

Satellite information fusion with low-altitude unmanned aerial vehicle flights for estimating weed coverage.

Bussi, Ulises^{1,3}, Sauczuk, Martín¹, Mandile, Guillermo¹, Poggio, Santiago L.², and Oliva, Damián^{1,3}

¹ Universidad Nacional de Quilmes, Departamento de Ciencia y Tecnología Roque Sáenz Peña 352, Bernal Buenos Aires, Argentina (B1876BXD)

² IFEVA, Universidad de Buenos Aires, CONICET. Facultad de Agronomía, Cátedra de Producción Vegetal. Av. San Martín 4453, Buenos Aires, Argentina (C1417DSE)

³ Consejo Nacional de Investigaciones Científicas (CONICET), Argentina

Abstract. The detection, geolocation, and classification of weeds in agricultural fields is a problem of interest associated with Precision Agriculture (PA). The main contribution of this work is to describe a workflow (feasible to automate) based on open-source software tools and open information to: 1) measure the spatiotemporal evolution of weed patches through satellite images, and 2) register high-resolution images (taken at low altitude) on top of the satellite image to identify the weeds that compose the detected patches. To merge the satellite and low-altitude information, the following problems must be solved: 1) correct distortions in the acquired images; 2) develop an image formation model that allows registering the low-altitude image on top of the satellite image, and 3) analyze green indices to measure patch coverage in both multispectral satellite images and RGB images obtained from a camera mounted on an unmanned aerial vehicle. Finally, the feasibility of merging information is demonstrated through an analysis of the correlation in the coverage measures obtained from satellite and low-altitude images.

Keywords: Weed, Sentinel-2, Unmanned Aerial Vehicle, UAV.

Fusión de información satelital con vuelos a baja altura de vehículos aéreos no tripulados para estimar la cobertura de malezas.

Bussi, Ulises^{1,3}, Sauczuk, Martín¹, Mandile, Guillermo¹, Poggio, Santiago L.², and Oliva, Damián^{1,3}

¹ Universidad Nacional de Quilmes, Departamento de Ciencia y Tecnología Roque Sáenz Peña 352, Bernal Buenos Aires, Argentina (B1876BXD)

² IFEVA, Universidad de Buenos Aires, CONICET. Facultad de Agronomía, Cátedra de Producción Vegetal. Av. San Martín 4453, Buenos Aires, Argentina (C1417DSE)

³ Consejo Nacional de Investigaciones Científicas (CONICET), Argentina

Resumen La detección, geolocalización y clasificación de malezas en campos agrícolas es un problema de interés asociado a la Agricultura de Precisión (AP). El aporte principal de este trabajo es describir un flujo de trabajo (factible de automatizar) basado en herramientas de software libre e información abierta para: 1) medir la evolución espaciotemporal de los parches de malezas a través de imágenes satelitales y; 2) registrar las imágenes de alta resolución (tomadas a baja altura) sobre la imagen satelital, para identificar las malezas que componen los parches detectados. Para fusionar la información satelital y de baja altura, se deben resolver los siguientes problemas: 1) corregir las distorsiones en las imágenes adquiridas; 2) desarrollar un modelo de formación de imágenes que permita registrar la imagen a baja altura sobre la imagen satelital; 3) analizar los índices de verde para medir la cobertura de los parches, tanto en las imágenes satelitales multiespectrales, como en las imágenes RGB obtenidas desde una cámara montada en un vehículo aéreo no tripulado. Finalmente, se muestra la factibilidad de realizar la fusión de información a partir de un análisis de la correlación en las medidas de cobertura obtenidas de las imágenes satelitales y de las de baja altura.

Keywords: Malezas, Sentinel-2, Unmanned Aerial Vehicle, UAV.

1 Introducción

La agricultura de precisión (AP) se define como la gestión de las parcelas agrícolas basada en tecnologías de la información, integrando datos de múltiples fuentes para optimizar la toma de decisiones en la producción agrícola. La AP incluye técnicas y métodos de manejo agronómico que utilizan información sitio-específica y consideran su heterogeneidad y variabilidad espacio-temporal. Actualmente, el manejo sitio-específico de las malezas es prioritario por la necesidad de reducir tanto costos como impactos socio-ambientales. El consumo de herbicidas en nuestro país, supera los 1200 millones de dólares anuales [1] y el impacto ambiental de los agroquímicos es una preocupación central de la sociedad. Estas

problemáticas promueven el interés en el desarrollo de técnicas para aplicar herbicidas en dosis variable con recetas optimizadas [2]. Esto implica la necesidad de conocer la distribución de los parches y determinar la identidad de las malezas que los componen.

Actualmente están disponibles imágenes del satélite Sentinel-2 [3] que permiten un estudio espacio-temporal de la evolución de estos parches. Luego de la detección y geolocalización de los parches en la imagen satelital, se necesita una identificación de las malezas a través de imágenes en alta resolución tomadas a baja altura a partir de vuelos de un vehículo aéreo no tripulado (UAV, por sus siglas en inglés).

Este trabajo propone desarrollar y describir una metodología simple, factible de automatizar y basada en tecnologías de software libre e información abierta que permita realizar este tipo de estudios.

En la Fig. 1 se muestra el sistema propuesto, la metodología parte de tomar una imagen NDVI satelital del campo a estudiar y realizar un vuelo adquiriendo un video sobre el mismo. En la Fig. 1A, en rojo se puede ver la trayectoria del UAV sobre el mapa NDVI satelital. Para los fotogramas adquiridos se calcula el campo visual de la cámara, en Fig. 1B se muestra una ampliación la imagen satelital con los campos de visión como rectángulos. Luego, se analizan las imágenes correspondientes a los fotogramas de interés (Fig. 1C), de dónde se pueden calcular índices de vegetación (Fig. 1D) que permitan determinar porcentajes de cobertura, presencia de malezas o analizar un zoom sobre la misma que permitan identificar los especímenes presentes (Fig. 1E). Una vez identificado los individuos, puede volverse a la imagen satelital para clasificar los parches.

En el marco de la agricultura de precisión, el uso de vehículos para la adquisición de información (sea visual, o de distintos tipos de sensores) toma un rol fundamental, puesto que permite sistematizar la toma de datos. Los vehículos aéreos no tripulados (UAVs) ofrecen ventajas adicionales tales como la posibilidad de automatizar los recorridos, la independencia al estado del terreno y la adquisición no invasiva (al encontrarse lejos del cultivo no puede dañarlo o afectar el desarrollo natural del mismo). En particular, en este trabajo se propone utilizar un UAV, con cuatro rotores y una serie de sensores que permiten la registración de variables de interés.

Existen trabajos previos, por ejemplo [4] , donde se estudia la registración de imágenes obtenidas con vuelos de un UAV para generar mosaicos de imágenes con la menor intervención posible. Existen múltiples procedimientos para realizarlo, dependiendo de la información disponible. El proceso de registración propuesto por [4] se basa en el contenido digital de cada imagen, sin utilizar meta-datos, como ser la localización geográfica.

