

IDENTIFICACIÓN Y CLASIFICACIÓN DE FRUTOS DE CAFÉ EN TIEMPO REAL, A TRAVÉS DE LA MEDICIÓN DE COLOR

Paula Jimena Ramos Giraldo*, Juan Rodrigo Sanz Uribe*, Carlos Eugenio Oliveros Tascón*

RESUMEN

RAMOS G., P.J.; SANZ U., J.R.; OLIVEROS T., C.E. Identificación y clasificación de frutos de café en tiempo real a través de la medición de color. *Cenicafé*, 61(4):315-326. 2010

Este artículo presenta un método para identificar cuatro estados de maduración de los frutos de café y su implementación en un circuito electrónico. El método se basa en la representación del color Hue-Saturation-Value (HSV), particularmente en las diferencias de los estados de maduración en el plano de las componentes "hue"(Matiz) y "saturation"(Saturación) (*hs*), y un algoritmo de aprendizaje que entrega puntos de corte para limitar las áreas de los cuatro estados de maduración sobre el plano *hs*. El algoritmo usa un proceso denominado Optimización de la Eficacia Global (OEG), para identificar las áreas que contienen el máximo número de lecturas de un estado de maduración, con el mínimo efecto de éstos en áreas vecinas. El sistema tiene dos funciones, una de calibración que usa el algoritmo para determinar el área correspondiente a cada estado de maduración sobre el plano *hs*, y otra de clasificación, que tiene la capacidad de identificar y separar hasta 50 frutos/s en tiempo real. El sistema tolera variaciones en la distancia de los frutos a la lente, humedad de los frutos y cambios suaves en la potencia luminosa de la fuente de luz. La eficacia de identificación varía entre 94,8% y 99,6%, lo cual hace que el dispositivo sea confiable para el control de calidad de la materia prima a procesar en el beneficio del café.

Palabras clave: Optoelectrónica, visión artificial, clasificador lineal, estados de maduración.

ABSTRACT

This paper presents a novel method and a hardware to identify coffee fruits in four different maturation stages. This method is based in the Hue-Saturation-Value (HSV) color representation, more particularly in differences in the H-S plane and it uses a learning algorithm that gives the cutoffs that limit four different areas in the H-S plane for every maturation stage. In order to obtain the limits, the algorithm uses an optimization process, named Optimized Global Efficacy (OGE), to identify the maximum number of fruits in the right area. The hardware has two functions; a calibration function that uses the mentioned algorithm to determine the areas in the H-S plane and the classification function that has the ability to identify and separated 50 fruits per second in real time. The system tolerates variations in the distance to the target, humid environment, wet fruits and smooth changes in illumination. The efficacy of the hardware varied from 94,8% up to 99,6%, which makes this device reliable for the propose that was developed.

Keywords: Optoelectronics, machine vision , classifier linear, maturation stages

* Investigador Científico I, Investigador Científico II e Investigador Principal, respectivamente. Ingeniería Agrícola. Centro Nacional de Investigaciones de Café, Cenicafé. Manizales, Caldas, Colombia.

El café es uno de los principales productos de exportación en Colombia, del cual dependen cerca de dos millones de personas. A nivel mundial, el café colombiano es reconocido como el café más suave del mercado, debido a una combinación de condiciones como la variedad, el suelo, el clima, las prácticas agrícolas, la recolección selectiva y el procesamiento.

Roa *et al.* (9) definen un fruto de café como una baya drupácea, conformada por la pulpa, epidermis o exocarpio, y el mucílago o mesocarpio, que es una sustancia azucarada y gelatinosa, que forma una capa de 0,5 mm de espesor promedio, la cual recubre los dos granos enfrentados por su cara plana (Figura 1). Las almendras o granos de café, de los cuales se extrae la bebida, se encuentran recubiertos por dos capas, una llamada pergamino o endocarpio, la cual es dura y frágil, y la segunda, una membrana fina llamada espermodermo o película plateada.

El color de la epidermis o pulpa varía de verde a rojo, por la aparición de sustancias como la clorofila y las antocianinas, respectivamente, las tonalidades presentes sobre la epidermis del fruto indican directamente el grado de maduración (4). Después de la floración, el café tarda en promedio 32 semanas para alcanzar el estado óptimo de maduración. El proceso de desarrollo de los frutos de café está dividido en tres etapas según Salazar *et al.* (10). En la primera etapa, que va desde la floración hasta la semana 8, el crecimiento es muy lento, los frutos verdes y se asemejan a la cabeza de un fósforo. En la segunda etapa, comprendida entre la semana 9 y la 26, el crecimiento es más rápido, los frutos empiezan a ganar peso y su color permanece verde presentando cambio de tonalidades. La tercera etapa, que va desde la semana 27 a la 32, el fruto cambia de color verde a rojo y adquiere su madurez fisiológica, estos frutos son aptos para ser cosechados.

Por último, se encuentra la postmaduración, etapa que se presenta si no se cosechan los frutos oportunamente (Figura 2).

Según Ramos *et al.* (8), durante el proceso de maduración los frutos de café experimentan un cambio de color correspondiente al cambio del espectro visible de 551,5 nm, cuando el fruto es inmaduro, a 616,18 nm, cuando el fruto es maduro, y se definen otras longitudes de onda cercanas a las mencionadas, para los estados de desarrollo intermedios. Adicional a esto, Carvajal *et al.* (1) determinaron cuantitativamente el color de los frutos en diferentes estados de desarrollo y concluyeron que existen diferencias de color definidas por componentes de luminosidad y cromaticidad sobre la superficie del epicarpio del fruto. Estos resultados potencian la relación entre la característica física del color con los estados de madurez presentados por los frutos de café, verificando la idea de relacionar los estados de madurez con alguna medida no destructiva sobre el fruto de café, que en este caso corresponde a la medida del color.

