	

Cuadro 5.10: Resumen de intervalos de confianza en cada punto. Kp Conf. = Keypoints confundidos. Lim. In. = límite inferior. Lim. Sup. = límite superior. Detecciones: el primer valor indica el número de imágenes en que fue detectado ese punto, el segundo valor corresponde al número de imágenes en que el keypoint en cuestión fue detectado el primero.

5.4 Emparejamientos automáticos

5.4.1 Metodología

En este apartado se realizan varias comparativas del emparejamiento de keypoints de las imágenes aéreas empleadas en el estudio. Para ello se realizará el matching de todas las imágenes test por separado con la imagen de referencia bajo unas condiciones. En el paso inicial se hace una tentativa de emparejamientos a partir del uso de SURF, implementado en OpenCV. La segunda condición es aplicar el filtrado de puntos singulares visto anteriormente y realizar el emparejamiento con los puntos singulares. Por último, la última condición es seleccionar puntos de interés, que visualmente parezcan más distinguibles y realizar matching con ellos. En todos los casos se hará un conteo del número de emparejamientos bien realizados, del número de emparejamientos mal realizados y el número total de emparejamientos. El estudio se realiza por subconjuntos de imágenes y también con el conjunto global de imágenes.

1. Este experimento es el más simple. Se procede a realizar matching de las imágenes de test, una por una, con la imagen de referencia. En un cuadro se mostrará el recuento final de los falsos positivos y de los verdaderos positivos. Se escribe el valor en porcentaje para poder comparar con los demás casos. Debido a que con frecuencia aparecen keypoints muy próximos entre sí, se han coloreado con el mismo color el keypoint A, por ejemplo, detectado en la imagen de referencia y el keypoint A, detectado

Análisis Fotogramétrico de Imágenes Adquiridas por un UAV

en la imagen de test i . Más adelante, se aclara la manera que se ha empleado para comprobar que los emparejamientos son correctos.

2. En el segundo experimento se reducen los puntos detectados automáticamente con la herramienta SURF en la imagen de referencia, según el procedimiento explicado en la sección 5.2. Se mostrará nuevamente en un cuadro los resultados obtenidos al aplicar la reducción de puntos. En la imagen de referencia se han seleccionado el 25% de los puntos detectados automáticamente por el programa SURF implementado en OpenCV.
3. En el último experimento, para comenzar se ha realizado a mano una selección de keypoints candidatos. Estos son los que visualmente se han determinado como los puntos más distinguibles en las imágenes. Estos 9 keypoints se comparan con todos los puntos que se detectan en las imágenes test y se procede a realizar los emparejamientos. Al igual que en los otros casos, se recuentan los emparejamientos correctos y los incorrectos. Nótese que el criterio de elección de keypoints se realiza a alto nivel. Es decir, se seleccionan puntos que perceptualmente son más singulares y que, tras el estudio, aparecen en muchas de las imágenes del conjunto. Añadir que se anota en los cuadros respectivos una columna que indica para la imagen i , el número de keypoints que se encontró en la imagen de referencia. El peor de los casos es que sea 0, pues quiere decir que no encontró ningún punto con el que se pudiera emparejar.

5.4.2 Primer subconjunto de imágenes

El presente subconjunto de imágenes corresponde a las imágenes tomadas a mayor altura.

5.4.2.1 Prueba 1

En el cuadro 5.11 vemos el recuento final de los falsos positivos y de los verdaderos positivos. Se escribe el valor en porcentaje para poder comparar con los demás casos. En la subfigura 5.10(a) se pueden apreciar los keypoints detectados en la imagen de referencia.

Según los resultados, podemos observar que más de la mitad de los emparejamientos son buenos, del orden del 67%. Además, existen unos 23 puntos, en términos medios, bien emparejados. Con lo cual, los emparejamientos buenos tienen un buen índice de valor. Sin embargo, lo deseable es que no aparezcan emparejamientos malos, pues añaden confusión en el matching final.

5. Estudio de las imágenes con la herramienta SURF

Img.	Kp	Emp	B	M	B (%)	M (%)
Ref.	298	-	-	-	-	-
1	282	34	23	11	68	32
2	287	37	25	12	68	32
3	198	34	22	12	65	35
Media	-	-	-	-	67	33

Cuadro 5.11: Emparejamientos automáticos con la herramienta SURF implementada en OpenCV. Img = Imagen. Kp = Keypoints detectados. Emp = Emparejamientos. B = Emparejamientos correctos. M = Emparejamientos Incorrectos.

Img.	Kp	Emp	B	M	B (%)	M (%)
Ref.	74	-	-	-	-	-
1	282	14	11	3	79	21
2	287	13	10	3	77	23
3	198	12	9	3	75	25
Media	-	-	-	-	77	23

Cuadro 5.12: Emparejamientos con la selección del 25% de los mejores puntos en la imagen de referencia con la herramienta SURF implementada en OpenCV para el subconjunto 1. Img = Imagen. Kp = Keypoints detectados. Emp = Emparejamientos. B = Emparejamientos correctos. M = Emparejamientos Incorrectos.

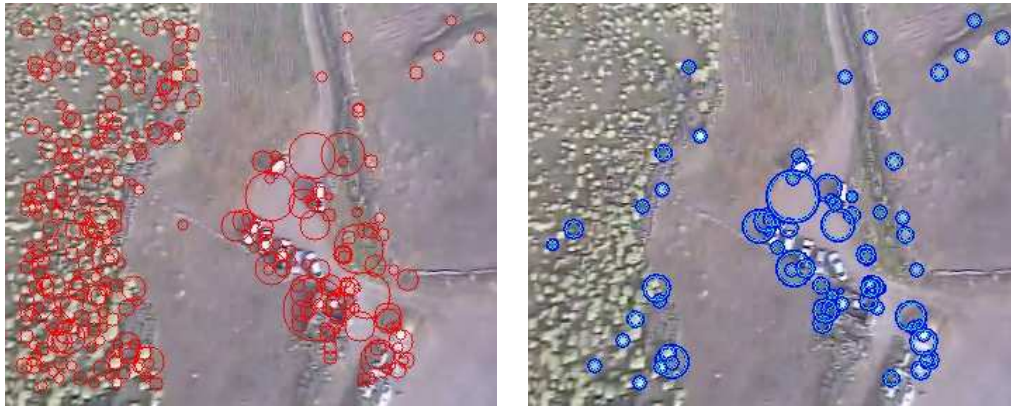
5.4.2.2 Prueba 2

La reducción de puntos en el segundo caso hace que los keypoints de la imagen de referencia sean 74. Se muestra en el cuadro 5.12 los resultados obtenidos al aplicar la reducción de puntos. En la figura 5.10 se puede visualizar los keypoints seleccionados en la imagen de referencia.