También existen programas comerciales de fotogrametría (más sofisticados y precisos) que producen la reconstrucción 3D de la superficie observada a través del vuelo del UAV (algunos ejemplos son los programas comerciales [5]: Pix4Dmapper Pro (versión *desktop* y versión *cloud*); Agisoft Photoscan; Photo-Modeler UAS; Drone2Map; 3DF Zephyr Aerial; ReCap 360 (versión *cloud*) [6,7] . Para facilitar el proceso de reconstrucción, estos métodos pueden usar informa-

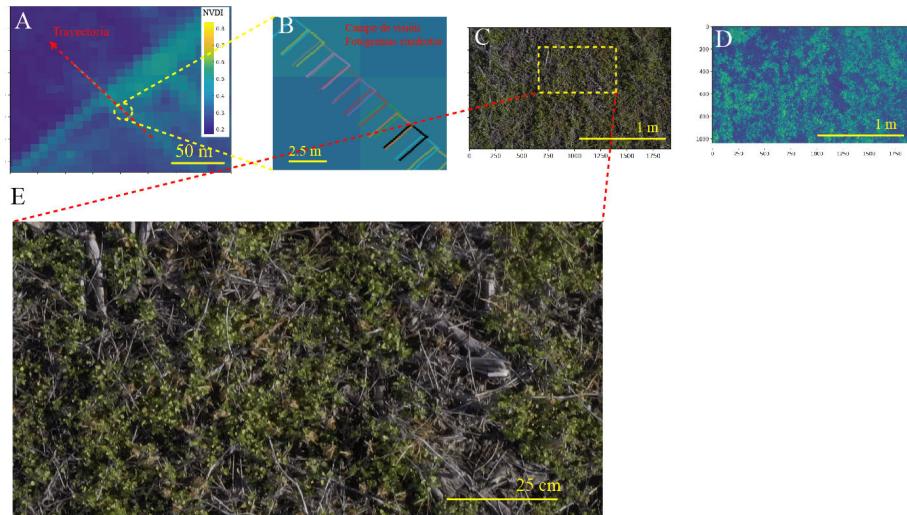


Figura 1. Sistema propuesto. A) Imagen satelital NDVI de Sentinel-2 y trayectoria del UAV (flecha roja). B) Ampliación de una zona y recuadro de campo visual del UAV. C) Imagen a 3m de altura desde la cámara del UAV. D) Índice de verde para cámaras RGB. E) Detalle de las malezas en mayor resolución.

ción de la trayectoria del vuelo (extraída del GPS y/o la IMU del UAV). Luego se analiza la secuencia de imágenes obtenidas y se buscan puntos de interés (*feature points* o puntos clave) que puedan identificarse en distintos fotogramas del vuelo. Además se puede introducir información *a priori* sobre la ubicación precisa de algunos puntos de la escena (puntos de calibración).

Si se conoce el comportamiento de los rayos (a través del modelo intrínseco de la cámara) y se tiene una medición inicial de la trayectoria del UAV; se pasa a una optimización que triangula los *feature points* asociados a un mismo punto sobre la superficie y se estima de forma robusta las transformaciones geométricas entre fotogramas (por ejemplo utilizando la técnica RANSAC). Un método tradicional para este análisis es el algoritmo de *bundle-adjustment* [8] que permite estimar la posición de los puntos 3D de la superficie y la trayectoria de la cámara. Con procesamientos adicionales se pueden estimar los modelos digitales de elevación, de superficie (DEM y DSM) y también generar ortofotografías [4,5].

Si el objetivo de la registración está relacionado con una reconstrucción 3D precisa de la superficie, entonces se requiere obligatoriamente la utilización de las técnicas antes mencionadas. Sin embargo, para realizar estos procedimientos es necesario porcentajes altos de solapamientos entre fotogramas sucesivos. Esto exige guardar gran cantidad de imágenes en alta resolución y un poder de cómputo elevado durante la fase de procesamiento *offline* para la generación de los mapas.

Se considera que para identificar las malezas que componen parches grandes (mayores a $100m^2$, que corresponde a la máxima resolución espacial de las

imágenes de Sentinel-2), no es imprescindible generar mosaicos de precisión (mapas de ortofotografías). Tampoco se considera necesario un posicionamiento más preciso que el obtenido con un GPS tradicional (alrededor de 3m [9]). Debido al costo computacional que implica la aplicación (por el solapamiento necesario antes mencionado) de estas técnicas de triangulación y reconstrucción 3D de la superficie en campos productivos reales; en este trabajo se propone evitar estimar el modelo digital de elevación, de superficie o de ortofotos, y en su defecto, utilizar el provisto por el Instituto Geográfico Nacional [10]. El método propuesto se basa en la suposición de que la región observada en cada fotograma obtenido por la cámara del UAV puede aproximarse por un plano con pendiente despreciable. Como se busca evitar la aplicación de técnicas de *feature-matching*, la incertezza espacial de las imágenes de alta resolución sobre la superficie de interés, quedará determinada por la incertezza del GPS y de la IMU. Se considera que este enfoque permite identificar las malezas en parches de grandes dimensiones (mayor a la incertezza del GPS) pudiendo responder una pregunta de interés para campos productivos de gran dimensión.

2 Modelo intrínseco y corrección de distorsiones

2.1 Modelo de *pin-hole*

Uno de los modelos de proyección más simple, que permite interpretar y predecir qué se verá en la imagen, es el modelo *pin-hole*. Este modelo consiste de tres pasos.

- Una roto-traslación que relaciona el sistema de referencia del mundo con el del UAV, donde se encuentra la cámara. Esta se aplica sobre las coordenadas de cada punto como se muestra en la ecuación (1) donde wX son las coordenadas de un punto en la trama del mundo, dX las del mismo punto en la trama del UAV y dT la transformación que relaciona ambas tramas.
- Una transformación a coordenadas homogéneas, proyectando cada punto sobre un plano a distancia unitaria del UAV, que se calcula como se muestra en la ecuación (2).
- Un reescalamiento y traslación para ubicar cada punto en las coordenadas del sensor digital de la cámara, hX . Que viene dado por la expresión de la ecuación (3).

$${}^dX = \begin{bmatrix} {}^dX \\ {}^dY \\ {}^dZ \\ 1 \end{bmatrix} = {}^dT {}^wX = {}^dT \begin{bmatrix} {}^wX \\ {}^wY \\ {}^wZ \\ 1 \end{bmatrix} \quad (1)$$

$${}^hX = \begin{bmatrix} {}^dX / {}^dZ \\ {}^dY / {}^dZ \\ 1 \end{bmatrix} \quad (2)$$

$${}^c X = C {}^h X = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} {}^h X \quad (3)$$

Los parámetros que interfieren en esta transformación, pueden separarse en dos grupos, los parámetros intrínsecos Γ (distancia focal $\{f_x, f_y\}$ y centro de imagen $\{c_x, c_y\}$) son aquellos asociados a la cámara en sí, no se modifican externamente y están comprimidos en la matriz C . Los parámetros extrínsecos Θ ($\{r_1, r_2, r_3, t_x, t_y, t_z\}$) dependen estrictamente de la pose de la cámara en el sistema de referencia y quedan representados en la matriz ${}^w T$. Los primeros tres (r_i) representan la orientación de la cámara medida en algún sistema de referencia, mientras que los restantes (t_i) representan la posición cartesiana de la cámara en la trama del mundo.

