

Análisis de imágenes multiespectrales para la detección de cultivos y detección de plagas y enfermedades en la producción de café

Arely Guadalupe Sánchez-Méndez, Simón Pedro Arguijo-Hernández

Instituto Tecnológico Superior de Misantla,
Posgrado en Sistemas Computacionales, Veracruz,
México

{162t0085, sparguijoh}@itsm.edu.mx

Resumen. En la presente investigación se muestra un método para la detección de cultivos de café y de presencia de plagas y enfermedades en la producción de estos cultivos, utilizando imágenes multiespectrales del satélite Landsat 8. Como área de estudio se considera la zona localizada entre las regiones cafetaleras Misantla y Coatepec, en el estado de Veracruz. En el desarrollo del método se involucran los procesos de: pre-procesamiento de imágenes; interpretación digital a través de proceso de clasificación supervisada, implementando el análisis de componentes principales; muestreo en campo para entrenamiento y validación y evaluación de resultados. Para evaluar la precisión de estos datos se aplicó una matriz de error, mostrando un valor aceptable en la precisión general. Esto permite considerar que el uso de imágenes Landsat por su resolución espectral y radiométrica es aceptable para el estudio del cultivo de café según lo observado en campo y lo obtenido en el procesamiento de la imagen, ya que genera información sobre áreas cultivadas que apoyará en la toma de decisiones en la agricultura.

Palabras clave: Landsat 8, clasificación supervisada, procesamiento de imágenes, café.

Analysis of Multispectral Images for the Detection of Crops and Detection of Pests and Diseases in Coffee Production

Abstract. This research shows a method for the detection of coffee crops and the presence of pests and diseases in the production of these plants, using multispectral images from the Landsat 8 satellite. The study area is located between the coffee regions of Misantla and Coatepec, in the state of Veracruz. The development of the method involves the processes of: pre-processing of images of the area involved; digital interpretation through the process of supervised classification, implementation of principal component analysis; field sampling for training, validation and evaluation of results. An error matrix was applied to evaluate the accuracy of the data, showing an acceptable value in the

overall accuracy. This allows us to consider that the use of Landsat images for their spectral and radiometric resolution is acceptable for the study of coffee plantations based on what is observed in the field and what is obtained in image processing, since it generates information of the cultivation areas that are used for decision making in agricultural processes.

Keywords: Landsat 8, supervised classification, image processing, coffee.

1. Introducción

Veracruz es el segundo estado con mayor producción de café en la República Mexicana. Los cultivos de café en Veracruz cubren una superficie de 139,000 hectáreas, principalmente con las variedades de alta calidad de la especie *Coffea arábica* [2]. En Veracruz, las zonas cafetaleras se dividen en 10 regiones: en el norte del estado se encuentra Huayacocotla y Papantla; en el centro están Atzalan, Misantla, Coatepec, Huatusco, Córdoba, Zongolica y en el sur comprende Tezonapa y Los Tuxtlas.

En México el cultivo de café orgánico es afectado por la roya *Hemileia vastatrix*. La alta severidad de afectación se traduce en niveles incontrolables de infección, defoliación severa, pérdidas en la producción, debilidad y muerte de los cafetos [1]. El café se produce en América Latina, en África y en Asia, y es uno de los productos agrícolas más comercializados en los mercados internacionales. La agroindustria del café se ha diversificado en todo el mundo y constituye una importante fuente de empleo, ingresos y divisas en muchos países productores. En los años recientes, su oferta mundial ha sido afectada por factores climatológicos adversos y plagas como la roya, lo que se ha reflejado en la elevada volatilidad de las cotizaciones internacionales de este producto [10].

Con la implementación del presente proyecto se pretende ayudar a detectar a tiempo la infestación de los cafetales por plagas como la roya, brindando a los agricultores información sobre el estado de sus tierras y cultivos utilizando la combinación de imágenes multiespectrales, reconocimiento de patrones e inteligencia artificial.