En este caso, al dejar los puntos más singulares según el criterio de eliminación de puntos no discriminantes, la tasa de acierto ha incrementado en un 10%. Este resultado es un indicio para pensar que el criterio desarrollado es acertado.

5.4.2.3 Prueba 3

Los keypoints elegidos manualmente corresponden en este caso a los 9 puntos de control comentados en la sección 5.3. Estos son los que visualmente se han determinado como los puntos más distinguibles en las imágenes. Nótese que el criterio de elección de keypoints se realiza a alto nivel. Es decir, se seleccionan



(a) Keypoints detectados con SURF (b) 25% de los keypoints más singulares detectados

Figura 5.10: En la subfigura (a) aparecen todos los keypoints detectados con la herramienta SURF implementada en OpenCV. En la subfigura (b) observamos los puntos singulares, con el fin de comparar la eliminación de puntos no discriminantes

puntos que a la vista aparentemente son más singulares. Los resultados se muestran en el cuadro 5.13 y en la figura 5.6 (sección 5.3) podemos ver los 9 puntos seleccionados manualmente. Vemos una columna adicional llamada *Encontrados* en el cuadro 5.13. Esta columna hace referencia a los keypoints que se detectaron en las imágenes de test y que además, visualmente, son detectados en la imagen de referencia. Por otro lado, la columna *Emparejamientos* hace referencia a aquellos keypoints que, tras ser encontrados en ambas imágenes, son emparejados en base al criterio de la distancia euclídea de sus descriptores. Es decir, dos puntos se emparejan en este caso cuando la distancia euclídea de sus descriptores es la mínima distancia euclídea en comparación con el resto de keypoints detectados.

Los resultados en este caso dan lugar a varios comentarios. Por un lado, al tener 9 puntos como máximo se realizarán 9 emparejamientos. En este subconjunto de imágenes test se encontraron los 9 puntos en cada una de las imágenes, salvo en un caso. Además, en términos medios, se han realizado 4/9 emparejamientos, es decir, un 44% de los emparejamientos posibles. Sin embargo, los puntos emparejados han sido los correctos. Todo ello apunta a un resultado esperanzador, pues si en lugar de 9 puntos señalados por un usuario, se señalaran más puntos, los resultados parecen indicar que se realizarían más emparejamientos exitosos. No obstante, las condiciones experimentales son escasas.

5. Estudio de las imágenes con la herramienta SURF

Img.	Kp	Encontrados	Emp	B	M	B (%)	M (%)
Ref.	9	-	-	-	-	-	-
1	282	9	3	3	0	100	0
2	287	8	3	3	0	100	0
3	198	9	6	6	0	100	0
Media	-	-	-	-	-	100	0

Cuadro 5.13: Emparejamientos con 9 keypoints seleccionados manualmente con la herramienta SURF implementada en OpenCV. Img = Imagen. Kp = Keypoints detectados. Emp = Emparejamientos. B = Emparejamientos correctos. M = Emparejamientos Incorrectos.

Img.	Kp	Emp	B	M	B (%)	M (%)
Ref.	207	-	-	-	-	-
1	246	16	9	7	56	44
2	366	15	8	7	53	47
Media	273	15.5	-	-	54.5	45.5

Cuadro 5.14: Emparejamientos automáticos con la herramienta SURF implementada en OpenCV. Img = Imagen. Kp = Keypoints detectados. Emp = Emparejamientos. B = Emparejamientos correctos. M = Emparejamientos Incorrectos.

5.4.3 Segundo subconjunto de imágenes

En este caso el subconjunto de imágenes corresponde a las imágenes tomadas a menor altura, a una altura más cercana al suelo.

5.4.3.1 Prueba 1

En el cuadro 5.14 vemos el resultado de la fase de emparejamiento automático. Adicionalmente, en la figura 5.11 podemos observar los keypoints detectados automáticamente por el algoritmo SURF y un emparejamiento automático con una imagen de test.

Los resultados en esta prueba indican que los keypoints se emparejan bien de la misma manera que se emparejan mal. Además, el número de emparejamientos, en comparación con los keypoints detectados es muy bajo, del orden de $\frac{15.5}{273} \approx 0.05$, es decir, del 5%. Debido a ello, las imágenes a menor escala o con el avión más cerca del suelo, tras aplicarle la calibración y el efecto de entrelazado, obtienen un aliasing perceptual tan notable y una baja calidad, que es difícil realizar en ellos un emparejamiento automático fiable.



(a) Keypoints detectados en la imagen de referencia

(b) Matching

Figura 5.11: Emparejamientos con la imagen de referencia y una imagen de test aplicando la herramienta SURF implementada en OpenCV en el subconjunto 2.

5. Estudio de las imágenes con la herramienta SURF

Img.	Kp	Emp	B	M	B (%)	M (%)
Ref.	51	-	-	-	-	-
1	246	8	3	5	37	63
2	366	7	3	4	43	57
Media	-	-	-	-	40	60

Cuadro 5.15: Emparejamientos con la selección del 25% de los mejores puntos en la imagen de referencia con la herramienta SURF implementada en OpenCV para el subconjunto 2. Img = Imagen. Kp = Keypoints detectados. Emp = Emparejamientos. B = Emparejamientos correctos. M = Emparejamientos Incorrectos.