Debido a las condiciones climáticas en la zona cafetera, la producción de las plantaciones se encuentra muy distribuida a lo largo del año. Por esta razón, los recolectores deben realizar varios pases en el año para realizar las labores de cosecha, encontrando durante cualquier pase, frutos de café en todos los estados de maduración. Es así como los recolectores deben ser muy cuidadosos para hacer esta labor de forma selectiva, lo que encarece la cosecha y la hace responsable de cerca del 40% de los costos totales de producción en una finca cafetera. No obstante, una cuidadosa recolección, los trabajadores accidentalmente desprenden frutos indeseables, lo que causa inconsistencias en la calidad de la materia prima, como lo reporta Puerta (7). Para separar estos frutos y obtener café de alta calidad, se utilizan equipos en las diferentes etapas del proceso de beneficio.

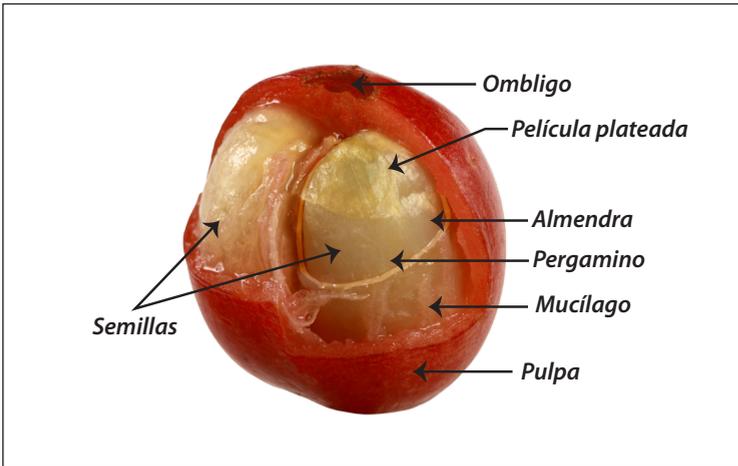


Figura 1.
Composición de un
fruto de café.

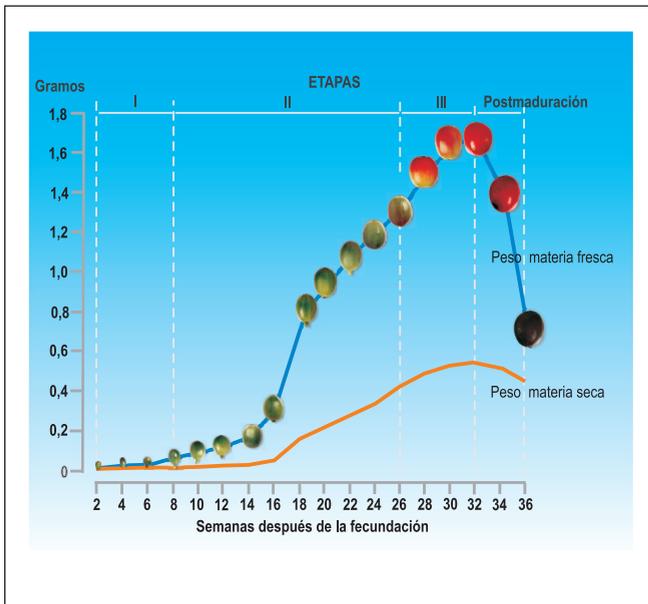


Figura 2.
Estados de
desarrollo de los
frutos de café.

Por ejemplo, antes del despulpado, los frutos de café se clasifican por densidad, por medio de un Separador Hidráulico de Tolva y Tornillo Sinfin (6), con esta máquina se retiran del proceso frutos vanos, secos o atacados por broca, los cuales flotan en la tolva de precipitación, además de piedras y objetos pesados, que se precipitan y quedan atrapados en una apéndice de la máquina. Luego del proceso de despulpado, se utiliza una zaranda donde los frutos despulpados

son clasificados por tamaño y se retiran del proceso frutos media cara, inmaduros de gran tamaño y enfermos, cuyo tamaño indica que aún tienen pulpa adherida (9).

En el estado del arte se encuentran investigaciones basadas en procesamiento digital de imágenes, en las cuales se ha logrado seleccionar frutos de café en condiciones controladas de laboratorio, Sandoval *et al.* (11, 12, 13), lograron porcentajes de clasificación

de 96,88%, a través del uso de cámaras digitales y computadores. En el estudio de Mosquera *et al.* (5), se identifican frutos de café en ramas, por medio del análisis multispectral de imágenes, requiriendo el uso de equipos sofisticados de visión por computador.

En la búsqueda de desarrollar un sistema sencillo, práctico y económico, Sanz *et al.* (14, 15), como antecedente a esta investigación, diseñaron un algoritmo de identificación basado en el espacio de color RGB; los resultados mostraron la capacidad de identificar cinco estados de maduración de frutos de café, sin embargo, el sistema mostró disminución de la sensibilidad con los cambios en la potencia luminosa de la fuente de iluminación.