La investigación se centrará en el análisis y procesamiento de imágenes multiespectrales para la detección de enfermedades y plagas en cultivos, esto a partir de que se ha demostrado que el análisis de imágenes satelitales representa una gran ayuda en el ámbito de la agricultura, debido a que se obtienen resultados significativos en comparación con el análisis de imágenes digitales que muestran únicamente el rango visible del espectro electromagnético [4]. Es importante mencionar que en las investigaciones analizadas no todas son respecto a las plagas y enfermedades en cultivos de café, aunque, en su mayoría, se enfocan en estos; los que no se enfocan en ello, son considerados debido a las técnicas utilizadas en el análisis de imágenes multiespectrales, no obstante, también consideran el análisis de la calidad de frutos.

Las tecnologías de información geográfica (TIG), incluyen la percepción remota, que actualmente utiliza imágenes de satélite multiespectrales que se aplican para estimar tipos de cubiertas sobre la superficie terrestre y la condición de las mismas [8]. De esta manera, se puede aplicar esta tecnología para identificar las superficies con plantaciones de café, diferenciar sus condiciones de desarrollo y discriminar aquellas superficies cuyas características en la imagen sugieran alguna restricción.

Dado que cualquier fenómeno que ocurra sobre la superficie de la tierra: vegetación, cultivos, cuerpos de agua, suelos, etc., se puede detectar y analizar mediante la tecnología de la percepción remota [3] es factible inferir la manifestación de la “salud” del cultivo o de su condición de desarrollo a través de la interpretación de imágenes de satélite multiespectrales.

Las imágenes multiespectrales son un arreglo de columnas y renglones que conforman una matriz de datos numéricos que representan la intensidad de la energía electromagnética reflejada o emitida por los objetos en la superficie de la Tierra. Las imágenes se pueden registrar en bandas individuales del espectro electromagnético, como ocurre en los satélites Landsat 8 que cuentan con 11 bandas espectrales; es decir, la misma escena capturada en diferentes bandas; banda 1, banda 2, banda 3, etc. [14].

Se utilizó la escena del satélite Landsat 8 adquirida el 4 de febrero del 2018; la cual fue descargada desde el visor de imágenes satelitales “Libra”. Se realizaron procesos de georreferenciación.

Con la finalidad de obtener información de las imágenes multiespectrales, se realiza un análisis mediante procesos digitales aplicando estadísticas de la imagen (clasificación supervisada) o mediante el análisis visual a través de un compuesto a color [5].

La presente investigación tomo por objetivo el análisis de imágenes multiespectrales satelitales con la finalidad de detectar cultivos de café además de plagas y enfermedades en los mismos.

2. Trabajos relacionados

Hacer uso de imágenes multiespectrales de alta resolución espectral y espacial, aplicando técnicas de procesamiento de imágenes y reconocimiento de patrones (visión computacional) en cultivos de café con la intención de obtener pronta detección específica de plagas en el que puedan poner en riesgo la producción de café y la seguridad de su cosecha [13].

Mahlein et al., relacionaron características foliares y reflectancia espectral de hojas de remolacha azucarera enfermas con mancha foliar de *Cercospora* y roya de hoja en diferentes etapas de desarrollo, utilizaron un espectrómetro de exploración de imágenes hiperespectrales (ImSpector V10E) con una resolución espectral de 2,8 nm de 400 a 1000 nm y una resolución espacial de 0,19 mm para la detección y monitorización continua de los síntomas de la enfermedad durante la patogénesis. No mencionan resultados cuantitativos, pero si concluyeron que la reflectancia espectral en combinación con la clasificación del mapeado del ángulo espectral permitió la diferenciación de los síntomas maduros en zonas que muestran todas las etapas ontogenéticas desde síntomas jóvenes hasta maduros [9].

Luaces et al., estudiaron el hecho de aprender funciones que fueran capaces de predecir si el valor de una variable objetivo continua puede ser mayor que un umbral dado. El objetivo de la publicación que estudiaron era alertar sobre la alta incidencia de la roya del café, principal enfermedad del cultivo de café en el mundo.

Realizan una comparación entre los resultados de su matriz de confusión, obteniendo resultados donde los costos de los falsos negativos son más altos que los de falsos positivos, y ambos son más altos que el costo de las predicciones de advertencia [11].