2.2 Distorsión

El modelo anterior asume que no existe ningún tipo de distorsión en el viaje de los rayos desde el objeto observado hasta el sensor. En general esto no se cumple, ya que existen distorsiones asociadas a la construcción de la lente (intencionales o accidentales). Generalmente, las distorsiones asociadas a las cámaras se caracterizan como tangenciales y radiales. Las primeras están asociadas al error de alineación entre el sensor digital de la cámara y la lente. Mientras que las segundas están relacionadas a cómo la lente curva los rayos que llegan.

Las distorsiones radiales pueden modelarse mediante un polinomio en la distancia al origen en las coordenadas homogéneas ($r_h = \sqrt{x_h^2 + y_h^2}$) que suele escribirse como:

$$r_d = r_h (1 + k_1 r_h^2 + k_2 r_h^4 + k_3 r_h^6) \quad (4)$$

Donde $\{k_1, k_2, k_3\}$ son los parámetros de distorsión radial.

Las distorsiones tangenciales pueden modelarse mediante una expresión que incluye información de la posición cartesiana y el radio en coordenadas homogéneas:

$$\begin{aligned} x_d &= x_h + 2p_1 x_h y_h + p_2 (r_h^2 + 2x_h^2) \\ y_d &= y_h + 2p_2 x_h y_h + p_1 (r_h^2 + 2y_h^2) \end{aligned} \quad (5)$$

Donde $\{p_1, p_2\}$ son los parámetros de distorsión tangencial.

Los parámetros anteriormente descritos se agrupan como parámetros intrínsecos de la cámara, dado que dependen de la construcción de la misma.

2.3 Calibración intrínseca de la cámara del UAV y corrección de distorsiones

En este estudio se utiliza el UAV DJI Mavic Pro 2. El tamaño de las imágenes utilizadas en este trabajo fue de 3840×2160 pixeles. Normalmente, las cámaras de los UAVs tienen campos de visión amplios y es común que existan fenómenos de distorsión (tipo barril) significativos. Por este motivo se decidió aplicar

un método de calibración y corrección de distorsiones. Para esto se tomaron 47 fotografías de un *chessboard* (Fig. 2A-B). Se estimaron los coeficientes de distorsión (descritos en la sección anterior) utilizando la función **calibrateCamera** de la librería de **OpenCV** para Python [11]. Para estimar las incertezas de los coeficientes, se utilizó la técnica de *bootstrap*, remuestreando sobre las imágenes de calibración. Los valores estimados fueron: $f_x = 4545 \pm 80$, $f_y = 4542 \pm 75$, $c_x = 1976 \pm 110$ y $c_y = 1070 \pm 40$ todas en unidades de pixeles.

Como ejemplo de análisis del efecto de las distorsiones, en la Fig. 2C se muestra una imagen de un campo experimental. Se ve que las líneas asociadas a una calle presentan una curvatura apreciable (elipses amarillas). Esta distorsión de la imagen debido a la lente de la cámara, puede ser corregida utilizando métodos de interpolación que corrigen las distorsiones y simulan una cámara tipo *pin-hole*. En la Fig. 2D se muestra la corrección utilizando el método de anti-distorsión de **OpenCV**. Comparando las líneas rojas con el borde de la calle se observa que la distorsión de barril es corregida.

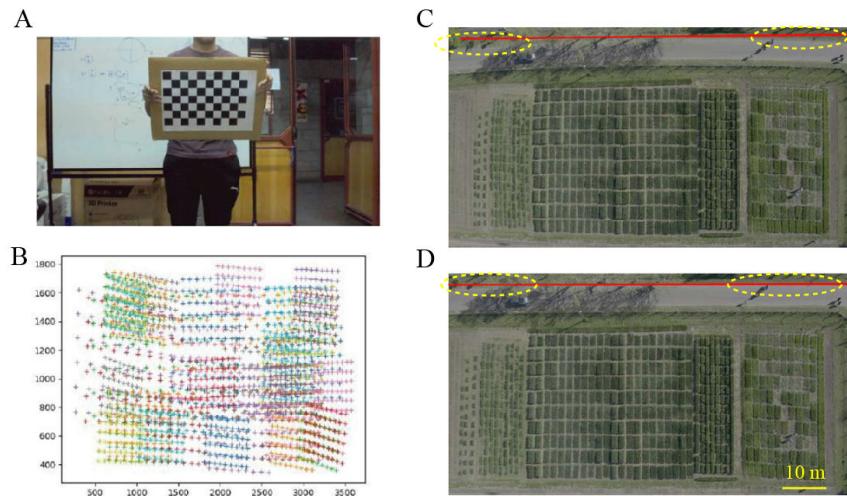


Figura 2. Modelo intrínseco y corrección de distorsiones en la cámara del UAV. A) Imagen de *chessboard* para la calibración. B) Vértices de los *chessboards* para 47 imágenes. C) Campo experimental, las elipses amarillas muestran la curvatura por distorsiones de la cámara. D) Corrección de distorsiones.

2.4 Selección de la altura y velocidad del vuelo

La ecuación (6) describe la relación entre el campo visual θ_{fov} y los parámetros geométricos como la altura de vuelo h , el tamaño de pixel W_{pix} del sensor CCD, la distancia focal de la cámara f_{pix} , y el ancho l sobre de la superficie observada (ver Fig. 3A).

$$\tan \frac{\theta_{fov}}{2} = \frac{W_{pix}/2}{f_{pix}} = \frac{l/2}{h} \quad (6)$$

$$h = f_{pix} \frac{d}{d_{pix}} \quad (7)$$

Adicionalmente, si se observa un objeto sobre una superficie de tamaño d cuya dimensión sobre la imagen es d_{pix} , se puede escribir esa relación como se muestra en la ecuación (7).

Teniendo en cuenta el tamaño del sensor CCD obtenemos que para un vuelo a 3m de altura la resolución en 0,065cm/pix. Mientras que para un vuelo a 20m de altura es 0,44cm/pix. Por lo tanto, la altura de vuelo queda determinada por el objetivo que se esté buscando. Si se quiere detectar parches del orden de algunos centímetros se puede volar a 20 metros de altura. Sin embargo, si se quiere identificar malezas de hoja pequeña se necesita volar a 3 metros.

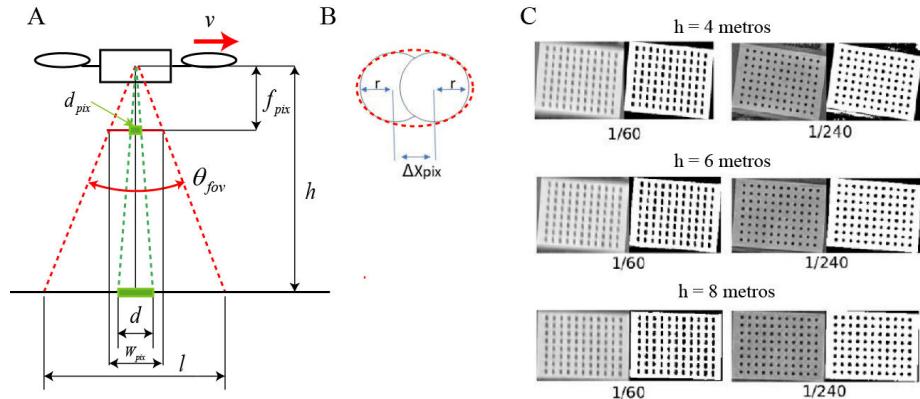


Figura 3. Relación geométrica entre los parámetros de vuelo y del sensor CCD. A) Esquema de relación entre parámetros del Modelo *pin-hole* de la cámara. B) Efecto del borroneo. C) Imágenes para ajuste del efecto de borroneo.

