

# Procesamiento de imágenes para la clasificación de café cereza

## Images processing classification for cherry coffee

Zulma Sandoval\*, Flavio Prieto\*\*

\*Msc. Ingeniería Automatización Industrial. Ing. Electrónica. Universidad Autónoma del Caribe.  
Grupo de investigación IET – UAC. E-mail: zulma.sandoval@uac.edu.co.

\*\* Doctor en Automática Industrial. Universidad Nacional de Colombia Sede Bogotá.  
Grupo de Automática de la Universidad Nacional. E-mail: faprieto@unal.edu.co

### RESUMEN

Se desarrolla un sistema que usa técnicas de visión artificial para clasificar frutos de café en ocho categorías según el estado de maduración en el que se encuentre. Las ocho categorías comprenden todo el proceso de maduración del café desde las etapas iniciales hasta sobremaduros y secos. A partir de un conjunto de 9 características que incluyen medidas de color, forma y textura obtenidas del análisis de una imagen del fruto, se logró implementar un clasificador Bayesiano cuyo desempeño, evaluado mediante el método de validación cruzada, corresponde al 96.88%.

**Palabras clave:** Clasificación, reconocimiento de patrones, café, procesamiento digital de imágenes, visión artificial.

### ABSTRACT

A machine vision system to grade coffee fruits to one of eight classes regarding its state of ripeness is proposed. Those eight classes cover the entire ripeness process, from the starting ripeness stage to over-matured and dried fruits (last state). Using a set of 9 measurements of color, shape and texture features on the fruit image, a Bayesian classifier was implemented with a 96.88% of performance, using cross-validation.

**Palabras clave:** Clasificación, reconocimiento de patrones, café, procesamiento digital de imágenes, visión artificial.

## 1. Introducción

Las condiciones climatológicas que se presentan durante todo el año en el territorio colombiano, hacen que el cafeto no fructifique en forma homogénea. En consecuencia, en cualquier época, es posible encontrar en la misma rama frutos en diferentes estados de maduración e incluso flores. La proporción de frutos maduros en una rama durante el periodo de mayor producción es inferior al 60%. Esta característica de fructificación del cafeto afecta la forma como se desarrolla la recolección, convirtiéndola en una tarea compleja y costosa. Para cosechar todos los frutos

se deben realizar entre 10 y 15 pases al año (Wallis, Montoya, Vélez, & Oliveros, 2004). Además, en cada pase se deben recoger uno a uno los frutos plenamente maduros evitando desprender los verdes o las flores. Esto hace que la eficiencia de la recolección, que se mide en kg por hora, sea baja. Por ejemplo, para el método tradicional manual selectivo la eficiencia por persona es de 10 kg de frutos desprendidos por hora (Wallis, Montoya, Vélez, & Oliveros, 2004). Por esta razón, la mano de obra necesaria para la cosecha incrementa los costos de recolección, de tal forma que representan entre el 35 y el 42 por ciento de los costos totales de producción (López, Oliveros, & Ramirez, 2006).

CENICAFÉ (Centro de Investigaciones de la Federación Nacional de Cafeteros) ha impulsado la investigación tendiente al mejoramiento y desarrollo tanto de sistemas tecnológicos como de estrategias, que permitan el incremento de la eficiencia de la recolección y en consecuencia una disminución en los costos. En esta institución, se han estudiado diferentes métodos de recolección que han mostrado un mejoramiento de la eficiencia con respecto al método selectivo en más de un 100% (23kg /h a 26.7kg/h) (Wallis, Montoya, Vélez & Oliveros, 2004; López, Oliveros & Ramírez, 2006). Sin embargo, se afecta la calidad de la recolección, que se mide como la proporción de frutos maduros con respecto a todos los frutos recogidos (masa cosechada). Para las técnicas estudiadas, la calidad de la recolección varía entre el 81 y el 88 por ciento. La presencia de frutos verdes, sobremaduros o secos en la masa cosechada afectan la calidad de la bebida y en consecuencia el precio de cotización en el mercado internacional (Marín, Arcila, Montoya, & Oliveros, 2003; Puerta, 1998). La acidez de la bebida se ve afectada cuando el porcentaje de frutos verdes en la masa cosechada sobrepasa el 2%, por otra parte, el amargo, el cuerpo y la impresión global se ven alterados cuando se sobrepasa el 2,5% (Isaza, Montoya, Vélez, & Oliveros, 2006). Por lo tanto, para preservar la calidad del café colombiano deben mantenerse porcentajes de frutos maduros superiores al 98% en la masa cosechada.