Huang et al., detectaron en los campos de trigo en temporadas de invierno, que éstos se encuentran afectados por la enfermedad llamada roya amarilla, la cual perjudica la producción de trigo, por lo cual, el objetivo de su estudio fue evaluar la exactitud del espectro óptico, el índice de reflectancia fotoquímica (PRI) para cuantificar el índice de dicha enfermedad y su aplicabilidad en la detección de la enfermedad mediante imágenes hiperespectrales.

Realizaron pruebas del PRI en tres temporadas, mostrando que, en invierno, con un porcentaje de determinación del 97%, el potencial del PRI es claro para cuantificar los niveles de óxido de color amarillo en el trigo y como base para el desarrollo de un sensor de imagen proximal de roya amarilla en los campos de trigo en invierno [16].

Devadas et al., evaluaron diez índices de vegetación ampliamente utilizados, basados en combinaciones matemáticas de mediciones de reflectancia óptica de banda estrecha en el rango de longitudes de onda visible e infrarrojo cercano por su capacidad para discriminar hojas de plantas de trigo de un mes infectadas con rayas amarillas.

No mencionan resultados cuantificables pero concluyen que ningún índice individual fue capaz de discriminar las tres especies de óxido entre sí, sin embargo, la aplicación secuencial del Índice de Reflectancia Antocianina para separar las clases sanas, amarillas y mixtas de óxido y roya de las hojas seguidas por el índice de absorción de la clorofila y el índice de reflectancia para separar las clases de roya de hojas y tallos, podrían constituir la base de la discriminación de las especies de óxido en el trigo en condiciones de campo [12].

Stefan et al., observaron la interacción planta-patógeno mediante mediciones simultáneas de reflexión y transmisión de imágenes hiperespectrales. Estos datos se analizaron estadísticamente utilizando el análisis de componentes principales, y se comparó con la estimación de la enfermedad visual y molecular, concluyendo que las mediciones basadas en reflectancia facilitan una detección temprana, y las mediciones de transmisión proporcionan información adicional para comprender y cuantificar mejor la compleja dinámica espacio-temporal de las interacciones planta-patógeno [15].

3. Materiales y métodos

3.1. Área de estudio

El área de influencia de la producción del café, es de carácter interestatal y las superficies ocupadas por café se distribuyen de manera dispersa, aunque en ciertas localidades tiene una mayor concentración.

Esta investigación se llevó a cabo en los municipios Misantla, zona de cultivo y producción del café. El área de trabajo se distribuye dentro de las regiones cafetaleras Misantla y Coatepec [10].

En este polígono es dónde se localiza una importante de concentración de cultivos de café, para estos municipios.

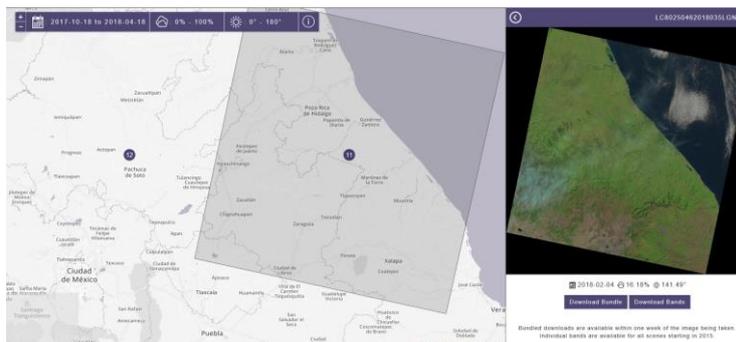


Fig. 1. Interfaz principal de Libra una vez situado el mapa sobre la zona de Veracruz. Se ofrece filtro de búsqueda por fecha, cobertura de nubes o ángulo solar (sobre el mapa). También podemos ordenar las imágenes disponibles por fecha, cobertura de nubes o ángulo solar (sobre la vista de resultados).

3.2. Imágenes empleadas

Para el reconocimiento por percepción remota de factores que restringen el desarrollo del cultivo del café, se utilizó una escena del satélite Landsat 8 del 04 de febrero de 2018, con un 16.18% de nubosidad y una inclinación del son de 141.49. Esta escena ha sido descargada del visor de imágenes satelitales Landsat 8, “Libra” (ver Fig. 1).