En esta investigación se desarrolló un sistema de identificación no destructivo de frutos de café de variedades de fruto rojo, con el cual se puede determinar el estado de madurez de cada fruto, a través de una relación entre el color presente en la pulpa del fruto, en componentes “hue” (Matiz) y “saturation” (Saturación), y su respectivo estado de madurez, con el fin de analizar los frutos en tiempo real, en un sistema electrónico dedicado, independiente de un computador. La relación encontrada entre los componentes matiz y saturación (*hs*), fue implementada en un sistema embebido, conformado por un sensor de color, una fuente de iluminación, un sistema basado en microprocesador y un algoritmo de clasificación de bajo costo computacional, que logra identificar los estados de madurez eficazmente y en tiempo real.

MATERIALES Y MÉTODOS

Montaje experimental. Consistió en un sensor de color tipo CMOS, referencia TAOS-TCS230, incluido dentro de un módulo AXE045, y

una corona de iluminación con 12 LEDs de alto brillo, de luz blanca. La óptica del módulo consistió en una lente de 2 mm de diámetro, que permite proyectar 4 mm² de un objetivo ubicado a una distancia de 25 mm sobre el sensor de color CMOS de 64 fotodetectores de 1 mm². El TCS 230 tenía un protocolo de comunicación de dos bits para definir el canal a leer, R para rojo, G para verde y B para azul, y contó con una única salida de señal digital, cuya frecuencia varió dependiendo de la intensidad de color detectada.

El montaje experimental incluyó un disco rotatorio con 22 cuencas, para alojar el mismo número de frutos, para simular el flujo de frutos frente al sensor, por medio de un motor DC y un sistema de control basado en un modulador de ancho de pulsos PWM (del inglés Pulse Width Modulation), para variar la velocidad de exposición de los frutos en frente del sensor.

El sistema completo fue controlado por computador, por medio de los módulos de tiempo real de LabView® 7.1. La Figura 3 muestra el montaje experimental desarrollado para caracterizar el color sobre los frutos de café.

Correlación color-estado de maduración.

Los estados de maduración sobre los cuales se realizó la investigación fueron inmaduro, pintón, maduro y sobremaduro. El estado inmaduro se asignó a los frutos con menos de 25 semanas de desarrollo, el pintón a aquellos frutos con 26 a 29 semanas, el maduro a frutos entre 30 y 33 semanas, y el sobremaduro a frutos con cerca de 34 semanas después de floración.

El montaje experimental se usó para correlacionar el color de la pulpa de los frutos con su estado de maduración. Para tal fin se caracterizaron 100 frutos de café por

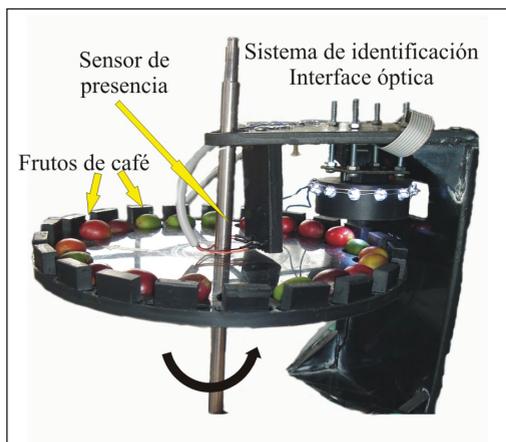


Figura 3. Montaje experimental desarrollado para caracterizar el color sobre los frutos de café.

estado de madurez (inmaduro, pintón, maduro, sobremaduro y seco), en total se requirieron 500 frutos para realizar la caracterización. Dichas muestras fueron seleccionadas por expertos antes de exponerlas frente al sistema optoelectrónico. Cada lugar en el disco tuvo la información del estado de maduración de cada fruto que allí se dispuso.

Se exploraron las diferencias en la representación del color HSV (del inglés Hue, Saturation y Value). Por lo tanto, se adquirieron las variables RGB (del inglés Red, Green and Blue) del sensor para cada fruto de café dispuesto en el disco, se tomaron los estados de maduración y se transformaron a las componentes de color HSV (2).

Como la variable h , en la representación de color HSV, es un círculo y el color rojo (maduro) tiene valores antes y después del cero, se realizó un corrimiento hasta el punto en que se aseguró que los frutos maduros permanecían en una ubicación del círculo sin sufrir particiones en el mismo (Figura 4).

Algoritmo de aprendizaje. El algoritmo encuentra los puntos de corte que limitan cuatro áreas diferentes en el plano h - s , para

cada estado de maduración. Para calcular los límites, el algoritmo usa un proceso denominado Optimización de la Eficacia Global (OEG), con el fin de identificar el máximo número de lecturas en el área correcta. Con esta optimización se encontraron los intervalos en los cuales se contienen el mayor número de lecturas correctas con el mínimo efecto de lecturas incorrectas, debido a los estados de maduración vecinos. Los pasos que siguió el algoritmo fueron:

Paso 1. Conversión de espacio de color de RGB a HSV, Ecuaciones <<1>>, <<2>> y <<3>>

$$v = \max\{r, g, b\} \tag{1}$$

$$h = \begin{cases} \frac{g-b}{6s}, \text{ para } v = r & \tag{2} \\ \left(2 + \frac{b-r}{s}\right) / 6, \text{ para } v = g & \tag{3} \\ \left(4 + \frac{r-g}{s}\right) / 6, \text{ para } v = b \end{cases}$$

$$\text{if } \{h < 0\} \rightarrow h = h + 1$$

$$s = v - \min\{r, g, b\}$$

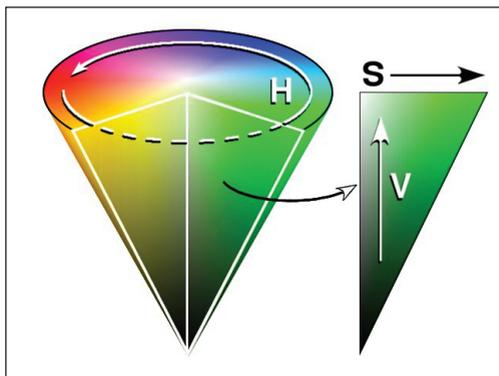


Figura 4. Esquema representación de color HSV.