Otro efecto relevante en la calidad de la imagen obtenida desde el UAV, es el efecto de borroneo (*blurring* en inglés). Una vez fijada la altura de vuelo, se procede a analizar como varía el borroneo según la velocidad de vuelo v , y el *shutter speed* Δt seleccionado en el CCD. Se propone alcanzar un ΔX_{pix} pequeño (Fig. 3B). Reemplazando $d_{pix} = \Delta X_{pix}$ y $d = v \cdot \Delta t$ en la ecuación anterior, obtenemos:

$$\Delta X_{pix} = f_{pix} \frac{v \Delta t}{h} \quad (8)$$

Para estudiar el efecto de borroneo, se voló el UAV sobre un conjunto de patrones de círculos (Fig. 3C) variando la altura h , velocidad v y *shutter speed*

Δt . Estos experimentos mostraron que la predicción de borroneo de la ecuación Ec. 8 sobreestima el borroneo en un factor $k = 2$. eligiendo $\Delta t = 1/240s$, $h = 3m$ y $\Delta X_{pix} < 5$ la velocidad estimada es $0,9m/s$.

3 Modelo geométrico simplificado del problema

En la sección anterior se describe un método de calibración que permite estimar los parámetros intrínsecos de la cámara y para el CCD. En esta sección se aborda la medición de los parámetros extrínsecos asociados a un vuelo que corresponde a la estimación de la pose (posición y orientación) de la cámara durante el vuelo. También se propone un modelo geométrico simple y se discute la fusión (alineación) de las imágenes satelitales y las obtenidas a baja altura con la cámara del UAV. Existen cuatro sistemas de referencias que son importantes para este problema (ver Fig. 4A-B):

1. Sistema de Referencia $\{W_1\}$: tiene asociadas las mediciones de $(lat_{oc}, long_{oc}, z_{oc}, \theta_{oc})$ del GPS y la IMU del UAV. Donde $(lat_{oc}, long_{oc})$ corresponden a la latitud y longitud, se obtienen del GPS (WGS84, EPSG:4326). El *logfile* asociado al vuelo del UAV también proporciona la altura respecto al suelo z_{oc} y el ángulo θ_{oc} (de la cámara) respecto a la dirección Norte. La cámara del UAV apunta en la dirección de la gravedad (con pose estabilizada).
2. Trama $\{W_2\}$: en las regiones estudiadas en este trabajo, está asociada al sistema de referencia EPSG:32721 y EPSG:32720 o sea, para cada punto del mapa conocemos sus coordenadas $(x_2, y_2) = {}^{w_2}(X, Y)$ en este sistema de referencia (en unidades de metros). Existe una transformación no lineal (denominada F_{NL}), que permite calcular ${}^{w_2}(X, Y)$ a partir de las coordenadas ${}^{w_1}(lat, long)$ en la trama $\{W_1\}$. Se representa esta relación matemática como: ${}^{w_2}(X, Y) = F_{NL}[{}^{w_1}(lat, long)]$ que se realiza con la función **transform** de la librería **pypyproj** [12].
3. Trama $\{S\}$: Este sistema de referencia corresponde a la imagen satelital. La posición de cada pixel ${}^s(u_S, v_S)$ de esta imagen, está relacionado con ${}^{w_2}(X, Y)$ a través de una transformación afín, ${}^{w_2}(X, Y) = F_{s \rightarrow w_2}[{}^s(u_S, v_S)]$. La información de esta transformación está incorporada como meta-dato en el archivo de formato **jpeg 2000** de la imagen satelital de Sentinel-2.
4. Trama $\{C\}$: Esta trama corresponde al sistema de referencia de la imagen en la cámara del UAV. Cada pixel ${}^c(u_c, v_c)$ de la imagen está relacionado con ${}^{w_1}(X, Y)$ a través de una transformación no lineal, ${}^{w_1}(X, Y) = F_{NL2}[(u_c, v_c)]$ (ver Fig. 4B), que se describe a continuación.

Como se muestra en la Fig. 4A, el UAV vuela sobre el campo (la trayectoria está representada por la línea punteada negra) con pose $(lat_{oc}, long_{oc}, z_{oc}, \theta_{oc})$ que se obtienen del *logfile* del UAV, mientras capture imágenes del suelo. Los rectángulos azules representan el campo visual de la cámara, junto con su sistema de referencia ($\{C\}$), el rectángulo verde representa la trama asociada a la imagen

satelital mientras que los ejes negros representan la trama del mundo $\{W_2\}$. Como puede verse, estas últimas dos tramas se diferencian por un reescalamiento y traslación. La información del *logfile* permite conocer la pose de la cámara (parámetros extrínsecos). Asumiendo que existen pendientes despreciables en la superficie del campo y por lo tanto, se puede aproximar localmente al suelo con un plano. Bajo este supuesto es posible convertir puntos de la imagen a las otras tramas de interés.

Lo primero que se hace es calcular los extremos de la imagen orientados con el norte usando la información del ángulo θ_{oc} (cruces rojas en la Fig. 4A). Asumiendo que la imagen formada puede describirse por el modelo *pin-hole* con distancia focal f_{pix} y que el UAV se encuentra a una altura z_{oc} , se puede retroproyectar los rayos asociados a los vértices de la imagen sobre el suelo en un sistema de referencia arbitrario en metros despejando de la ecuación (7). Por último se transforma EPSG:4386 utilizando la inversa de la proyección elipsoidal [13].

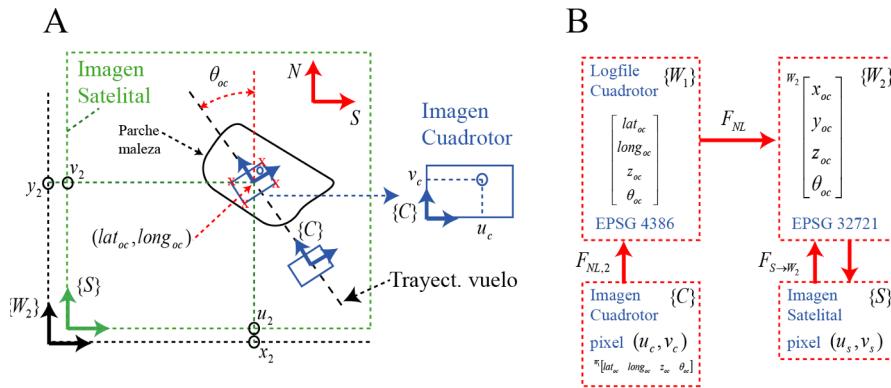


Figura 4. Modelo geométrico simplificado, sistemas de coordenadas y geolocalización. (A) Esquema de vuelo del UAV junto a todas las tramas de interés. (B) Esquema de la relación de coordenadas y sistemas de referencia. Más detalles en el texto.