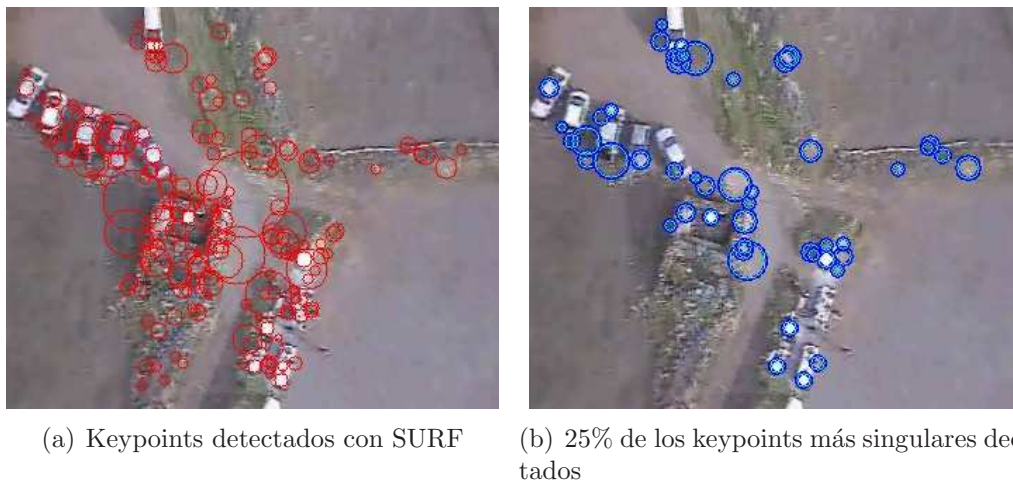


Figura 5.12: En la figura (a) aparecen todos los keypoints detectados con la herramienta SURF implementada en OpenCV. En la figura (b) observamos los puntos singulares, con el fin de comparar la eliminación de puntos no discriminantes

5.4.3.2 Prueba 2

En el segundo caso, se ha reducido los puntos de control quedándose la imagen de referencia con 51 keypoints. El cuadro 5.15 recoge los resultados obtenidos al aplicar la reducción de puntos. En la figura 5.12 se puede comparar la reducción de puntos en la imagen de referencia.

Para este caso, al aplicar la reducción de puntos, los resultados han empeorado. El número de emparejamientos ha caído a la mitad con respecto al número de emparejamientos en el caso contrario. Debido a ello, los resultados referentes a los emparejamientos buenos y malos vuelven a ser pobres y peores que en el caso anterior. Aquí los emparejamientos buenos han decrementado en un 10%.



Figura 5.13: 12 keypoints señalados manualmente en la imagen de referencia del subconjunto 2 de imágenes

5.4.3.3 Prueba 3

En el tercer caso, se han señalado manualmente 12 puntos en base a la observación de cuáles podrían ser los puntos más deterministas en este caso. Se han observado los keypoints que se emparejaban en los otros apartados y se han señalado de manera manual. Se muestra en la figura 5.13 los keypoints seleccionados manualmente en la imagen de referencia. El cuadro 5.16 muestra los resultados de este apartado. En este caso, 6 keypoints seleccionados se emparejan con la imagen 1 y los otros 6 keypoints, con la imagen 2.

Los resultados en este caso son ideales, sin embargo existen tres inconvenientes. El primero es que los puntos señalados manualmente han sido señalados tras observar los puntos que se emparejaban correctamente. Esto dista mucho de la búsqueda correcta de puntos automáticos. Es por ello, que aparecen 6 puntos emparejados en la imagen 1 y 6 puntos emparejados en la imagen 2. El segundo inconveniente es que los puntos correctamente emparejados son muy pocos (6 en cada caso). Para un correcto matching es necesario alrededor de 9 puntos. El tercer inconveniente, el cual se muestra en todo el trabajo, es que el conjunto de imágenes es escaso.

5. Estudio de las imágenes con la herramienta SURF

Img.	Kp	Encontrados	Emp	B	M	B (%)	M (%)
Ref.	12	-	-	-	-	-	-
1	246	6	6	6	0	100	0
2	366	6	6	6	0	100	0
Media	-	-	-	-	-	100	0

Cuadro 5.16: Emparejamientos con 12 keypoints seleccionados manualmente con la herramienta SURF implementada en OpenCV. Img = Imagen. Kp = Keypoints detectados. Emp = Emparejamientos. B = Emparejamientos correctos. M = Emparejamientos Incorrectos.

Img.	Kp	Emp	B	M	B (%)	M (%)
Ref.	298	-	-	-	-	-
1	282	34	23	11	68	32
2	287	37	25	12	68	32
3	198	34	22	12	65	35
4	246	21	9	12	43	57
5	207	37	24	13	65	35
6	366	21	12	9	58	42
Media	269	31	-	-	61.2	38.8

Cuadro 5.17: Emparejamientos automáticos con la herramienta SURF implementada en OpenCV (III). Img = Imagen. Kp = Keypoints detectados. Emp = Emparejamientos. B = Emparejamientos correctos. M = Emparejamientos Incorrectos.

5.4.4 Combinación de los subconjuntos de imágenes

5.4.4.1 Prueba 1

En este apartado se combinan los dos subconjuntos de imágenes. La imagen de referencia escogida ha sido la misma que se escogió en el subconjunto 1 de imágenes. Ver en la figura 5.2 al inicio del presente capítulo, la imagen con un recuadro rojo. Esta combinación la haremos para completar el estudio de esta sección con el conjunto completo de imágenes.

En el cuadro 5.17 vemos el recuento final de los falsos positivos y de los verdaderos positivos. Estos son emparejados automáticamente, con el programa SURF.