Donde: r , g y b son las intensidades de rojo, verde y azul, expresadas en forma decimal de 0 a 1, y h , s y v son valores decimales de tono (*hue*), saturación (*saturation*) y luminancia (*value*). Las funciones mín. y máx. representan valores mínimos y máximos, respectivamente.

Paso 2. Aplicación del corrimiento sobre la variable h , para asegurar que los valores de tonos rojos se encontraban en una sola sección del plano cartesiano, sin cortes. El valor definido para el corrimiento fue de 0,9 determinado empíricamente como resultado de la caracterización (Ecuación <<4>>).

$$\begin{aligned} \text{if}\{h \geq 0,9\} &\rightarrow h = h - 0,9 \\ \text{if}\{h < 0,9\} &\rightarrow h = h + 0,1 \end{aligned} \quad \ll 4 \gg$$

Paso 3. Construcción de una matriz de frecuencias acumuladas de 8 x 100 dividida en dos partes. La parte superior con cuatro vectores de frecuencia acumulada de la variable h de cada estado de maduración y los restantes correspondieron a los cuatro vectores de frecuencia acumulada de la variable s , como se observa en la Ecuación <<5>>. Se cambió la escala de las variables h y s , para expresarlas con valores entre 0 y 100.

$$F = \begin{bmatrix} f_i(h_1) & \sum_{k=1}^2 f_i(h_k) & \dots & \sum_{k=1}^{100} f_i(h_k) \\ f_u(h_1) & \sum_{k=1}^2 f_u(h_k) & \dots & \sum_{k=1}^{100} f_u(h_k) \\ f_m(h_1) & \sum_{k=1}^2 f_m(h_k) & \dots & \sum_{k=1}^{100} f_m(h_k) \\ f_o(h_1) & \sum_{k=1}^2 f_o(h_k) & \dots & \sum_{k=1}^{100} f_o(h_k) \\ f_i(s_1) & \sum_{k=1}^2 f_i(s_k) & \dots & \sum_{k=1}^{100} f_i(s_k) \\ f_u(s_1) & \sum_{k=1}^2 f_u(s_k) & \dots & \sum_{k=1}^{100} f_u(s_k) \\ f_m(s_1) & \sum_{k=1}^2 f_m(s_k) & \dots & \sum_{k=1}^{100} f_m(s_k) \\ f_o(s_1) & \sum_{k=1}^2 f_o(s_k) & \dots & \sum_{k=1}^{100} f_o(s_k) \end{bmatrix} \quad \ll 5 \gg$$

Donde: f se refiere a la frecuencia de las ocurrencias de los subíndices i , u , m y o , que se relacionan con los estados de maduración inmaduro (i), pintón (u), maduro (m) y sobremaduro (o).

La Figura 5 muestra el plano h - s y la ubicación de los diferentes estados de maduración, al igual que las gráficas de frecuencia acumulada de cada una de las variables en su eje correspondiente. El algoritmo toma ventaja de la ubicación de cada estado de maduración por separado.

Paso 4. Determinación del rango de la clase de interés (estado de maduración de interés). Para este desarrollo, una clase de interés, correspondió a la clase que es influenciada por clases vecinas. La clase de interés de h fue pintón y sus vecinos fueron inmaduro y maduro. La clase de interés en la variable s fue sobremaduro y sus vecinos fueron inmaduro, pintón y maduro. El rango de la clase de interés se determinó por P_{\min} , el cual correspondió a la posición donde el vector de frecuencia acumulada tuvo por primera vez un valor diferente de cero, y P_{\max} , que correspondió a la posición donde el vector de frecuencia acumulada tuvo por primera vez el valor de uno (Figura 5).

Paso 5. División del rango de observación en dos conjuntos de puntos. Se crearon dos vectores de puntos, x y y , donde el punto medio estuvo dado por la Ecuación <<6>>; cuando el vector tiene un número impar de elementos el vector x tiene un elemento más que y (Ecuación <<7>>).

$$P_{\text{mid}} = \text{int}\left(\frac{P_{\text{max}} + P_{\text{min}}}{2}\right) \quad \ll 6 \gg$$

$$\begin{aligned} x &= [P_{\text{min}} \quad P_{\text{min}} + 1 \quad \dots \quad P_{\text{mid}}] \\ y &= [P_{\text{mid}} + 1 \quad P_{\text{mid}} + 2 \quad \dots \quad P_{\text{max}}] \end{aligned} \quad \ll 7 \gg$$

Paso 6. Cálculo de la eficacia como la diferencia entre el valor de la frecuencia