El satélite Landsat 8 transporta dos instrumentos OLI y TIRS, que corresponden a las siglas en inglés para Operational Land Imager (OLI) y Thermal Infrared Sensor (TIRS). El sensor OLI provee acceso a nueve bandas espectrales que cubren el espectro desde los 0.433 micrómetros a los 1.390 micrómetros, mientras que TIRS registra de 10.30 micrómetros a 12.50 micrómetros. Por esto, la escena está compuesta por 11 imágenes, cada una es una banda de la escena con su resolución espectral correspondiente. Con ello es posible capturar la radiación proveniente de la superficie terrestre en once bandas espectrales, cada una para registrar características de objetos en la superficie: suelos, vegetación, agua, etc. [7].

3.3. Muestro en campo para entrenamiento y validación

Las imágenes fueron georreferenciadas mediante el uso de Google Earth mediante los valores de la imagen Landsat 8 (ver Fig. 2).

3.4. Combinación de bandas

Se realizó diferentes combinaciones de bandas con la finalidad de analizar la combinación más adecuada para realizar la detección de los cultivos e identificación de enfermedades en los cultivos (ver Fig. 3). Las combinaciones más relevantes realizadas en el procesamiento de imágenes satelitales del Landsat 8, son mostradas en la Tabla 1, donde se identifican los diversos usos que se les da a cada combinación de bandas.

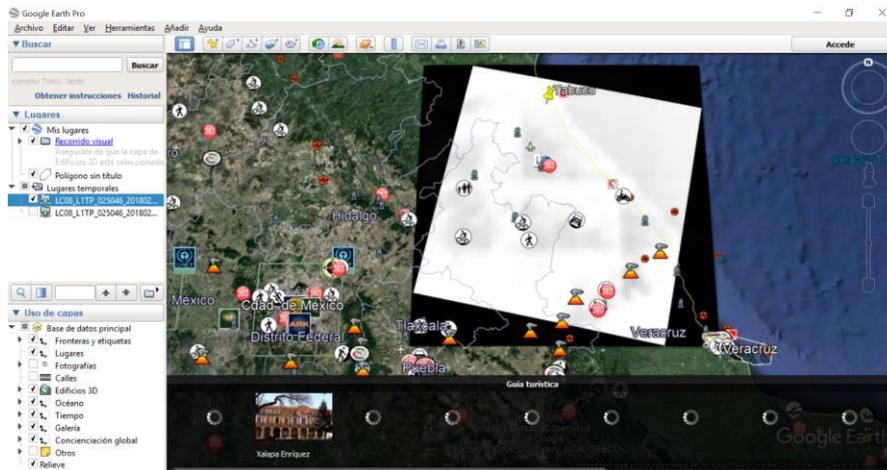


Fig. 2. La imagen es posicionada en el mapa geográfico, considerando que contempla parte de los estados de Veracruz, Tlaxcala, Puebla e Hidalgo.

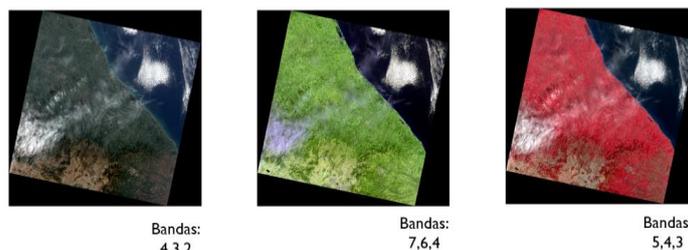


Fig. 3. La imagen es posicionada en el mapa geográfico, considerando que contempla parte de los estados de Veracruz, Tlaxcala, Puebla e Hidalgo.

Tabla 1. Combinaciones de bandas del satélite Landsat 8.

Uso	Combinación de bandas
Color natural	4, 3, 2
Falso color (urbano)	7, 6, 4
Color infrarrojo (vegetación)	5, 4, 3
Agricultura	6, 5, 2
Penetración atmosférica	7, 6, 5
Vegetación saludable	5, 6, 2
Tierra / Agua	5, 6, 4
Natural con remoción atmosférica	7, 5, 3
Infrarrojo de onda corta	7, 5, 4
Análisis de vegetación	6, 5, 4

3.5. Cálculo del índice de vegetación de diferencia normalizada (NDVI)

Los valores del NDVI varían entre -1 y 1, donde el cero corresponde a un valor aproximado de no vegetación [6]. Valores negativos representan superficies sin vegetación, mientras valores cercanos a 1 contienen vegetación densa. El NDVI se calcula mediante la ecuación (1):

$$NDVI = \frac{(NIR - R)}{(NIR + R)}, \quad (1)$$

donde R y NIR se refieren a los valores de reflectancia medidos por las bandas del rojo (R) e infrarrojo cercano (NIR).