Un primer comentario tras los resultados observados en la tabla es el bajo número de emparejamientos realizados en comparación con el número de keypoints detectados en las imágenes. El orden es del $\frac{31}{269} \approx 0.12$, es decir, del orden del 12%. Las imágenes 4 y 6 son las que peores resultados muestran, am-

Análisis Fotogramétrico de Imágenes Adquiridas por un UAV

Img.	Kp	Emp	B	M	B (%)	M (%)
Ref.	74	-	-	-	-	-
1	282	14	11	3	79	21
2	287	13	10	3	77	23
3	198	12	9	3	75	25
4	246	7	3	4	43	57
5	207	18	10	8	56	44
6	366	11	4	7	36	64
Media	-	-	-	-	61	39

Cuadro 5.18: Emparejamientos con la selección del 25% de los mejores puntos en la imagen de referencia con la herramienta SURF implementada en OpenCV (III). Img = Imagen. Kp = Keypoints detectados. Emp = Emparejamientos. B = Emparejamientos correctos. M = Emparejamientos Incorrectos.

bas pertenecen al subconjunto 2. En cambio, la imagen de referencia pertenece al subconjunto 1. Sin embargo en términos medios, se consigue que el número de emparejamientos correctos sea mayor al número de los emparejamientos incorrectos. Estos resultados muestran que la tarea se debe centrar principalmente en dos partes. Una parte es la de eliminar los falsos positivos. Otra parte es aumentar el número de emparejamientos en relación a los keypoints detectados.

En la figura 5.14 se muestran la imagen de referencia y una imagen test con diferentes círculos coloreados con diferentes colores. Un círculo con el mismo color significa un keypoint encontrado y emparejado en ambas imágenes. La táctica de los colores ha servido para facilitar el conteo de los emparejamientos correctos e incorrectos.

5.4.4.2 Prueba 2

En el cuadro 5.18 se muestran los resultados cuando se reduce el 25% de los puntos detectados automáticamente. Para este caso, se detectaron 74 keypoints. En la figura 5.6 se puede visualizar los keypoints seleccionados en la imagen de referencia.

Realizando una comparación de estos resultados con respecto a los otros, en términos medio podemos decir que no existen diferencias entre los emparejamientos buenos y los malos. Sin embargo, se aprecia en las columnas B (%) y M (%) que existen dos tipos de imágenes. Uno comprende las imágenes 1, 2 y 3 y el otro las imágenes 4, 5 y 6, las cuales, efectivamente corresponden a los dos subconjuntos. Si se compara el comportamiento del subconjunto de imágenes en este cuadro y en el cuadro 5.15 (prueba 2, segundo subcon-

5. Estudio de las imágenes con la herramienta SURF



(a) Imagen de referencia

(b) Imagen de test



(c) Matching

Figura 5.14: Emparejamientos con la imagen de referencia y una imagen de test aplicando la herramienta SURF implementada en OpenCV.

Análisis Fotogramétrico de Imágenes Adquiridas por un UAV

Img.	Kp	Encontrados	Emp	B	M	B (%)	M (%)
Ref.	9	-	-	-	-	-	-
1	282	9	3	3	0	100	0
2	287	8	3	3	0	100	0
3	198	9	6	6	0	100	0
4	246	2	1	1	0	100	0
5	207	4	3	3	0	100	0
6	366	2	0	0	0	0	0
Media	-	-	-	-	-	83.3	0

Cuadro 5.19: Emparejamientos con 9 keypoints seleccionados manualmente con la herramienta SURF implementada en OpenCV (III). Img = Imagen. Kp = Keypoints detectados. Emp = Emparejamientos. B = Emparejamientos correctos. M = Emparejamientos Incorrectos.

junto), se comprueba que existen más emparejamientos ahora y, además, más keypoints correctamente emparejados. Esto apunta a que los malos resultados mostrados en el cuadro 5.15 se debe a los pocos emparejamientos que se realizaron. Es decir, se deduce que a mayor número de emparejamientos, la probabilidad de que hayan más emparejamientos buenos que malos es elevada. Por otra parte, en este caso, la imagen de referencia es una imagen tomada a más altura que las imágenes que pertenecen al segundo conjunto.

5.4.4.3 Prueba 3

Los puntos señalados manualmente en el último caso, son los mismo puntos que se señalaron en la prueba 3 con el primer subconjunto de imágenes. Los 9 puntos señalados se comparan con todos los puntos que se detectan en la secuencia de imágenes de test. Se incide nuevamente en que los puntos seleccionados son puntos que visualmente indican ser singulares. Además, tras observar el comportamiento del emparejamiento automático, aparecen puntos que suelen emparejarse a menudo. Éste ha sido un motivo para seleccionar los 9 puntos de control. Los resultados se muestran en el cuadro 5.19. En la figura 5.6 (sección 5.3) podemos ver los 9 puntos seleccionados manualmente, los cuales se comparan con todos los detectados por la función SURF en las imágenes test.

La columna “Encontrados” nos indica el número de keypoints encontrados en las imágenes de test, de los 9 posibles puntos que se puede encontrar. Lo ideal es que en todas las imágenes de test se encuentren 9 puntos y además que sean 9 emparejamientos correctos (columna “Emp”). Para comparar el número de emparejamientos, con relación a los emparejamientos posibles, se

5. Estudio de las imágenes con la herramienta SURF

	Emp. Automát.		25% Mej. Kp.		Selec. Manual	
	B (%)	M (%)	B (%)	M (%)	B (%)	M (%)
1	68	32	79	21	100	0
2	68	32	77	23	100	0
3	65	35	75	25	100	0
4	43	57	43	57	100	0
5	65	35	56	44	100	0
6	58	42	36	64	0	0
Media	61.2	38.8	61	39	83.3	0

Cuadro 5.20: Resumen de los emparejamientos en 3 experimentos con la herramienta SURF implementada en OpenCV. Emp. Automát. = Emparejamiento Automático. 25% Mej. Kp. = 25% de los mejores keypoints. Selec. Manual = Selección de puntos manual. B = Emparejamientos correctos. M = Emparejamientos Incorrectos.

deduce que hay dos tipos de comportamientos, debido a los dos tipos de escala que se han tratado. Uno para las tres primeras imágenes de test, donde se emparejan la mitad de posibles puntos. El otro comportamiento es en las últimas tres imágenes de test, donde también se emparejan la mitad de puntos. La diferencia está en que en el primer subconjunto se pueden realizar más emparejamientos que en el segundo. Lo realmente satisfactorio de este caso es que los emparejamientos que se realizan son siempre correctos.

Por otro lado, en el cuadro 5.20 se muestran los porcentajes de los puntos bien emparejados y de los puntos mal emparejados en los tres casos: emparejamiento automático de todos los puntos; selección de keypoints en la imagen de referencia, por un lado de los 25% mejores puntos y por otro lado con 9 puntos seleccionados manualmente.