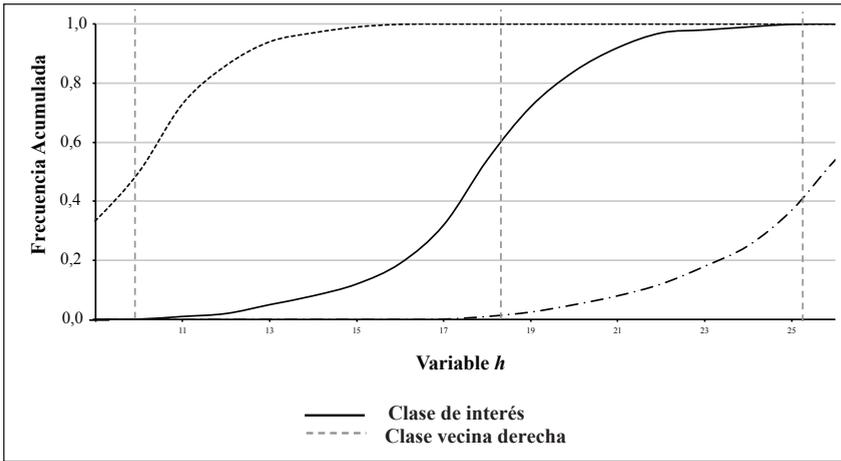


Figura 5. Frecuencias acumuladas de la variable tono (h) en el rango de observación.

acumulada de un punto del vector de y con el valor de la frecuencia acumulada de un punto del vector de x , de la clase de interés.

Con la ecuación de eficacia (Ecuación <<8>>), se construyó una matriz de eficacias de la clase de interés E_{int} , como se muestra en la Ecuación <<9>>. Esta matriz tiene todas las posibles combinaciones posibles entre los puntos seleccionados de x y y .

$$\eta_{int}(x_i, y_j) = \sum_{k=x_1}^{x_i} f_u(h_k) - \sum_{k=y_1}^{y_j} f_u(h_k) \quad \ll 8 \gg$$

$$E_{int} = \begin{bmatrix} \eta_{int}(x_1, y_1) & \eta_{int}(x_1, y_2) & \dots & \eta_{int}(x_1, y_j) \\ \eta_{int}(x_2, y_1) & \eta_{int}(x_2, y_2) & \dots & \eta_{int}(x_2, y_j) \\ \vdots & \vdots & \ddots & \vdots \\ \eta_{int}(x_i, y_1) & \eta_{int}(x_i, y_2) & \dots & \eta_{int}(x_i, y_j) \end{bmatrix} \quad \ll 9 \gg$$

Paso 7. Cálculo de la eficacia de las clases vecinas usando la Ecuación <<10>>, sobre el mismo rango de observación, donde el subíndice n es la clase vecina, que en este caso correspondió a inmaduros (i) o maduros (m).

$$\eta_{nei}(x_i, y_j)_n = \sum_{k=x_1}^{x_i} J_n(h_k) - \sum_{k=y_1}^{y_j} J_n(h_k) \quad \ll 10 \gg$$

Con la Ecuación <<10>> se construyó la matriz de eficacia de cada clase vecina, E_{nei} , como se muestra en la Ecuación <<11>>.

$$E_{nei} = \begin{bmatrix} \eta_{nei}(x_1, y_1)_n & \eta_{nei}(x_1, y_2)_n & \dots & \eta_{nei}(x_1, y_j)_n \\ \eta_{nei}(x_2, y_1)_n & \eta_{nei}(x_2, y_2)_n & \dots & \eta_{nei}(x_2, y_j)_n \\ \vdots & \vdots & \ddots & \vdots \\ \eta_{nei}(x_i, y_1)_n & \eta_{nei}(x_i, y_2)_n & \dots & \eta_{nei}(x_i, y_j)_n \end{bmatrix} \quad \ll 11 \gg$$

Paso 8. Cálculo de la intersección entre la clase de interés y una clase vecina. Se calculó como el rango donde se encuentran lecturas de la clase vecina dentro del rango de observación. El índice de intersección de la clase vecina izquierda se calculó como se muestra en la Ecuación <<12>> y el índice de intersección de la clase vecina derecha se calculó como se presenta en la Ecuación <<13>>.

$$O_l(x_i, y_j) = \eta_{nei}(x_i, y_j)_n \quad \ll 12 \gg$$

$$O_r(x_i, y_j) = 1 - \eta_{nei}(x_i, y_j)_n \quad \ll 13 \gg$$

Paso 9. Cálculo de la eficacia global, η_G , como el peso de la eficacia de la clase de interés con cada índice de intersección de las clases vecinas (Ecuación <<14>>)

$$\eta_G(x_i, y_j) = \eta_{int}(x_i, y_j) O_l(x_i, y_j) O_r(x_i, y_j) \quad \ll 14 \gg$$

Con la Ecuación <<14>> se construyó la matriz de eficacia global, E_{global} , que se presenta en la Ecuación <<15>>.

$$E_{\text{global}} = \begin{bmatrix} \eta_G(x_1, y_1) & \eta_G(x_1, y_2) & \dots & \eta_G(x_1, y_j) \\ \eta_G(x_2, y_1) & \eta_G(x_2, y_2) & \dots & \eta_G(x_2, y_j) \\ \vdots & \vdots & \vdots & \vdots \\ \eta_G(x_i, y_1) & \eta_G(x_i, y_2) & \dots & \eta_G(x_i, y_j) \end{bmatrix} \quad \ll<15>>$$

Paso 10. Determinación del valor más alto de la eficacia global. La posición (x, y) , en la matriz, de este valor indicó los límites óptimos del intervalo que contenía mayor cantidad de información de un estado de madurez y una mínima influencia de los estados vecinos (Ecuación <<16>>).

$$\begin{aligned} [L_{\text{inf}}, L_{\text{sup}}] &= \text{posición} \{ \max(E_{\text{global}}) \} \\ L_{\text{inf}} &= \text{Límite inferior} \rightarrow \text{Columna} \\ L_{\text{sup}} &= \text{Límite superior} \rightarrow \text{Fila} \end{aligned} \quad \ll<16>>$$

Para la clase de interés pintón y la variable h , la superficie creada de la matriz de eficacia global se presenta en la Figura 6, donde el valor mayor se encuentra en la posición $(18, 38)$, e indica los límites inferior en 18 y superior en 38.