4 Flujo de trabajo del sistema

En la Fig. 5 se muestra el flujo de trabajo del sistema propuesto.

4.1 Imágenes satelitales de Sentinel-2

Se programó una función que descarga automáticamente una imagen satelital centrada en $(lat_{campo}, long_{campo})$ en una fecha genérica, y devuelve un recorte de un sector de radio R (en este caso 1200 metros). Las imágenes satelita-

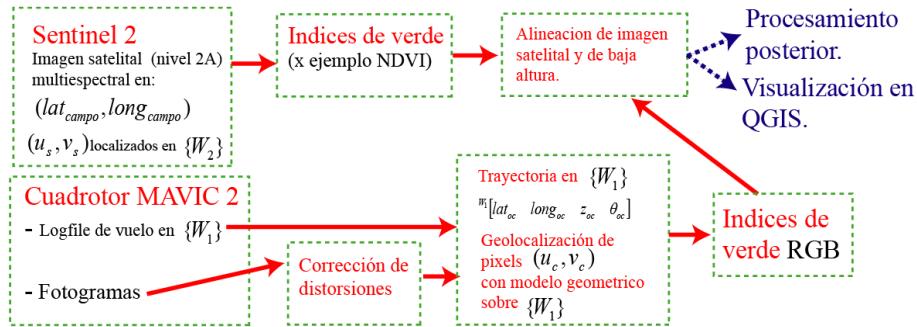


Figura 5. Diagrama de flujos del sistema propuesto.

les de la región de interés son descargadas utilizando la función **SentinelAPI.download_all** de la librería **SentinelSat**. Tienen un tamaño de 109,8 km en cada dimensión (10980×10980 pixeles, en formato **jpeg 2000**). En el proceso de recorte del campo de interés, se utilizan nuevamente las transformaciones entre $\{W_1\}$ y $\{W_2\}$ descritas en la sección anterior. Este trabajo se centra en las bandas con resolución espacial de 10 metros.

4.2 Imágenes adquiridas desde el UAV

En este trabajo se utilizó el UAV *DJI Mavic Pro 2*. Debido al campo visual amplio de la cámara, se introducen distorsiones importantes. Para corregir estas distorsiones se calibra la cámara y se obtienen los parámetros intrínsecos como se describió en la sección 2.3. Finalmente se aplica la función **undistort**, generándose una imagen equivalente a la adquirida por una cámara *pin-hole* con las mismas características que la imagen original pero sin distorsiones.

4.3 Registro de vuelo y de video

Al tomar videos con el UAV *DJI Mavic Pro 2* puede obtenerse información adicional que se adquiere en dos formatos distintos, el **registro de video** y el **registro de vuelo**.

El registro de video se presenta como un archivo de *subtitle* que se genera automáticamente con el mismo nombre que la grabación y se guarda en la memoria interna del UAV. Este archivo de texto cuenta con información de parámetros al momento de adquirir la imagen del video. La información pertenece a tres categorías distintas: parámetros de cuadro, parámetros de configuración de la cámara y parámetros de localización. En la Tabla 1 se describen los nombres y significado de cada uno de los parámetros.

La frecuencia de estos datos coincide con la del video, en este caso $30Hz$ (un cuadro cada 33ms).

Del conjunto de datos mencionado anteriormente, son de importancia para este trabajo la información de localización, que se utiliza para georreferenciar las

Cuadro 1. Descripción de parámetros en el registro de video del UAV DJI Mavic Pro 2.

Categoría	Nombre	Descripción
Parámetros de cuadro	FrameCnt	Número de cuadro en el video.
	DiffTime	Duración del cuadro en ms.
	DateTime	Fecha y hora de captura del cuadro.
Parámetros de configuración de cámara	iso shutter fnum ev	Sensibilidad del sensor. Velocidad de obturación de la cámara. Proporción diámetro de apertura sobre distancia focal. Valor de exposición.
Parámetros de localización	latitude longitude altitude	Latitud del UAV al tomar imagen. Longitud del UAV al tomar imagen. Altura del UAV sobre el nivel del mar.

imágenes, así como la fecha y hora correspondiente a cada *frame*, que se utiliza para sincronizar con el registro de vuelo.

El registro de vuelo es un archivo almacenado remotamente en la memoria del dispositivo (*smartphone* o *tablet*) anexado al control del UAV. Este registro hace las veces de caja negra de la misma, conteniendo una amplia variedad de información de la misma, lo que permite depurar y encontrar problemas de ser necesario. Este archivo cuenta con información temporal, de localización detallada, incluyendo velocidad, ángulos de orientación, estado del sensor de GPS, ángulos de orientación del estabilizador de cámara, información de la conexión remota con el control y estado de baterías del UAV entre otros parámetros que se compilan en un total de 45 celdas. Esta información se registra con una frecuencia aproximada de 10Hz (cada 100ms). En este trabajo, del registro de vuelo solo se utiliza información de la altura del UAV y el *heading* (ángulo de orientación medido respecto del norte).

4.4 Fusión de información satelital-UAV

Al grabar videos durante los vuelos del UAV sobre los cultivos de interés, se obtiene una gran cantidad de datos redundantes: las imágenes consecutivas pueden compartir un gran área de intersección (valores superiores al 90 % de redundancia). Como se mencionó anteriormente, en este trabajo se busca contar con información de todo el área abarcada por el vuelo, pero intentando disminuir la redundancia. Para esto se propone reducir el número de imágenes utilizadas, admitiendo un área máxima de intersección deseada (I_{thresh}) entre imágenes, que permita evitar segmentos sin información en el recorrido. De esta manera, se optimiza el uso de los recursos y se obtiene la información necesaria sin pérdida de calidad.

Para encontrar el conjunto de cuadros que cumplan un área máxima de intersección I_{thresh} , se parte de tomar cada imagen de un video, y utilizando la

información del registro de vuelo, de video y el modelo descrito en la Sección 3, calcular la posición en coordenadas GPS de las esquinas de la imagen. Si se trata del primer cuadro, este se guarda junto con la posición de las esquinas en una lista, para cada cuadro siguiente se analiza si el área comprendida por este se interseca y en qué proporción con los cuadros guardados. Si el área de intersección con alguno de los cuadros guardados es mayor que I_{thresh} , se descarta, en caso contrario, se guarda en la lista y se pasa al cuadro siguiente.

Hay algunas implicaciones con este algoritmo, la comparación con todos los cuadros guardados hace que sea extensivo en uso de tiempo. Esto principalmente porque se supone que el vuelo del UAV es libre, si se tomase la simplificación de que el vuelo no volverá a pasar por un lugar (por ejemplo moviéndose en línea recta) podría compararse solo con el cuadro anterior. Para resolver esta situación se propone definir un mínimo salto de cuadros mf_s (por sus siglas en inglés), es decir, en vez de analizar todos los cuadros, cuando el cuadro i es analizado el siguiente cuadro a tratar será el $i + mf_s$ con $mf_s \geq 1$. Este supuesto es válido considerando que el UAV se mueve relativamente lento por lo que no habrá un gran cambio de escena entre cuadros consecutivos y permite reducir el número de comparaciones. En la Fig. 6 puede verse un diagrama de flujo del algoritmo utilizado.