Este último cuadro muestra unos resultados que no pueden ser concluyentes por sí mismos, es decir, necesitan de la explicación del comportamiento en el resto de casos para poder ser entendido. Sin embargo, se concluye con que el criterio de eliminación de puntos singulares es aceptable pero no determinante. Es decir, la idea de eliminación es acertada a pesar de que existe una mezcla de efectos (diferentes escalas, ángulos en los que son tomadas las imágenes, momentos diferentes en los que son tomadas las imágenes, diferentes iluminaciones en las imágenes, etc. . .). En todos los casos, lo mejor que ha resultado es realizar una selección manual de puntos, pues los emparejamientos son siempre correctos. Intencionadamente, se ha añadido al conjunto de imágenes de test dos imágenes muy parecidas, estas son la imagen 1 y 2, corresponden con las dos imágenes localizadas en la parte central superior de la figura 5.1. Por ello, el comportamiento es prácticamente igual en todos los cuadros de las

Análisis Fotogramétrico de Imágenes Adquiridas por un UAV

secciones 5.4.2 y 5.4.4. Esto indica una conclusión interesante. Si queremos emparejar imágenes tomadas por el avión no tripulado, el cual planea lentamente, con poco tiempo de toma entre una y otra, es probable que el índice de puntos emparejados correctamente sea elevado. Por ello, realizar una tarea de stitching, por ejemplo, podría ser interesante plantearla teniendo en cuenta estas condiciones.

Para terminar, en el rango de cuadros 5.14 - 5.16 referentes al estudio con el segundo subconjunto de imágenes, y en los cuadros 5.17 - 5.19 referentes al estudio con la combinación de los subconjuntos de imágenes, se observa un comportamiento curioso. Cuando se trabaja con el segundo subconjunto de imágenes, la imagen de referencia es tomada en la misma escala en la que se trabaja. Cuando se trabaja con todas las imágenes, las filas correspondientes a las imágenes 4, 5 y 6 son imágenes del subconjunto 2, en todos los cuadros. En este caso la imagen de referencia es una imagen tomada a mayor altura. Se observa que los resultados de las imágenes del segundo subconjunto son un poco mejores cuando la imagen de referencia es una imagen tomada a más altura.

5. Estudio de las imágenes con la herramienta SURF

Capítulo 6

Conclusiones y líneas futuras

Resumen: En el presente trabajo se ha realizado un estudio de emparejamiento de imágenes aéreas adquiridas por el avión aéreo no tripulado, según sus descriptores visuales, en dos etapas: Una etapa de preprocesado de las imágenes y otra etapa de filtrado de descriptores no determinantes. Se procede a apuntar las conclusiones relevantes de este trabajo. Así mismo se discuten posibles mejoras y líneas futuras.

6.1 Conclusiones

1. Se ha hecho un estudio y una búsqueda de los diferentes métodos que se emplean en la detección de puntos de interés automáticos en imágenes aéreas en el capítulo 2. Adicionalmente, se muestran en el mismo capítulo los diferentes métodos de detección visual más usados, divididos en: métodos monoescala y métodos multiescala. Se han encontrado bastantes trabajos dónde se explica el montaje y las especificaciones del vehículo aéreo, así como su instrumental y sensores. Además muchos autores coinciden en emplear la herramienta SURF (Bay *et al.*, 2008), razón por la cual se la ha elegido también para este estudio.
2. En el capítulo 3 se muestra el diseño electrónico del instrumental que se aloja en el vehículo aéreo no tripulado. La parte del circuito que se refiere al conexionado de la cámara con el OSD y el emisor, así como el receptor con el digitalizador conectado al PC, ha sido montado en el laboratorio para obtener los parámetros de calibración de la cámara. La figura 3.2 de dicho capítulo, muestra el circuito montado en el laboratorio. También se ha procedido a diseñar un algoritmo que, a partir del software comercial

6. Conclusiones y líneas futuras

del digitalizador empleado, sea capaz de captar la secuencia de vídeo de la cámara de abordo del avión, y la guarde en formato de datos IPL de OpenCV.

3. La secuencia de imágenes no está corregida de los efectos de cabeceo y alabeo. Se emplean fotogramas que detecta la cámara, la cual está instalada solidaria con el UAV. Nótese que el efecto de guiñada no precisa de corrección, pues el efecto es una imagen girada con respecto a un eje perpendicular a la imagen.
4. La calibración de la cámara se ha llevado a cabo mediante un método de calibración clásico, el cual emplea un patrón. Debido a la altura a la que han sido tomadas las imágenes, no se produce mucha distorsión en ellas. Todo apuntó a que el método más indicado a utilizar fuese el método de calibración de Zhang. Los resultados de la corrección de las imágenes, así como de la calibración son mostrados en el capítulo 4.
5. Se ha notado que la secuencia de fotogramas sufrían un efecto conocido como entrelazado. Debido a ello, se ha diseñado un filtro para desentrelazar una imagen. Se ha realizado, por tanto, este filtrado consiguiendo así separar una imagen en dos. Esto filtrado, al eliminar filas de píxeles y ponderar columnas de píxeles, la resolución de una imagen disminuye en $1/4$.
6. En el capítulo 5 se muestra el conjunto de fotogramas previamente calibrados y desentrelazados con el que se procede a realizar el estudio de sus descriptores. Se ha visto la necesidad de separar el conjunto de imágenes en dos según la escala. El hecho se justifica con ampliar el estudio de los descriptores y optimizar la elección de estos.
7. En la tarea de elegir los mejores keypoints, se ha realizado un algoritmo que elimina aquellos keypoints cuyos descriptores tienen algún parecido con otro. Por ejemplo, muchos puntos de interés automáticamente seleccionados coinciden con zonas de arbustos, tales como tabaibas. El intento de eliminación de keypoints confirmó esta hipótesis, pues muchos keypoints falsos o confusos fueron eliminados. Sin embargo, este método puede ser modificado para robustecerlo, como se planteará en las líneas futuras. A pesar de ello, podemos decir que este criterio de puntos no discriminantes es bueno cuando hay aliasing perceptual.
8. Se han emparejado puntos de control entre imágenes, seleccionando manualmente los keypoints. Se ha calculado la distancia euclídea de los descriptores. En términos medios, los puntos que son detectados como

puntos iguales en dos imágenes distintas, se emparejan cuando la distancia euclídea de sus descriptores es 0.4.