Finalmente, el clasificador desarrollado obtiene los puntos de corte para la variable h y para la variable s . Para la variable h se definió el criterio para clasificar tres estados de maduración, maduro, pintón, e inmaduro, así:

$$\begin{cases} \text{Maduro} & 0 < h < 0,18 \\ \text{Pintón:} & 0,18 < h \leq 0,38 \\ \text{Inmaduro:} & 0,38 < h < 0,73 \end{cases} \quad \left. \begin{array}{l} \\ \\ \end{array} \right\} \text{for } s \geq 0,36$$

Para la variable s se definió el criterio para clasificar el estado sobremaduro, así:

Sobremaduro: $0 \geq 0,36$
independiente de la variable h .

El algoritmo de aprendizaje encontró el intervalo óptimo y separó los estados de maduración por medio de un clasificador lineal simple.

Criterio de decisión. Con el sistema de identificación desarrollado, sobre un fruto de café se tomaron entre 7 y 11 imágenes; cabe anotar que fue común encontrar algunas imágenes de un mismo fruto, pertenecientes a estados de maduración diferentes. Con el fin de determinar el estado de maduración global del fruto, se usó la moda estadística sobre la maduración asignada a las lecturas sobre un mismo fruto. De esta forma el estado de maduración global del fruto correspondió al estado de maduración que más se repitió sobre un mismo fruto.

Clasificación en tiempo real e implementación de un sistema embebido. El algoritmo desarrollado fue implementado en un sistema embebido, basado en un microcontrolador dsPIC30F4012, el cual realizó todas las operaciones con una base de tiempo de 20 ns, con capacidad de analizar 50 frutos/s. El sistema embebido tuvo dos funciones, la función de calibración, en la cual el dispositivo aprendió los criterios de identificación, y la función de clasificación en la cual el dispositivo identificó el estado de maduración de cada fruto que pasó enfrente del sistema de identificación.

Evaluación del clasificador. El clasificador se evaluó con los datos empleados en el entrenamiento y con datos obtenidos sobre frutos de café iluminados con una fuente de luz de menor potencia. Las imágenes usadas correspondían a frutos de café con diferentes humedades y un rango de distancia que varió entre 20 mm y 35 mm del fruto a la lente.

Evaluación del clasificador implementado en el sistema embebido. Para verificar el desempeño del sistema de identificación se consideraron diversas velocidades de exposición de 1, 10, 30 y 50 frutos por segundo, y diversos tamaños de los frutos a analizar. Expertos seleccionaron 20 muestras de 20 frutos de café cada una, y las proporciones

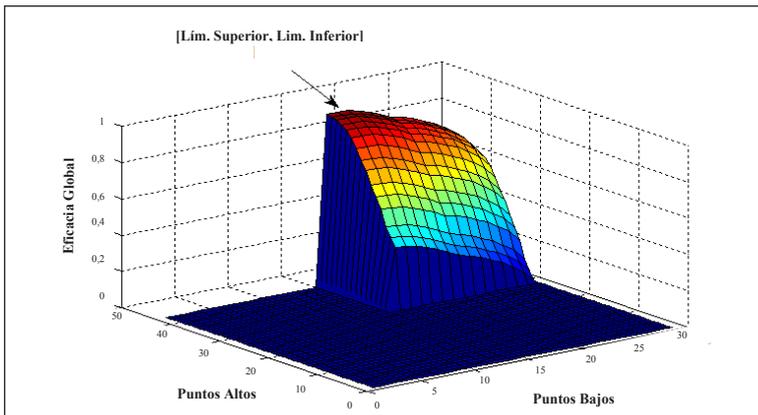


Figura 6. Superficie que muestra el punto óptimo de corte de la eficacia global.

de estados de madurez en cada muestra variaron, de acuerdo como se presenta en la Tabla 1, usando el sistema embebido desarrollado.

En cada prueba se registró la proporción de estados de madurez de la muestra, la velocidad de exposición, el número de lecturas realizadas en cada estado de madurez a cada fruto de café y el estado de maduración global determinado por el sistema; además, cada fruto fue etiquetado para verificar y validar la información registrada. Cada prueba se repitió tres veces.

RESULTADOS Y DISCUSIÓN

Se encontraron diferencias notables entre los estados de maduración en el plano hs . Dichas diferencias fueron independientes a los cambios suaves en la potencia de la fuente luminosa, a la distancia al objetivo (rango entre 20 mm y 35 mm), a la humedad del medio ambiente y a la presencia de una película de agua sobre la superficie del fruto.

El número de imágenes por segundo de cada fruto varió dependiendo de la velocidad de exposición del fruto frente del sensor de color. En la Figura 7 se muestra el plano $h-s$ y los puntos que representan los diferentes estados de maduración extraídos

con el sistema optoelectrónico descrito, cada punto en el plano corresponde a una lectura de color sobre un fruto de café; dependiendo de la velocidad de exposición y del tamaño del fruto, puede variar el número de lecturas. En la Figura 7 se identifican secciones para cada estado de maduración, maduro, pintón e inmaduro sobre la variable h , y los frutos sobremaduros son diferenciables en la variable s .

