

Análisis de imágenes para el estudio de soja

Mónica G. Larese, Rafael Namías, Pablo M. Granitto

Unidad de Investigación en Imágenes

Centro Internacional Franco-Argentino de Cs. de la Información y de Sistemas
(CIFASIS-CONICET-UNR)

Bv. 27 de Febrero 210 Bis, S2000EZP Rosario, Argentina

(0341) 423 7248 int. 303

{larese,namias,granitto}@cifasis-conicet.gov.ar

Resumen

La línea de investigación presentada surge de dos motivaciones paralelas. La primera consiste en desarrollar una metodología y un software para identificar objetivamente, a través del análisis de imágenes digitales, la coloración verde de granos/semillas de soja como indicador de la cantidad de clorofila presente en los mismos. El nivel de clorofila es un factor influyente en la calidad fisiológica de las semillas para su utilización en los procesos industriales (extracción de aceites) así como también de su capacidad fértil para la siembra. Actualmente el análisis se realiza de manera subjetiva, o bien requiere de la utilización de costosos equipamientos de laboratorio, por lo cual se requiere de nuevos métodos que permitan realizar el procedimiento de manera rápida, confiable y económica. La segunda motivación tiene por objetivo desarrollar un sistema de reconocimiento automático de imágenes foliares de soja que permita la clasificación e identificación de las distintas variedades de dicho cultivo. Debido a que no existen diferencias evidentes a simple vista en cuanto a la forma, color y textura de las hojas para las distintas variedades, interesa realizar el análisis detallado y automático de la estructura del sistema de nervaduras buscando identificar posibles marcadores que permitan diferenciar las variedades.

Palabras clave: Análisis de imágenes, soja,

granos, semillas, hojas, nervaduras, extracción de características, clasificación, reconocimiento de patrones.

Contexto

Las tareas de investigación se realizan de manera conjunta entre el Instituto Nacional de Tecnología Agropecuaria (INTA) Estación Experimental Oliveros, y el CIFASIS (CONICET-UNR). INTA participa brindando el asesoramiento técnico pertinente al problema de aplicación, así como también mediante la disponibilidad de equipamiento específico (equipo de video, ambiente de germinación, instrumentales de laboratorio, equipamiento de transporte y almacenaje). Por su parte, CIFASIS dispone de equipos fotográficos digitales, clusters de computadoras para el procesamiento y sistemas de calibración de color.

La investigación referente a la identificación de la coloración en granos/semillas se encuentra inserta en un proyecto conjunto entre ambas instituciones, el que se denomina “Desarrollo de un equipo medidor de intensidad de color verde en granos y semillas de soja”. Asimismo, el análisis de imágenes foliares se realiza en el marco de un plan de trabajo de beca posdoctoral de CONICET, denominado “Reconocimiento automático de especies y variedades vegetales

agrícolas en Argentina a partir de imágenes de nervaduras”.

Introducción

El cultivo de soja en la República Argentina ha tenido en los últimos años una gran expansión territorial que fue acompañada, además, por un importante incremento en el rendimiento. El gran crecimiento del cultivo fue consecuencia de una situación favorable de precios internacionales, acompañada de la difusión de nuevos cultivares con altos rendimientos potenciales para zonas de producción con diversas condiciones ambientales. Sin embargo, los granos/semillas de soja no siempre se producen bajo condiciones ambientales favorables, viéndose afectados negativamente los rendimientos y/o la calidad de la producción cuando se manifiestan diversos tipos de estrés ambiental durante el ciclo total del cultivo [2].

Las temperaturas superiores a 32° C con continuidad de horas, situaciones de déficit hídrico prolongado y/o el manejo inadecuado del lote de producción generan condiciones estresantes que provocan un desbalance fisiológico seguido de la muerte prematura de las plantas [8]. Este desbalance fisiológico es un factor determinante para la aparición de un fenómeno denominado como problemática de producción de granos y semillas verdes. La coloración verde es consecuencia de la permanencia de pigmentos clorofílicos en distintos tejidos vegetales [2].

Este problema se ha convertido en una característica negativa de los lotes, tanto para la siembra como para la industrialización. Las consecuencias para la industria aceitera son de enorme importancia económica, puesto que el proceso de producción sufre una serie de perturbaciones debido a la presencia de estos granos verdes en el granel. Los granos verdes poseen un mayor contenido de humedad y una consistencia más elástica que dificulta el proceso de extracción de aceites. Por otro lado, el aceite obtenido de lotes con granos verdes presenta un alto

contenido de clorofila, por encima de los estándares aceptados para la comercialización internacional [1], transmitiendo una coloración verde al aceite similar al aceite de oliva. Como consecuencia de ello se encarece el proceso industrial debido a la necesidad de clarificar el aceite. Otro punto importante a tener en cuenta es que los aceites con altos contenidos de clorofila son más sensibles al enranciamiento.