9. Por otro lado, se ha realizado un estudio de emparejamientos automáticos de keypoints. Dicho estudio se ha dividido en dos partes: Emparejamiento automático simple, emparejamiento automático controlado. A su vez, la segunda parte se subdivide en eliminación de puntos no determinantes y en selección manual de puntos candidatos. Ambos controles se realizan sobre la imagen de referencia. Estos experimentos se han realizado tres veces: dos experimentos con los dos subconjuntos de imágenes a dos escalas distintas, y otro experimento con todas las imágenes juntas. Como cabe esperar, los mejores resultados se logran cuando manualmente se seleccionan puntos de interés, en base a la percepción visual. Sin embargo, la eliminación de puntos no determinantes es clave en este apartado. Los resultados muestran que el criterio de eliminación de puntos no determinantes, en algunos casos incrementa mejoras del orden del 10% en las tasas de aciertos, en otros casos las mejoras decrecientan en un 10% también. Por lo tanto, el criterio empleado no introduce robustez ni fiabilidad, pero si una línea a explorar.

6.2 Líneas futuras

1. Actualmente, se carecen de datos tales como la altura de vuelo o las coordenadas de orientación del avión en el espacio asociados a las imágenes para realizar matching sobre las mismas. Se pretende dotar de sensores que den más información del vuelo. Actualmente el glider posee un GPS y un horizonte artificial, sus datos no han sido tomados en cuenta en este trabajo. Por ejemplo, añadiendo un sistema inercial se pueden captar datos de la aceleración lineal y de las velocidades de giro. Se estima que estos datos en conjunto enriquecerían la posición del fotograma en el mosaico final, el cual debe ser georeferenciado.
2. Una mejora en la cámara de vídeo podría proporcionar información visual de mayor calidad. Se pretende realizar emparejamientos de imágenes con un conjunto de imágenes de mejor calidad. De esta manera se podría conseguir un mosaico de calidad, con respecto a la cartografía de GRAFCAN. Posteriormente, la tarea sería la de enganchar ese mosaico en una ortoimagen de GRAFCAN.
3. Se ha desarrollado en este trabajo un algoritmo de eliminación de puntos no discriminantes. A tenor de las conclusiones expuestas, la idea y el desarrollo auguran una posible ayuda al emparejamiento automático de

6. Conclusiones y líneas futuras

puntos en las imágenes a partir de los descriptores de las mismas. Por ello se pretende mejorar este criterio apuntando estas ideas futuras:

- Actualmente cuando se describen dos puntos de una manera parecida se eliminan ambos. Se sospecha que este criterio elimine puntos buenos. Es decir, en este trabajo se eliminan puntos que podrían ser estables. Por ello se piensa en buscar algún criterio para eliminar uno de los dos, no los dos.
- Un posible criterio sería la distancia, en píxeles, a la que se encuentran los puntos. Por ejemplo, si los puntos están muy cerca es posible que se trate de una zona confusa y en tal caso sería conveniente eliminar los dos. Por el contrario, si los puntos se encuentran a una distancia considerable, es posible que ambos puedan ser buenos para futuros emparejamientos.
- Otra idea que generaría robustez sería la orientación. Si se detectan varios puntos en una imagen, es deseable encontrar en la otra imagen esos puntos con la misma orientación.
- En algunos casos, es posible encontrar un keypoint y que sus píxeles adyacentes también sean puntos de control. Como si de keypoints superpuestos se tratara. En tal caso, sería necesario imponer un criterio de discriminación empleando información que da el algoritmo de SURF. Por ejemplo, la orientación; el ángulo, el cual es recomendado por OpenCV para seleccionar puntos clave más importantes; el espacio - escala en el cual fue encontrado el keypoint o incluso la octava a la que fue encontrado.

4. Las condiciones en las que se han hecho este estudio han sido excesivamente adversas debido a los siguientes motivos:

- Falta de datos.
- Calidad muy mala en las imágenes.
- Cámara de bajo coste.
- etc ...

El nivel de exigencia de este trabajo no es real, en comparación con muchos de los trabajos mencionados en el capítulo 2. Debido a ello, los futuros esfuerzos apuntan a lograr una colección de imágenes con una cámara giroestabilizada para reducir los efectos de cabeceo y alabeo. Con dicha colección, la cual se desea que sea tomada a la misma distancia del suelo, se pretende resolver el problema de la creación de un mosaico.

Bibliografía

- Método de calibración de cámaras implementado en las librerías de opencv.
URL <http://opencv.willowgarage.com>.
- John D. Anderson. *Fundamentals of Aerodynamics*. McGraw-Hill, Boston, third edition edition, 2001.
- M. Ballesta, A. Gil, O. Reinoso, and D. Ubeda. Analisis de detectores y descriptores de características visuales en slam en entornos interiores y exteriores. *In To appear in RIAII (Revista Iberoamericana de Automatica e Informatica Industrial)*, 2010.
- Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool. Speeded-up robust features (surf). *Comput. Vis. Image Underst.*, 110:346–359, June 2008. ISSN 1077-3142. doi: 10.1016/j.cviu.2007.09.014. URL <http://portal.acm.org/citation.cfm?id=1370312.1370556>.
- Federico Consentino. Development of an integral system for aero photogrammetric flight planning, data acquisition and imagery ortho-mosaicking. In *AeroDreams UAV*, Buenos Aires, Argentina, Julio 2008.
- J. L. Crowley and A. C. Sanderson. Multiple resolution representation and probabilistic matching of 2-d gray-scale shape. *IEEE Trans. Pattern Anal. Mach. Intell.*, 9:113–121, January 1987. ISSN 0162-8828. doi: 10.1109/TPAMI.1987.4767876. URL <http://dl.acm.org/citation.cfm?id=28748.28757>.
- B. Etkin and L.D. Reid. *Dynamics of flight: Stability and control*. John Wiley and Sons, New York, third edition edition, 1996.
- Martin A. Fischler and Robert C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, 24:381–395, June 1981. ISSN 0001-0782. doi: <http://doi.acm.org/10.1145/358669.358692>. URL <http://doi.acm.org/10.1145/358669.358692>.