Evaluación del clasificador. La Tabla 2 muestra resultados por encima del 87,55%, con los mayores valores para los frutos sobremaduros, seguido por los frutos maduros, inmaduros y pintones. Este análisis fue realizado usando Tablas de contingencia o matriz de confusión (3), con la cual se registra el óptimo desempeño del sistema de identificación para todos los estados de maduración, independiente de las variaciones de brillo, distancia y potencia luminosa. La Tabla 2 muestra que la eficacia de identificación de dos conjuntos de datos, definida por Kohavi *et al.* (3); el conjunto A corresponde a los datos de entrenamiento con una fuente de alta luminosidad y el conjunto B corresponde a los datos tomados con una fuente de potencia luminosa media. La eficacia de identificación varió entre 87,6% y 98,17%, lo que hace el clasificador confiable para el propósito que fue desarrollado.

Tabla 1. Descripción porcentual de las muestras de café evaluadas.

Muestra	Inmaduro(%)	Pintón(%)	Maduro(%)	Sobremaduro(%)
1	20	10	55	15
2	15	35	45	5
3	15	20	55	10
4	10	15	65	10
5	5	10	75	10
6	10	55	5	30
7	30	45	0	25
8	20	55	5	20
9	15	65	5	15
10	15	75	0	10
11	55	20	10	15
12	45	15	35	5
13	55	15	20	10
14	65	10	15	10
15	75	5	10	10
16	5	5	10	80
17	5	5	35	55
18	5	10	20	65
19	5	10	15	70
20	5	5	10	80

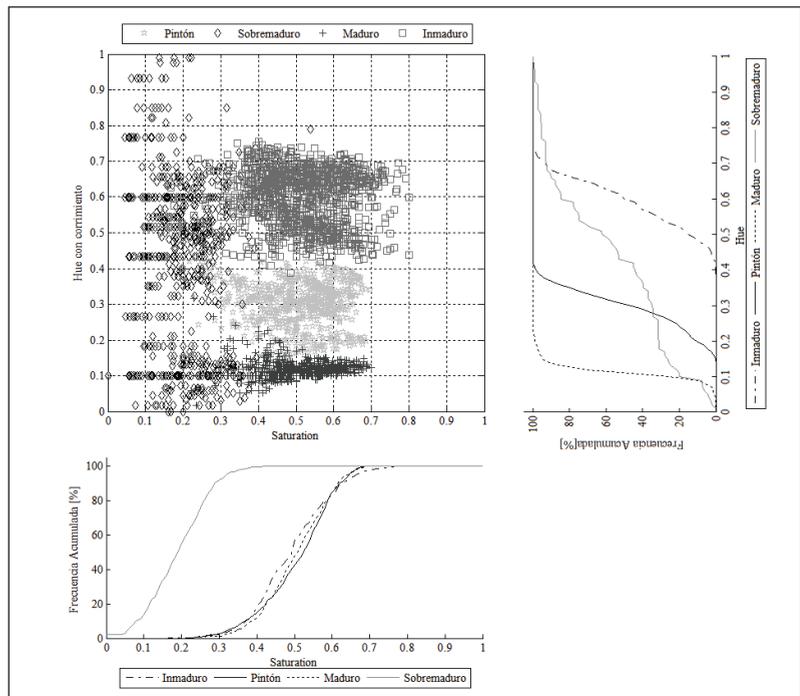


Figura 7. Comportamiento de las componentes h y s para determinar el estado de madurez de frutos de café.

Evaluación del clasificador implementado en el sistema embebido. A través de las Tablas de contingencia se encontró que el desempeño del sistema de identificación es

alto para todos los estados de madurez, superior al mostrado en la evaluación inicial del clasificador (Tabla 2). La Tabla 3 muestra que el sistema tiene una eficacia que varía

entre 94,81% y 99,62%, como resultados globales.

Además, se verificó que el sistema trabaja correctamente en tiempo real, pues la eficacia de identificación es alta?, independiente de la velocidad de exposición de los frutos y las condiciones reales de los frutos analizados. Los porcentajes de error presentados en la Tabla 3, fueron generados por el tipo de clasificador implementado, sin embargo éste presenta ventajas de consumo mínimo de recursos computacionales, lo cual hace que el sistema funcione en tiempo real a altas velocidades de exposición de los frutos.

Se puede concluir que:

La relación *hs*, de las componentes “hue y “saturation” del espacio de color HSV, entrega una clara diferenciación entre estados de maduración, con una alta tolerancia a los cambios de iluminación y a las condiciones de brillo de los frutos de café, en comparación con otros resultados obtenidos en el espacio de color RGB (14, 15). Adicionalmente, este tipo de relación permitió diseñar e implementar un clasificador sencillo con bajos requerimientos

computacionales logrando clasificaciones en tiempo real con altos porcentajes de eficacia de identificación.

El sistema de identificación logró eficacias mayores al 95%, para todos los estados de maduración, independiente de la velocidad de exposición de los frutos frente al sensor de color, además de otras condiciones reales con las que ingresan los frutos de café al beneficiadero.

El algoritmo de aprendizaje potencia el uso de sistemas microprocesados de bajo costo y de fácil adquisición en el mercado nacional, pues los requerimientos computacionales son mínimos, lo que garantiza que el sistema funcione adecuadamente en tiempo real.

Debido al bajo costo del dispositivo electrónico y la simplicidad de los componentes utilizados, y el bajo requerimiento computacional de los algoritmos implementados, este sistema es una opción para realizar control de calidad en línea para los caficultores colombianos y de otros países, además para productores de productos similares al café.

Tabla 2. Eficacia de identificación del clasificador con la información obtenida en la caracterización. Estadísticas obtenidas del total de imágenes clasificadas correctamente en función al número total de imágenes.