La selección de este salto mínimo es manual y viene dada por la configuración de vuelo considerando que, a mayor velocidad se reduce mf_s , puesto que habrá menor intersección entre cuadros si el UAV se mueve rápido. Y a mayor altura se sube este valor, ya que la escena no cambiará mucho entre cuadros si se observa a mucha distancia.

Una vez obtenido el conjunto de cuadros y la posición de sus esquinas es posible: extraer las imágenes del video para el posterior análisis (calculando los índices que se presentan en la siguiente sección), calcular la transformación que permita mapear la imagen a algún software de información geográfica para visualizarla sobre otros mapas como imágenes satelitales o de NDVI y extraer de los mapas de NDVI los pixeles correspondientes a cada imagen para comparar índices, entre otras cosas.

5 Índices obtenidos a partir de las imágenes satelitales y del UAV (RGB)

5.1 Índices obtenidos a partir de las imágenes Sentinel-2

El índice de verde utilizado para las imágenes satelitales fue el *Normalized Difference Vegetation Index* (NDVI), definido en (9). Donde B4 hace referencia a la banda 4 que capta las longitudes de onda en el rango de 650nm a 680nm (rojo) y la B8 el rango de 780nm a 900nm (infrarrojo cercano).

$$NDVI = \frac{B8 - B4}{B8 + B4} \quad (9)$$

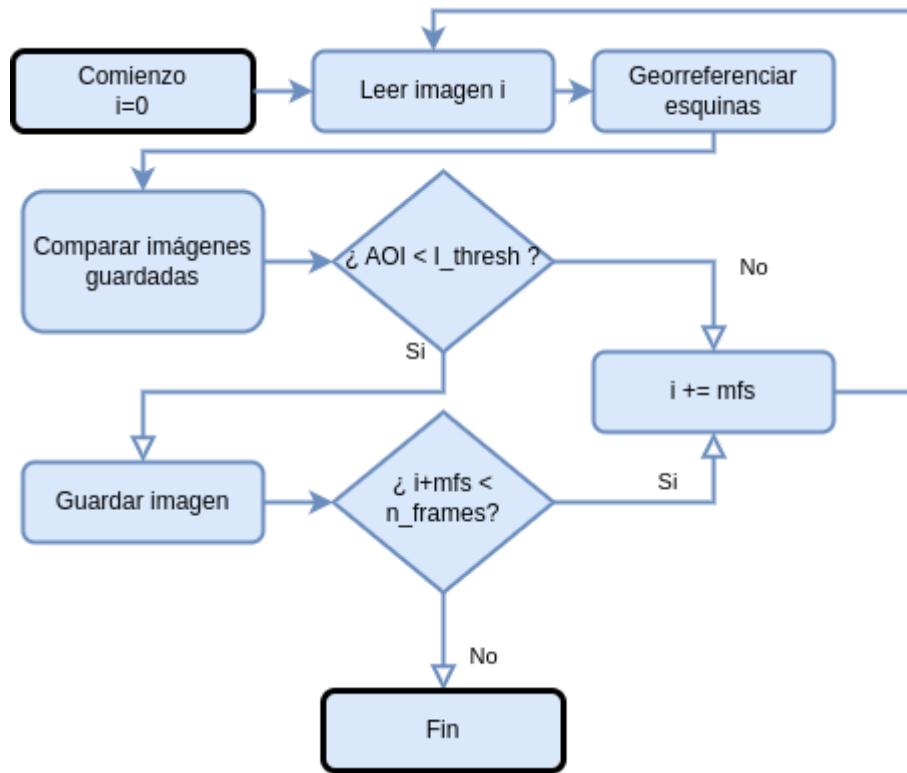


Figura 6. Diagrama de flujo del algoritmo para obtener los cuadros con menor área de intersección.

5.2 Índices obtenidos a partir de las imágenes RGB del UAV

Índice de exceso de verde (gEx): Este índice (Fig. 1D) se calcula tomando la información de la imagen RGB (Fig. 1C) en el espectro visible y para cada pixel se define como combinación lineal de los canales R, G y B:

$$gEx = 2G - R - B \quad (10)$$

Este índice queda definido para cada pixel con valores en el rango de $[-512, 512]$. Para poder realizar una comparación entre estos índices y los valores obtenidos desde las imágenes satelitales, se busca reducir este cálculo a un único número por imagen del UAV. Para esto se proponen las siguientes expresiones.

Promedio del exceso de verde (g_ex): se calcula el promedio de todos los pixeles en cada imagen de índice de exceso de verde gEx . Esta propuesta considera tanto los puntos que son clasificados como vegetación (muy verdes, con valores de índice > 0) como los que son clasificados como no vegetación (valores < 0), además de la intensidad de los mismos.

$$g_ex = \text{mean}(gEx) \quad (11)$$

Promedio enmascarado del exceso de verde (g_ex_mask): compara el índice con un umbral, para luego seleccionar solo los píxeles cuya clasificación sea “vegetación” y calcular un promedio sobre estos píxeles (en particular en este trabajo se escogió $\text{thresh} = 0$).

$$g_{\text{ex_mask}} = \text{mean}(gEx_i, \text{ si } gEx_i > \text{thresh}) \quad (12)$$

6 Análisis de los índices de cobertura de malezas sobre imágenes satelitales y del UAV

En la Fig. 7 se muestran imágenes NDVI satelitales de un campo previo (Fig. 7a) y luego de la cosecha (Fig. 7b-d). La región de análisis se ubica en la zona de la ciudad de Chivilcoy, provincia de Buenos Aires, Argentina, coordenadas 34,8681309S, 60,1103311W. En la imagen del 8 de Mayo de 2021 pueden verse sectores con alto índice NDVI en las regiones rodeadas por la elipse punteada, lo que indica vegetación en esas zonas. Dado que es una imagen satelital obtenida luego de la cosecha, y que en fechas anteriores (6 de Abril y 16 de Abril) no se presenta vegetación puede inferirse que se trata de parches de maleza que crecieron en esa ubicación. Por lo tanto, se planifica un vuelo para identificar dichos parches.

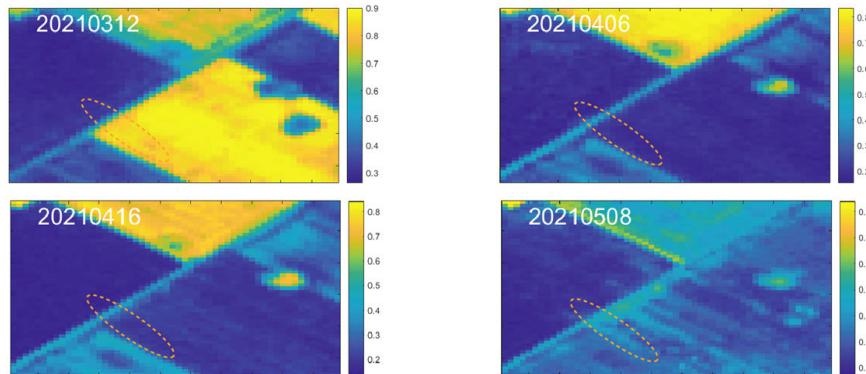


Figura 7. Medición de índice NDVI a partir de imágenes satelitales de Sentinel-2 en el periodo previo y posterior a la cosecha (desde el 12 de marzo de 2021 hasta 8 de mayo del 2021).