BIBLIOGRAFÍA

- Martin A. Fischler and Robert C. Bolles. Readings in computer vision: issues, problems, principles, and paradigms. chapter Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, pages 726–740. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1987. ISBN 0-934613-33-8. URL <http://dl.acm.org/citation.cfm?id=33517.33575>.
- Arturo Gil, Oscar Martinez Mozos, Monica Ballesta, and Oscar Reinoso. A comparative evaluation of interest point detectors and local descriptors for visual SLAM. *Machine Vision and Applications (MVA)*, 21(6):905–920, October 2010. ISSN 0932-8092.
- C. Harris and M. Stephens. A combined corner and edge detector. In *Proceedings of the 4th Alvey Vision Conference*, page 147–151, 1988.
- R. Hartley and A. Zisserman. Multiple view geometry in computer vision. Cambridge University Press, 2000.
- E. Hecht and A. Zajac. *Optics*. Addison Wensley New York, 2nd edition edition, 1979.
- José Isern-González. *Estudio experimental de métodos de calibración y auto-calibración de cámaras*. PhD thesis, Universidad de Las Palmas de Gran Canaria, 2003.
- G. Ladd, A. Magchaudhuri, T.J. Earl, G Bland, and M. Mitra. Rectification, georeferencing, and mosaicking of images acquired with remotely operated aerial platforms. In *Proceedings of Annual Conference of American Society for Photogrammetry and Remote Sensing*, Reno, Nevada, 2006.
- David G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, Nov 2004.
- F. López. *Formación de mosaicos ortorectificados utilizando imágenes aéreas digitales*. PhD thesis, Tesis de Licenciatura en Geografía. Facultad de Filosofía y Letras, UNAM., 2006. pp 1-103.
- Luis Merino Cabañas and Anibal Ollero Baturone. Procesamiento de imágenes aéreas para detección y monitorización empleando vehículos aéreos. Technical Report 23, XXIII Jornadas de Automatica. Jornadas de Automatica, La Laguna, Tenerife, 2002.
- Krystian Mikolajczyk and Cordelia Schmid. Indexing based on scale invariant interest points. In *In Proc. ICCV*, pages 525–531, 2001. URL <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.25.984>.

Análisis Fotogramétrico de Imágenes Adquiridas por un UAV

- Hans Moravec. Obstacle avoidance and navigation in the real world by a seeing robot rover. In *tech. report CMU-RI-TR-80-03, Robotics Institute, Carnegie Mellon University and doctoral dissertation, Stanford University*, number CMU-RI-TR-80-03. September 1980.
- Peter Pesti, Jeremy Elson, Jon Howell, Drew Steedly, and Matt Uyttendaele. Low-cost orthographic imagery. In *Proceedings of the 16th ACM SIGSPATIAL international conference on Advances in geographic information systems, GIS '08*, pages 24:1–24:8, New York, NY, USA, 2008. ACM. ISBN 978-1-60558-323-5. doi: <http://doi.acm.org/10.1145/1463434.1463465>. URL <http://doi.acm.org/10.1145/1463434.1463465>.
- Voicu Popescu, Paul Rosen, Laura Arns, Xavier Tricoche, Chris Wyman, and Christoph Hoffmann. The general pinhole camera: Effective and efficient non-uniform sampling for visualization.
- Jorge Prado-Molina, Armando Peralta Higuera, and José López García. Utilización de imágenes aéreas digitales de pequeño formato para la evaluación de desastres. *Mapping*, pages 70–73, 2009. ISSN 11319100. URL http://www.mappinginteractivo.com/plantilla-ante.asp?id_articulo=1550.
- Markus Quaritsch, Emil Stojanovski, Christian Bettstetter, Gerhard Friedrich, Hermann Hellwagner, Bernhard Rinner, Michael Hofbaur, and Mubarak Shah. Collaborative microdrones: applications and research challenges. In *Proceedings of the 2nd International Conference on Autonomic Computing and Communication Systems, Autonomics '08*, pages 38:1–38:7, ICST, Brussels, Belgium, Belgium, 2008. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering). ISBN 978-963-9799-34-9. URL <http://portal.acm.org/citation.cfm?id=1487652.1487690>.
- J. Shi and C. Tomasi. Good features to track. In *Novena conferencia del IEEE sobre Computer Vision and Pattern Recognition*, 1994.
- Patrice Y. Simard, Léon Bottou, Patrick Haffner, and Yann LeCun. Boxlets: a fast convolution algorithm for signal processing and neural networks. In *Proceedings of the 1998 conference on Advances in neural information processing systems II*, pages 571–577, Cambridge, MA, USA, 1999. MIT Press. ISBN 0-262-11245-0. URL <http://portal.acm.org/citation.cfm?id=340738>.
- Girish Singh and Zia ur Rahman. Hazard detection on runways using image processing techniques. In *Proc. SPIE 6957, 69570D*, 2008. doi: [doi:10.1117/12.782569](https://doi.org/10.1117/12.782569).

BIBLIOGRAFÍA

Stephen M. Smith and J. Michael Brady. Susan — a new approach to low level image processing. *Int. J. Comput. Vision*, 23:45–78, May 1997. ISSN 0920-5691. doi: 10.1023/A:1007963824710. URL <http://dl.acm.org/citation.cfm?id=258049.258056>.

Webster's. *New World College Dictionary*. Macmillan, fourth edition edition, 1999.

S. Yahyanejad, D. Wischounig Strucl, M. Quaritsch, and B. Rinner. Incremental mosaicking of images from autonomous, small-scale uavs. In *7th IEEE International Conference on Advanced Video and Signal-Based Surveillance (AVSS10)*, pages 329–336, 2010.

Fenghui Yao, Guifeng Shao, Ali Sekmen, and Mohan Malkani. Real-time multiple moving targets detection from airborne ir imagery by dynamic gabor filter and dynamic gaussian detector. *EURASIP J. Image and Video Processing*, 2010, 2010.