Otros inconvenientes causados por la presencia de altos niveles de granos verdes en el proceso industrial incluyen un mayor desgaste mecánico de los equipos; aumento de mano de obra de mantenimiento y control; aumento de espacio de almacenamiento para productos fuera de especificación y un incremento del costo financiero debido a la existencia de productos fuera de especificación en tanques de stock.

Si el destino de la producción es la siembra, la ocurrencia de semillas verdes en el lote de simientes puede afectar negativamente la calidad del mismo y dificultar la implantación del cultivo obligando a realizar resiembras.

El valor de uso de semillas que poseen un tinte verdoso para la siembra depende del estado de madurez alcanzado por el embrión. Si el embrión alcanzó su madurez fisiológica, entonces existen posibilidades de que se desencadene la germinación y produzca una plántula normal, aún cuando los cotiledones conserven cierto grado de inmadurez. Estas semillas presentan un grave inconveniente durante el acondicionamiento ya que poseen forma y tamaño semejante a las semillas maduras y por lo tanto, no pueden separarse del resto mediante una clasificación normal, desmejorando el aspecto general del lote [1].

Existen métodos que facilitan la observación e identificación del color que podrían aplicarse en semillas tales como: Espectro Análisis, Método de Reflectancia por Infrarrojo cercano (NIR) y métodos que emplean equipamientos especiales como Colorímetros y Espectrofotómetros [5]. El elevado costo de los equipamientos mencionados, sumado a la inminente necesidad de los analistas

de identificar de forma rápida y precisa las diferentes intensidades de color verde en los granos y semillas de soja, pone de manifiesto la necesidad de contar con una metodología sencilla y económica que brinde soluciones prácticas a los analistas.

Una solución rápida y económica consiste en el desarrollo de un equipo y un método de software que facilite la identificación y clasificación de los granos y semillas verdes presentes en un lote según las distintas tonalidades de color verde y que a su vez relacione dicha característica del color con parámetros de calidad física y fisiológica.

Por otra parte, el interés en realizar el análisis de hojas de soja se origina en que, en la actualidad, los especialistas agrónomos realizan la clasificación de las plantas agrícolas mediante el análisis manual de las hojas. Además de requerir más tiempo de procesamiento, el análisis manual es subjetivo, tedioso y más propenso a errores. En otras situaciones, las características distintivas entre especies o incluso entre variedades de la misma especie) no son visibles a simple vista, o son desconocidas *a priori* por los expertos. Éste es el caso de la planta de soja, para la cual los especialistas aún no han encontrado los posibles rasgos distintivos visibles en las hojas que permitan diferenciar las distintas variedades dentro de la misma especie.

Las hojas son los órganos vegetales más utilizados para la identificación de plantas. Esto se debe a su característica exógena y sus propiedades que pueden ser observadas fácilmente a nivel macroscópico. Por lo general, distintas especies vegetales poseen hojas con características diferentes.

Los rasgos distintivos que presentan las hojas incluyen su color y textura. Sin embargo, para la identificación de plantas no se pueden considerar únicamente estos rasgos debido a que pueden ser compartidos por otras especies vegetales. La consideración, además, de la forma y tamaño de las hojas permite acotar un poco más las posibilidades de clasificación. Más aún, analizar la estructura del sistema de nervaduras podría ayudar notablemente a

diferenciar las especies.

En el caso de la planta de soja, resulta de particular interés para los especialistas disponer de un método que permita distinguir los individuos correspondientes a las distintas variedades mediante el análisis digital automático de las hojas. Debido a que no existen diferencias evidentes a simple vista en cuanto a la forma, color y textura de las hojas, es de interés realizar el análisis detallado de la estructura del sistema de nervaduras, buscando identificar posibles marcadores que podrían permitir desarrollar un sistema de reconocimiento automático de variedades de soja basado en esta información foliar.

En el ambiente agrícola frecuentemente se requiere recolectar muestras para su análisis posterior en laboratorio. Dados los actuales avances tecnológicos, sin embargo, es posible que algunos de estos análisis puedan realizarse directamente durante el trabajo de campo, permitiendo obtener resultados más rápidamente y brindando mayores facilidades a los especialistas. Si bien hoy en día existen técnicas de identificación más avanzadas, tales como el análisis del ADN (Ácido Desoxirribonucleico), éstas técnicas son mucho más costosas y además requieren de equipamiento especial del cual no se puede disponer fácilmente al aire libre. Por el contrario, el desarrollo de un sistema de identificación foliar automático requeriría contar solamente con un PDA (*Personal Digital Assistant*) con conexión a Internet.

Líneas de investigación y desarrollo

La línea de investigación cuenta con dos ejes principales, los cuales tienen por objetivo principal, respectivamente:

- Desarrollar una metodología y un software que permita identificar objetivamente, a través del análisis de imágenes digitales, la coloración verde de granos/semillas de soja como indicador de la cantidad de clorofila presente en los mismos.

- Desarrollar un sistema de reconocimiento automático de imágenes foliares de soja que permita la clasificación e identificación de las distintas variedades, utilizando para ello información contenida en las nervaduras.