En la Fig. 8 se muestra un ejemplo de la trayectoria del UAV, volando a 3 metros de altura, sobre un parche de malezas. En la Fig. 8A se muestra la trayectoria elegida en rojo, sobre la imagen NDVI satelital de la zona de estudio. Sobre la trayectoria se seleccionan dos zonas de situación límite en lo que a valor de cobertura corresponde (círculos rojos), una zona de alta densidad de malezas y otra de baja densidad. En las Fig. 8B-C se muestra un fotograma de baja

altura, en alta resolución, para una región de alta cobertura de malezas. La Fig. 8B muestra la imagen original extraída del UAV. En Fig. 8C se muestra el índice de verde calculado para la misma imagen, este índice permite resaltar las zonas con plantas basado en la coloración de las mismas, aumentando el contraste entre vegetación y fondo. En las Fig. 8D-E se muestra un fotograma para una región de baja cobertura de malezas. La Fig. 8D muestra la imagen original, mientras en Fig. 8E se muestra el índice de verde. En la Fig. 8F se muestra un recorte con mayor detalle de la misma imagen. Es posible apreciar que la maleza principal corresponde a *Stellaria media* (capiquí), que es una especie de planta herbácea anual otoño-invernal de la familia Caryophyllaceae.

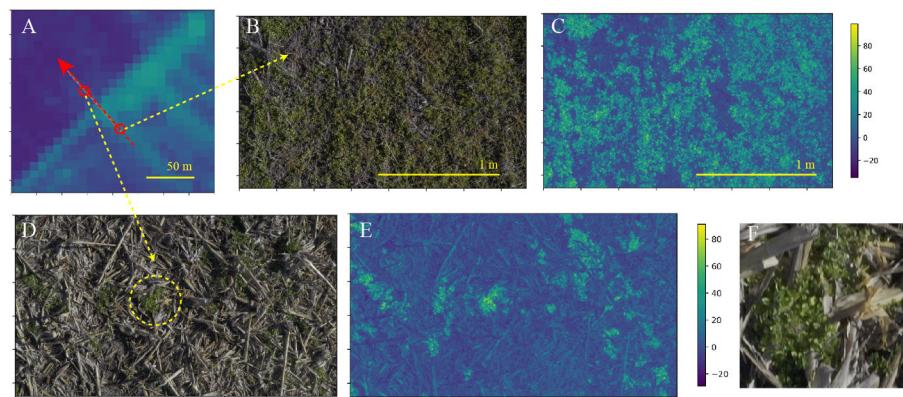


Figura 8. Ejemplo imágenes originales del UAV y medición de índices sobre las mismas.

Se compararon los valores obtenidos por los índices de verde contra el valor de NDVI para un vuelo del UAV. En la Fig. 9A-B se muestran, en azul, los índices de verde (g_ex y g_ex_mask) obtenidos de las imágenes del UAV (RGB) en función del NDVI obtenido del satélite promediado sobre el campo visual proyectado. En naranja se muestra un ajuste lineal realizado sobre los datos. En la Fig. 9C-D se muestran en trazo azul, los índices g_ex y g_ex_mask (C y D respectivamente) en función del avance sobre la trayectoria de vuelo del UAV. Debido a que el campo visual de la cámara del UAV es mucho menor al valor del pixel de la imagen de Sentinel-2, se promediaron los 3 índices sucesivos medidos a baja altura. Se calculó el coeficiente de correlación entre el índice NDVI obtenido de la imagen satelital y los índices de verde basados en la imagen RGB. El mayor coeficiente de correlación se obtuvo relacionando el índice NDVI con el índice g_ex_mask , con $r^2 = 0,92$.

Este resultado y la correlación entre datos muestra la factibilidad del método, confirmando qué, la geolocalización de la información del UAV es precisa: puede verse que el comportamiento de aumento o descenso de los índices varía de forma similar para los calculados desde imágenes satelitales y los hallados con imágenes tomadas a baja altura desde el UAV. El valor de correlación entre

imágenes de los distintos sistemas es aceptable, mostrando que es posible seleccionar los parches desde imágenes satelitales y realizar un posterior estudio en detalle con imágenes de alta resolución tomadas desde un UAV. Por último, que puede aplicarse una metodología simple, tanto en procedimiento como en costo computacional, para realizar este estudio. El costo computacional se califica como pequeño dado que el procedimiento es secuencial y los algoritmos aplicados no requieren de optimizaciones.

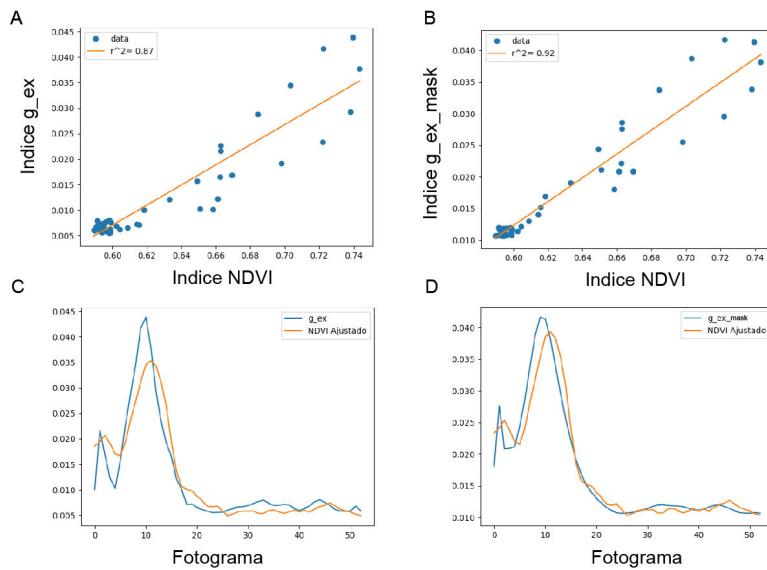


Figura 9. Resultados de comparación de índices satelitales y del UAV (RGB).

7 Discusión

A partir de los resultados obtenidos existen varios puntos que analizar:

En lo que respecta al modelo de cámara seleccionado (Sección 2), el modelo *pin-hole* es ampliamente utilizado en la bibliografía y no suele ser cuestionado. Como valor agregado puede destacarse el uso de la técnica de *bootstrap*. Lo que permite encontrar, además de un conjunto de valores esperados, las incertezas para cada parámetro ajustado. Este cálculo podría permitir en el futuro, propagar las incertezas en el mapeo sobre la superficie de observación.

En lo que respecta al modelo de distorsiones, en la Fig. 2C-D se muestra una comparativa entre las imágenes originales y las corregidas, demostrando la mejoría a la hora de corregir distorsiones ópticas. Este efecto necesita ser corregido en la mayoría de imágenes obtenidas con UAVs.