Z. Zhang. A flexible new technique for camera calibration. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(11):1330 – 1334, nov 2000. ISSN 0162-8828. doi: 10.1109/34.888718.

Zhengyou Zhang. Flexible camera calibration by viewing a plane from unknown orientations. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, volume 1, pages 666 –673 vol.1, 1999. doi: 10.1109/ICCV.1999.791289.

Glosario

A

aerial surveillance Aerial surveillance o vigilancia aérea es la reunificación de la supervisión de imágenes, por lo general de las imágenes visuales o de vídeo, tomadas desde un vehículo en el aire, como un vehículo aéreo no tripulado, helicóptero o avión espía.

B

balanceo (roll) rotación de la aeronave sobre el centro de gravedad en dirección vertical (Etkin and Reid, 1996), (Anderson, 2001), p. 6.

C

cabeceo (pitch) movimiento de nariz y talón de una aeronave sobre el centro de gravedad en la dirección vertical al cuerpo (Etkin and Reid, 1996), (Anderson, 2001), p. 6.

cenital Plano cenital es aquel en el que la óptica de la cámara se encuentra perpendicular respecto del suelo y la imagen obtenida ofrece un campo de visión orientado de arriba a abajo.

cockpit Monitor virtual de vuelo que reproduce con el máximo realismo la cabina de vuelo. En ella se visualiza la imagen junto con datos tales como niveles de batería, posición y orientación del glider UAV, etc, p. 28.

D

diferencia de Gaussianas La diferencia de Gaussianas se trata de un algoritmo de mejora de una imagen en escala de grises. Consiste en la

GLOSARIO

sustracción de una versión borrosa de la imagen a partir de la original en escala de grises. Las imágenes borrosas son obtenidas mediante la convolución del original en escala de grises con los núcleos de Gauss usando diferentes desviaciones típicas, p. 14.

F

fotogrametría El arte o proceso de reconocimiento o de medición, como en la confección de mapas, tomando fotografías, especialmente fotografías aéreas (Webster's, 1999).

FPV First-person view (FPV), en castellano se emplea el término vuelo en primera persona, es un método usado para controlar un vehículo de modelismo de radio control a partir de la vista que tendría un piloto. Este término es usado para aplicaciones de pilotaje de vehículos aéreos no tripulados (VANT/UAV) o para aircraft de radiocontrol. El vehículo es pilotado remotamente a través de una cámara de abordo, la cual es alimentada de manera inalámbrica a unas gafas de realidad virtual o a un monitor. Con ello el piloto consigue la misma perspectiva visual que un piloto sentado en el aeromodelo, p. 28.

G

georefenciar Es alinear algo a un sistema de coordenadas terrestre centrado en la tierra. Con imágenes aéreas, georefenciar es el alineamiento de una imagen rectificadas según un sistema de coordenadas terrestre, por ejemplo, empleando la latitud y la longitud, p. 9.

guiñada (yaw) rotación del cuerpo de la aeronave en el plano horizontal (Etkin and Reid, 1996), (Anderson, 2001), p. 6.

I

imagen ortográfica Se trata de una imagen pre-existente que ya ha sido rectificadas y georeferenciadas con una conocida precisión. Suelen estar disponible en fuentes estatales y federales, p. 9.

IMU *Unidad de medida Inercial*. Instrumento que mide los incrementos angulares de rotación y de velocidad (aceleraciones) en los tres ejes ortogonales fijos a su cuerpo, p. 9.

M

mosaico Es el resultado de unir imágenes pequeñas. Se realiza uniendo dos imágenes según unos puntos de control. Estos puntos de control son internos a cada imagen y no tienen una correlación conocida con un sistema de coordenadas externo, p. 9.

O

ortografía Es una presentación fotográfica de una zona de la superficie terrestre, en la que todos los elementos presentan la misma escala, libre de errores y deformaciones, con la misma validez de un plano cartográfico.

OSD *On screen display*. Se trata de un dispositivo que realiza una conmutación controlada entre la señal de vídeo y la señal correspondiente al GPS u otra información que posteriormente es monitorizada. De ésta forma, sólo se ven afectadas algunas líneas y por un determinado tiempo por la inclusión de la información que se desee mostrar en el cockpit, p. 28.

P

Pinhole lens Las cámaras con este tipo de lentes se conocen, como cámara estenopeica. Es una cámara fotográfica sin lente, que consiste en una caja a prueba de luz con un pequeño agujero donde entra luz, p. 33.

R

rectificación Es el proceso de restar los ángulos de rotación de una imagen (pitch - cabeceo, roll - balanceo y yaw - guiñada) para dejar la imagen alineada con los ejes de una imagen ortogonal, p. 9.

remote sensing Remote Sensing o la teledetección es la adquisición de información a pequeña o gran escala de un objeto o fenómeno, mediante el uso de la grabación o en tiempo real de detección de dispositivos inalámbricos o en contacto físico con el objeto (por ejemplo, como por medio de aviones, naves espaciales, satélites, boyas, o barco). En la práctica, la teledetección es la disputa de recogida de datos mediante el uso de una variedad de dispositivos para recopilar información sobre un objeto dado o de una zona.

GLOSARIO

rubber-sheeting En la cartografía, rubber-sheeting o deformación elástica se refiere al proceso por el cual una capa que está distorsionada puede ser ensamblada con una capa adyacente geográfica y así hacer coincidir las imágenes, tales como las imágenes de satélites que están en los mapas digitales, tales como Google Map, MapQuest, Microsoft Virtual Earth o Yahoo. La deformación elástica es necesario porque las imágenes y los datos vectoriales rara vez coinciden correctamente debido a diversas razones, tales como el ángulo en que fue tomada la imagen, la curvatura de la superficie de la tierra, los movimientos de menor importancia en la plataforma de imágenes (por ejemplo, como un satélite o una aeronave), y otros errores en las imágenes, p. 8.

S

stitching El stitching o foto stitching es el proceso por el cual se combinan múltiples imágenes para producir una imagen panorámica o una imagen de alta resolución, normalmente mediante un programa de ordenador, p. 4.