Estado de maduración	Conjunto A	Conjunto B
Sobremaduro	98,17%	95,77%
Maduro	97,21%	92,80%
Pintón	91,95%	88,85%
Inmaduro	90,84%	87,55%

Tabla 3. Eficacia de identificación del clasificador implementado en el sistema embebido. Estadísticas obtenidas del total de imágenes clasificadas correctamente en función al número total de imágenes

Experto	Clasificador			
	Sobremaduro	Maduro	Pintón	Inmaduro
Sobremaduro	99,62%	0,31%	0,08%	0,00%
Maduro	1,69%	97,80%	0,51%	0,00%
Pintón	0,00%	3,74%	94,81%	1,45%
Inmaduro	0,00%	2,22%	2,57%	95,21%

Como el dispositivo electrónico no es usado al 100% de sus potencialidades, existe la posibilidad de incluir una nueva característica para crear cafés especiales con diferentes proporciones de frutos en diferentes estados de maduración.

El algoritmo de aprendizaje puede definir puntos de corte para subestados de maduración que se encuentren dentro del estado de madurez “maduro”, con los cuales fuese posible crear productos de alta calidad y de forma consistente.

AGRADECIMIENTOS

Los autores expresan sus agradecimientos al Departamento Nacional de Ciencia, Tecnología e Innovación – Colciencias, por el apoyo económico en el desarrollo de esta experimento, a la Federación Nacional de Cafeteros de Colombia y a los compañeros de la Disciplina de Ingeniería Agrícola de Cenicafé.

LITERATURA CITADA

1. CARVAJAL, J.J.; ARISTIZÁBAL, I.D.; OLIVEROS, C.E.; MEJÍA, J.W. Colorimetría del fruto de café (*Coffea arabica* L.) durante su desarrollo y maduración. Rev. Fac. Nal. Agr. Medellín. 64(2): 6229-6240. 2011.
2. GONZÁLEZ, R.C.; WOODS, R.E. Digital image processing. 3 ed. New Jersey (Estados Unidos), Pearson : Prentice Hall, 2008. 954 p.
3. KOHAVI, R.; PROVOST, F. “Glossary of terms”. Special issue on applications of machine learning and the Knowledge Discovery process. Machine Learnig, 30 (2). 271 – 274.
4. MARÍN L., S.M.; ARCILA P., J.; MONTOYA R., E.C.; OLIVEROS T., C.E. Escala de maduración para los frutos del cafeto *Coffea arabica* L. Avances Técnicos Cenicafé No. 315:1-8. 2003.
5. MOSQUERA, J.C., C.A. ISAZAYA. SEPÚLVEDA. 2007. Procesamiento de imágenes ópticas de frutos café en

cereza por medio de filtros acusto-ópticos. Ingeniería y Desarrollo, Universidad del Norte 21: 93-102.

6. OLIVEROS T., C.E.; SANZ U., J.R.; RAMÍREZ G., C.A.; MEJÍA G., C.A. Separador hidráulico de tolva y tornillo sinfin. Avances Técnicos Cenicafé No. 360:1-8. 2007.
7. PUERTA Q., G.I. Influencia de los granos de café cosechados verdes, en la calidad física y organoléptica de la bebida. Revista Cenicafé 51(2):136-150. 2000.
8. RAMOS, P.J.; VALDÉS, C.; GÓMEZ, E.; SANZ, J. R.; SOLARTE, E. Características Espectrales de la Luz Reflejada por Frutos de Café (*Coffea arabica*). Revista Colombiana de Física, Vol. 38, No. 2, 2006. pp. 822-825.
9. ROA M., G.; OLIVEROS T., C.E.; ÁLVAREZ G., J.; RAMÍREZ G., C.A.; SANZ U., J.R.; DÁVILA A., M.T.; ÁLVAREZ H., J.R.; ZAMBRANO F., D.A.; PUERTA Q., G.I.; RODRÍGUEZ V., N. Beneficio ecológico del café. Chinchiná, Cenicafé, 1999. 273 p.
10. SALAZAR G., M.R.; CHAVES C., B.; RIAÑO H., N.M.; ARCILAP, J.; JARAMILLOR, A. Crecimiento del fruto de café *Coffea arabica* var. Colombia. Revista Cenicafé 45(2):41-50. 1994.
11. SANDOVAL, Z.; PRIETO, Y F. 2007. Caracterización de café cereza empleando técnicas de visión artificial. Revista Facultad Nacional Agronomía Medellín 60(2): 4105-4127.
12. SANDOVAL, Z.; PRIETO, F. Procesamiento de imágenes para la clasificación de café cereza. Prospectiva Vol. 7, No. 1, Enero - Junio de 2009. Pp: 67-73.
13. SANDOVAL, Z.; PRIETO, F.; BETANCUR, J.; Digital Image Processing for Classification of Coffee Cherries. In: Electronics, Robotics and Automotive Mechanics Conference (CERMA), 2010. Pp: 417 – 421.
14. SANZ, J.R.; RAMOS, P.J.; C.E. OLIVEROS. Algorithm to Identify Maturation Stages of Coffee Fruits. Advances in electrical and Electronics Engineering – IAENG Special Edition of Word Congress on Engineering and Computer Science 2008. IEEE Publisher.
15. SANZ, J.R.; RAMOS, P. J.; C.E. OLIVEROS. Optoelectronic Analyzer of Coffee Fruits. Word Congress on Engineering and Computer Science 2009”. San Francisco, California, USA, Octubre 10-14 de 2008.