3.6. Segmentación y clasificación de las imágenes

La identificación de los endmembers pertenecientes a los cultivos de café georreferenciados en la escena requirió un análisis de los componentes principales (ACP) con el fin de eliminar la redundancia propia de los datos utilizados. La naturaleza multiespectral o multidimensional de las imágenes puede ajustarse mediante la reconstrucción de un espacio vectorial con un número de ejes o dimensiones igual al número de componentes asociados con cada pixel.

Esta transformación genera un conjunto de bandas que corresponden a cada valor propio y son organizadas de acuerdo a la estimación del ruido en las imágenes multiespectrales; por esta razón, para obtener un resultado confiable en la caracterización de los perfiles espectrales de los cultivos es necesario remover dicho ruido de la imagen; sin embargo, es necesario evitar la pérdida de datos, tanto como sea posible.

4. Experimentos y resultados

4.1. Valores del NDVI

De acuerdo a los valores del NDVI, la vegetación verde fotosintéticamente activa se encuentra entre 0,2 a 0,8 y los cultivos tienden a estar entre 0,4 y 0,9 dependiendo en gran parte del índice de área foliar y de la disposición en el terreno. Igualmente, en el NDVI influye el porcentaje de cobertura del suelo, y se presenta la mejor correlación cuando la cobertura está entre el 25% y el 80%. Los valores bajos presentados por los pastos se deben posiblemente a poca cobertura, por debajo del 15%, en cuyo caso el NDVI no indica con precisión el grado de biomasa de la vegetación, ya que está afectado por la reflectancia del suelo desnudo [5]. En la Fig. 4, se observa la manera en que los valores de reflectancia de cada tipo de cobertura se comportan en las 11 bandas que componen la escena del satélite Landsat 8.

Se observan claramente 3 grupos. Cada pixel está representado por su valor de reflectancia en cada banda, por lo que cada pixel se representa en 11 puntos (azul, verde, amarillo, púrpura, café, naranja, negro, blanco, cian, magenta, rojo, y vino), cada color representa una banda.

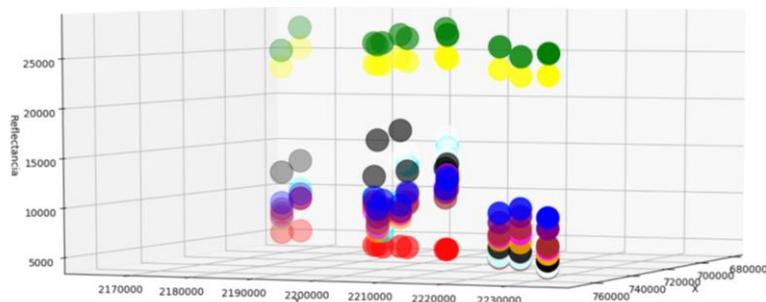


Fig. 4. Selección de 14 píxeles elegidos de la imagen satelital que involucran 3 tipos de cobertura (vegetación, zonas urbanas y cuerpos de agua). Se observa el comportamiento en cada una de las bandas de la escena del satélite Landsat 8.

5. Conclusiones y trabajos futuros

Utilizando las imágenes multiespectrales del área que corresponde una región cafetal importante del estado de Veracruz y realizando el procesamiento y análisis de la escena, se realiza el proceso marcado por el método propuesto como una alternativa confiable para la identificación de cultivos de café, y en consecuencia se debe establecer su utilidad en otros cultivos y en diversas condiciones ambientales. Si se tiene completa disponibilidad de las imágenes multiespectrales, se puede explorar el uso de aquellas tomadas de diferente estado fenológico para caracterizar mejor el comportamiento espectral del cultivo. Es importante mencionar que, considerando la resolución espacial de las imágenes empleadas, los resultados e información obtenida es aceptable, aunque si es posible acceder a imágenes satelitales con mayor resolución espacial y espectral los resultados serían mejores en gran medida.