En lo que al modelo de borroneo respecta (Sección 2.4), este permite, a partir del conocimiento de los parámetros de la cámara, estimar condiciones de vuelo del UAV, de forma tal que la adquisición de información visual sea de buena calidad. La optimización de calidad de imágenes permite, como se mostró en este trabajo, detectar el tipo de maleza presente en las imágenes (Fig. 8F).

En lo que respecta al algoritmo para encontrar las imágenes que cubren el área de interés (minimizando la cantidad de imágenes a analizar); permite seleccionar un grupo de imágenes óptimo sobre el que se calculen los índices de color, reduciendo los costos de procesamiento y almacenaje de información.

En lo que respecta a índices de color utilizados en las imágenes adquiridas desde el UAV (sección 5.2), se mostró el alto grado de correlación entre los índices utilizados y los valores de NDVI extraídos de imágenes satelitales. Esto demuestra que la geolocalización correcta y permite relacionar imágenes RGB de baja altura con imágenes multiespectrales adquiridas de forma satelital.

Existe una gran cantidad de índices RGB que podrían evaluarse para analizar la vegetación, a la hora de comparar con los valores de imágenes satelitales multiespectrales [14]. La principal diferencia es que, generalmente, los índices propuestos en la bibliografía trabajan a nivel píxel en las imágenes mientras que los seleccionados en este trabajo condensan la información de cada imagen a un único valor.

8 Conclusiones

En este trabajo se propone una metodología simple para estudiar la dinámica espacio-temporal de la cobertura de malezas a través de imágenes satelitales y; luego determinar la identidad de las malezas con imágenes de alta resolución, obtenidas con vuelos a baja altura. La hipótesis de este trabajo asume que, para identificar los parches de malezas de gran tamaño (mayores a 10 metros, que corresponde a la máxima resolución espacial de las imágenes de Sentinel-2), no es necesario la construcción de modelos digitales de elevación y superficie por medio de técnicas avanzadas de fotogrametría. Estas técnicas computacionales son costosas en términos del número de imágenes necesarias y del procesamiento computacional requerido para las reconstrucciones tridimensionales. Si bien estas tecnologías serían una solución mucho más precisa para la localización y análisis de los parches de malezas, tienen el problema de requerir grandes volúmenes de información y procesamiento (debido al alto solapamiento que requieren este tipo de algoritmos). Además, un mapeo total de la superficie implicaría realizar largos recorridos sobre el campo en contradicción con la corta autonomía de la mayoría de los UAVs. Por lo tanto, se propone un estudio previo utilizando imágenes satelitales y luego un muestreo local en alta resolución de las zonas de interés.

El aporte principal de este trabajo fue describir un sistema basado en herramientas de software libre e información satelital abierta para estudiar la identidad y dinámica espacio-temporal de la cobertura de malezas en campos productivos. Es importante destacar que las imágenes a baja altura pueden adquirirse

con otros dispositivos distintos a un UAV, como ser cámaras montadas sobre un vehículo terrestre.

Los resultados presentados en la Fig. 9, muestran que existe una clara correlación entre las medidas de cobertura obtenidas de las imágenes satelitales y las obtenidas a baja altura con el vuelo del UAV, validando el procesamiento propuesto.

A futuro podrían aplicarse mejoras al método presentado, algunas ya fueron discutidas en la sección anterior cuando se trató cada uno de los puntos principales del presente trabajo, sin embargo existen modificaciones que implican cambios mayores ya sean de hardware o de software. Entre estas modificaciones se puede mencionar: 1) una mejor localización por medio de un GPS de mayor precisión o por la generación de mosaicos y alineaciones utilizando puntos clave geolocalizados. 2) Aplicar algoritmos de detección y segmentación de malezas que utilicen otros descriptores (además del índice de verde). Se podrían utilizar descriptores obtenidos de una red neuronal de convolución y/o redes neuronales como SegNet [15]. Esto permitiría automatizar y mejorar la identificación de cuáles son las malezas que existen sobre cada trayectoria realizada sobre los parches y cómo se distribuyen.

Agradecimientos

Se agradece a la Dra. Sandra Torrusio por la sugerencia de utilizar las imágenes de Sentinel-2.

Se agradece a los Ing. Braian Soullier y Nicolás Cuedo, por sus colaboraciones en la toma de imágenes para la calibración de la cámara.

Referencias

1. Cid, R., Moltoni, A. F., Fuica, A. M. (2010). El camino hacia una agricultura con maquinarias autónomas. el control de malezas. Curso Internacional de Agricultura de Precisión. 9. Expo de Máquinas Precisas. 4. 2010 07 14-16, 14, 15 y 16 de julio de 2010. Manfredi, Córdoba. AR.
2. Baruffaldi, B., Cavalieri, J. I., Baruffaldi, J. M. (2020). Segmentación de imágenes RGB para aplicación selectiva sobre barbecho. In XII Congreso de AgroInformática (CAI 2020)-JAIIO 49 (Modalidad virtual).
3. Sentinel-2, <https://sentinels.copernicus.eu/web/sentinel/user-guides>. Last accessed 19 Jun 2021
4. Castello, F. A., Albornoz, E., Martínez, C. E. (2018). Desarrollo de un sistema integral de adquisición de imágenes por dron, procesamiento y análisis para agricultura de precisión. In X Congreso de AgroInformática (CAI)-JAIIO 47 (CABA, 2018).
5. Romanos, P., Lovatto, M., Rossi, B. D. (2018). Criterios y recomendaciones para selección de software aplicativos para drones. In X Congreso de AgroInformática (CAI)-JAIIO 47 (CABA, 2018).
6. PIX4D, Pix4D Mapper 3.2 User Manual, 2016.
7. AGISOFT, Agisoft PhotoScan User Manual Professional Edition, Version 1.2,2016.

8. Triggs, B., McLauchlan, P. F., Hartley, R. I., Fitzgibbon, A. W. (1999, September). Bundle adjustment—a modern synthesis. In International workshop on vision algorithms (pp. 298-372). Springer, Berlin, Heidelberg.
9. Zogg, J. M. GPS Essentials of Satellite Navigation. u-blox, 2009.
10. IGN DEM,
<https://www.ign.gob.ar/NuestrasActividades/Geodesia/ModeloDigitalElevaciones/Introduccion>.
Last accessed 19 Jun 2021
11. OpenCV, https://docs.opencv.org/4.x/dc/dbb/tutorial_py_calibration.html. Last accessed 19 Jun 2021
12. Pyproj, <https://pypi.org/project/pyproj/>. Last accessed 19 Jun 2021
13. : Code of Federal Regulations (Annual Edition). Title 47: Telecommunication. 73 (208). October 1, (2016)
14. Cuaran, J., Leon, J. (2021).Crop monitoring using unmanned aerial vehicles: a review. Agricultural Reviews, 42(2), 121-132.
15. Sa, I., Chen, Z., Popović, M., Khanna, R., Liebisch, F., Nieto, J., Siegwart, R. (2017). weednet: Dense semantic weed classification using multispectral images and mav for smart farming. IEEE robotics and automation letters, 3(1), 588-595.