

Desarrollo de un método de detección de Sigatoka negra utilizando atributos intrínsecos del huésped y el anfitrión por medio de técnicas de visión por computadora

“Development of a method for Black Sigatoka detection using intrinsic attributes from host and guest through computer vision techniques”

Problema

Uno de los riesgos más importantes que sufre el cultivo del banano en costa rica, cuyas exportaciones representaron para el 2019 un monto de 1000 millones de dólares anuales, es el del hongo que produce la enfermedad conocida como Sigatoka negra.

El control constante de esta enfermedad requiere conocimientos especializados, y de inspecciones visuales frecuentes en los campos de producción.

El nivel de afectación de Sigatoka negra se determina en niveles que van de 0 a 5, donde cero significa la ausencia de la enfermedad y cinco su mayor nivel de afectación en el que es imposible salvar el cultivo.



El síntoma para determinar la presencia de la enfermedad es la presencia de manchas que empiezan con tamaños muy pequeños de uno o dos milímetros y de un color blanco. Estas manchas van creciendo y cambiando de color hacia tonos de marrón, y finalmente negros para los niveles 4 y 5.

Una de las características para determinar el nivel de infección, en especial de los niveles 1 al 3, es la cantidad de manchas en un área determinada.

Para el estudio de la enfermedad y la efectividad de fungicidas es importante tener una medición del nivel de enfermedad de la forma más precisa posible.

El desarrollo de la investigación trata entonces de identificar las áreas exactas donde la enfermedad se encuentra presente, para habilitar en un trabajo futuro el conteo de las manchas. Lograr estas metas de una forma automatizada aseguraría una mayor precisión en la determinación del nivel de la enfermedad, dadas las características determinísticas de los procesos automatizados.

Solución

En términos generales lo que se desea es poder identificar las áreas en una hoja que presentan Sigatoka negra a partir de imágenes tomadas en el campo.

Visión por computadoras

Para atacar el problema se utilizaron técnicas de visión por computadora. Estas técnicas tienen que ver con la obtención de imágenes, preprocesamiento, extracción de características, segmentación, procesamiento de alto nivel, y toma de decisiones.

Imágenes, análisis y pruebas

Se tomaron fotografías de varias plantas enfermas en niveles del 1 al 3 en los campos de laboratorio de la Corporación Bananera Nacional, CORBANA.

Estas fotografías fueron editadas una a una de forma manual con la ayuda de experto en la identificación de la enfermedad. En estas imágenes se marcaron (pintaron) las áreas enfermas de la hoja. Las imágenes fueron luego procesadas y se eliminaron las áreas no marcadas. Esto dejó, entonces, una imagen llamada binaria pues solo tiene dos colores, negro para las áreas que fueron marcadas por el experto, y blanco, para todo lo demás.

El siguiente paso fue el análisis de las secciones enfermas. Para esto, se toma una imagen binaria y la misma imagen original. De la imagen original, se toman los píxeles que se encuentran en las mismas ubicaciones que las áreas marcadas en la imagen binaria. Estos son los píxeles de la imagen donde está la enfermedad.

Para una imagen, el número de estas secciones puede rondar el orden de las centenas.

Este mismo trabajo se realizó en todas las imágenes tomadas.

Una vez coleccionados todos los píxeles enfermos se hace un análisis de las características de dichos píxeles. Este análisis se hizo en el espacio de colores RGB.

En análisis de las características extraídas permitió desarrollar una fórmula que se utilizó para determinar si un determinado píxel correspondía a Sigatoka negra o no. Dicha fórmula toma en cuenta los cambios de los valores del color en los canales individuales (R, G, ó B). A cada canal se le da un peso diferente de aporte de información para determinar si el píxel corresponde a la enfermedad o no.

Procesamiento de imágenes

Es importante mencionar que la fórmula desarrollada necesita de una serie de variables con valores de entrada iniciales heurísticos. Cada una de estas variables podía tener valor en determinados rangos. Esto hizo que en la fase de experimentación se debieran de correr las pruebas en todas las imágenes con cada configuración (valores de entrada) para la fórmula. Esto llevó a la ejecución del algoritmo más de medio millón de veces. La corrida de los experimentos fue únicamente posible gracias a la disposición casi exclusiva del clúster *Kabré* del Centro Nacional de Alta Tecnología y su súper poder de procesamiento paralelo.

Análisis de resultados comparación con SCM

Los resultados del algoritmo desarrollado fueron comparados con los resultados obtenidos de otro método probabilístico que se desarrolló con fines similares llamado SCM, por sus siglas en inglés de *Statistical Color Model*. El algoritmo desarrollado, IABSEG por sus siglas en inglés de *Intrinsic-Attributes-Based SEGmentation*, presentó una mejor exactitud al compararlos con varias métricas ($F_{0.5}$, F_1 , F_2 , y MCC).

Agradecimientos

CONICIT

PhD. Luis Alexander Calvo Valverde

PhD. Pablo Alvarado Valverde

CORBANA

Centro Nacional de Alta Tecnología