El avance de la investigación presentada muestra una parte de los resultados que se esperan obtener implementando todo el proceso del método propuesto, aunque estos resultados preliminares muestran altas posibilidades para la detección adecuada de los cultivos de café y algunas restricciones en la producción, tales como enfermedades y/o plagas.

Además, como trabajo futuro se planea evaluar por medio de diferentes algoritmos de reconocimiento de patrones (redes neuronales, árboles de decisión, etc.), las mismas imágenes. Con la finalidad de obtener el algoritmo que dé una mayor precisión en clasificación de los cultivos.

Referencias

1. Chemura, A., Mutanga, O., Dube, T.: Separability of coffee leaf rust infection levels with machine learning methods at sentinel-2 msi spectral resolutions. *Precision Agriculture*, p. 23 (2016)

2. PRONATURA Veracruz.: Regiones cafetaleras de Veracruz. Aroma de la Biodiversidad. Pronatura México A.C. (2017)
3. Boken, V. K., Cracknell, A. P., Heathcote, R. L.: Oxford University Press, Cary, NC (2005)
4. Castro, W. M.: Aplicación de la tecnología de imágenes hiperespectrales al control de calidad de productos agroalimentarios de la región de Amazonas (Perú). Master's thesis, Universidad Politécnica de Valencia, Instituto Universitario de Ingeniería de Alimentos para el Desarrollo (2015)
5. García, S. A., Martínez, L. J.: Método para identificación de cultivos de arroz (*Oryza sativa* L.) con base en imágenes de satélite. *Agronomía Colombiana* 28(2), pp. 281–290 (2010)
6. Jensen, J. R.: *Introductory Digital Image Processing: A Remote Sensing Perspective*. Prentice-Hall, Englewood Cliffs (2005)
7. Landgrebe, D.: Hyperspectral image data analysis. *IEEE Signal Processing Magazine* 19, pp. 17–28 (2002)
8. Lencinas, J. D., Mohr-Bell, D.: Estimación de clases de edad de las plantaciones de la provincia de Corrientes, Argentina, con base en datos satelitales Landsat. *Bosque* 28(2), 106–118 (2007)
9. Mahlein, A. K., Steiner, U., Hillnhutter, C., Dehne, H. W., Oerke, E. C.: Hyperspectral imaging for small-scale analysis of symptoms caused by different sugar beet diseases. *Plant Methods* 8, pp. 3 (2012)
10. Morales, A. C. M.: Retos del productor cafetalero frente al contexto económico y político, en la región de Coatepec. In: XIII Congreso Mundial de Sociología Rural de la Asociación Internacional de Sociología Rural (IRSA) (2012)
11. Luaces, O., Antunes, L. H., Meira, C., Bahamonde, A.: Using nondeterministic learners to alert on coffee rust disease. *Expert Systems with Applications* 38, pp. 14276–14283 (2011)
12. Devadas, R., Lamb, D. W., Backhouse, D., Simpfendorfer, S.: Evaluating ten spectral vegetation indices for identifying rust infection in individual wheat leaves. *Precision Agriculture* 10, pp. 459–470 (2009)
13. Roman-Gonzalez, A., Vargas-Cuentas, N. I.: Análisis de imágenes hiperespectrales. *Revista Ingeniería Desarrollo* 9(35), pp. 14–17 (2013)
14. Smith, R.B.: *Remote Sensing of Environment (RSE)*. MicroImages (2012)
15. Thomas S., Wahabzada, M., Kuska, M. T., Rascher, U., Mahlein, A. K.: Observation of plant–pathogen interaction by simultaneous hyperspectral imaging reflection and transmission measurements. *Functional Plant Biology* 44, pp. 23–34 (2016)
16. Huang, W., Lamb, D.W., Niu, Z., Zhang, Y., Liu, L., Wang, J.: Identification of yellow rust in wheat using in-situ spectral reflectance measurements and airborne hyperspectral imaging. *Precision Agriculture* 8(4-5), pp. 187–197 (2007)