La metodología de trabajo propuesta para analizar la coloración verde en granos/semillas incluye tareas de campo y de laboratorio a cargo de INTA (recolección de semillas, separación y clasificación de semillas verdes, creación de una escala color de referencia, etc.). Esto permitirá armar muestras artificiales que contengan distintos porcentajes de semillas verdes con diferentes niveles de intensidad de coloración para poner a prueba el equipamiento.

CIFASIS se ocupará de desarrollar un método analítico que permita estandarizar color. Se realizará la detección y segmentación de semillas, y se evaluará dentro de cada una la proporción de cada uno de los colores individualizados. Cada semilla se clasificará, luego, a partir de la información de color recolectada. Para el procesamiento se utilizarán métodos de *machine learning* junto con algoritmos de procesamiento digital de imágenes.

Para la identificación de variedades de soja a partir del análisis de nervaduras de las hojas también se trabajará con una base de datos provista por INTA, la cual está actualmente constituida por imágenes color de hojas de soja de diferentes cultivares. Cada cultivar corresponde a una variedad distinta dentro de la misma especie.

Como primera medida se realizará el pre-procesamiento de las imágenes para ajuste de contraste y filtrado de ruido [3, 7]. Luego se llevará a cabo la segmentación, y se utilizarán técnicas de procesamiento y análisis de imágenes para extraer y representar eficientemente la estructura de nervaduras. Las nervaduras de las hojas presentan una gran variabilidad intra-clase (variedades de soja distintas) y una pequeña variabilidad inter-clase (individuos de la misma variedad).

Sobre la representación obtenida se utilizarán técnicas de aprendizaje automatizado que permitan extraer las características más relevantes y lograr un nivel de separación óptima entre las clases consideradas. Estas técnicas incluyen métodos de selección de variables, clustering y clasificación [4].

Resultados y Objetivos

Con el desarrollo del equipo medidor de intensidad de color verde en lotes de soja se espera:

- Lograr un equipamiento y un software que sea una herramienta útil para el Control de Calidad de granos y semillas.
- Que el equipamiento permita cuantificar objetivamente la presencia de granos y semillas con diferentes tonalidades de color verde en lotes de soja.
- Brindar una estimación de la calidad fisiológica de las semillas verdes de soja.

En cuanto al análisis de imágenes de hojas y nervaduras, los objetivos planteados son los siguientes:

- Desarrollar un sistema de reconocimiento automático que permita, mediante técnicas de análisis de imágenes y aprendizaje automatizado, la identificación de las hojas de soja.
- Identificar el conjunto de características (*features*) pertenecientes a las hojas que sean más relevantes para este propósito, incluyendo el análisis de la estructura del sistema de nervaduras.
- Evaluar cuantitativamente el desempeño del método desarrollado sobre imágenes reales de hojas de soja.

Formación de Recursos Humanos

Con referencia a estos proyectos, se encuentran en curso una tesis doctoral y un trabajo de investigación posdoctoral.

El grupo de trabajo se encuentra compuesto por los siguientes integrantes:

Por parte de INTA Oliveros:

- Ing. Agr. MSc. Miriam Arango (Director)
- Ing. Agr. MSc. Carina Gallo (Co-Director)
- Ing. Agr. MSc. PhD Roque Craviotto (Co-Director)
- Marcelo Fared (Colaborador)
- Marta Montero (Colaborador)

Por parte de CIFASIS:

- Dr. Pablo Granitto (Director)
- Lic. Rafael Namías (Beca Doctoral ANPCyT, en curso)
- Dr. Mónica Larese (Beca Posdoctoral CONICET, en curso)

Referencias

- [1] R.M. Craviotto and M.R. Arango. Calidad de semilla 2001. odisea de la soja. *Revista APOSGRAN*, 2001.
- [2] C. Gallo. Calidad fisiológica y efecto de la presencia de semillas verde de soja (*Gycine max* (L.) Merrill) en lotes destinados a simiente. *Maestría en Cs. Agrop. con Mención en Tecnol. de Semillas, Fac. de Cs. Agrop., Univ. Nac. de Cba., Argentina*, 2008.
- [3] R. Gonzalez and R. Woods. *Digital image processing*. Prentice Hall, 2nd. edition, 2002.
- [4] T. Hastie, Tibshirani R., and J. Friedman. *The Elements of Statistical Learning, Second Edition*. Springer, 2009.
- [5] L. Guiping L. Jinwei and X. Fen. Rapeseed seeds colour recognition by machine vision. 2008.
- [6] H.-C. Lee. *Introduction to Color Imaging Science*. 2005.
- [7] S.E. Umbaugh. *Computer imaging: digital image analysis and processing*. CRC Press, 2005.
- [8] B. Wiebold. Soybean plants killed before maturity possess grain that remains green. (<http://www.plantsci.missouri.edu/-soyx/-soyfacts/green.PDF>), 2002.