Con el fin de mantener los niveles de eficiencia sin afectar la calidad de la bebida, se ha pensado en la posibilidad de incluir una etapa de clasificación después de la recolección. Un sistema que utilice el análisis de las características de tamaño, forma y color del fruto del café a partir de la observación de una imagen, para clasificarlos según el estado de maduración en que se encuentre. Las ventajas de utilizar un sistema de clasificación que emplee procesamiento de imágenes con respecto a sistemas de tipo óptico o electromecánico son principalmente, la posibilidad de diferenciar un gran número de etapas de maduración y la reducción del contacto físico con el fruto que pueda causar daño mecánico. Clasificar los frutos de café cosechados en diferentes estados de maduración mejora la producción de café en dos sentidos, por un lado, se obtendría un café de muy buena calidad con porcentajes superiores al 98% de frutos maduros y por el otro, la diversificación de la producción al elaborar café de cada una de las otras etapas, es decir café con características especiales de aroma, sabor y textura aprovechando la reciente aparición de nichos de mercado para los cafés con valor agregado o diferenciados.

La clasificación de café usando procesamiento de imágenes ha sido estudiada con anterioridad en varios trabajos (Sandoval & Prieto, 2007), (Betancur, Prieto &

Osorio, 2006), donde se obtuvieron resultados promisorios con índices de desempeño superiores al 95%, sin embargo, se debe alcanzar el índice de clasificación superior al 98%. En el presente trabajo se obtiene un mejoramiento del desempeño del orden del 96.88%, esto se logra haciendo la clasificación por etapas de maduración del café según el trabajo presentado por Marín, Arcila, Montoya & Oliveros (2003) y no por número de muestras como lo veníamos desarrollando en pruebas pasadas (Sandoval & Prieto, 2007).

## 2. Adquisición de Imágenes

El proceso de selección de los frutos de café utilizados en la experimentación fue supervisado y desarrollado por Cenicafé y la Universidad de Caldas, en el marco de un proyecto de investigación. En este proyecto, se seleccionó un lote compuesto por aproximadamente 7500 plantas de café *coffea Arabica* variedad Colombia cereza roja, ubicado en la estación central Naranjal en Chinchiná, con condiciones de cultivo representativas de toda la región del eje cafetero, localizado a una altitud de 1400m, con una temperatura promedio de 21,3°C y una humedad relativa del 78%. De las 7500 plantas, se seleccionó el 13.3%, del cual se desprendieron semanalmente 100 frutos, a partir de la semana 26 después de la floración de más del 60% de las plantas y hasta la semana 33. Se tomaron 100 imágenes de dichos frutos con tres sistemas de adquisición diferentes. Como el interés en este trabajo es la clasificación de los frutos, en la adquisición se fotografió un fruto por imagen. Sin embargo, para realizar la clasificación a partir de imágenes con muchos frutos y condiciones de iluminación menos controladas, se deben utilizar algoritmos de segmentación de los frutos como lo presentados en (Betancur, Prieto & Osorio, 2006),.

El sistema de adquisición implementado está compuesto por una cámara de video a color (3 CCD), una tarjeta digitalizadora National Instruments, un computador y dos lámparas de luz blanca. Con este sistema se tomaron y almacenaron imágenes que corresponden a la vista superior de los frutos dispuestos individualmente sobre un fondo e iluminados con luz blanca difusa. Para variar el sistema de adquisición se cambió la posición de la iluminación así como el color del fondo. El número de imágenes usadas en éste estudio fueron 300 por cada una de las etapas de maduración, es decir se contó con una base de datos de 2400 imágenes de tamaño 160x160 píxeles con una resolución de 24 bits por píxel almacenadas en formato bmp.

## 3. Caracterización

La caracterización de los frutos de café se hace con base en sus propiedades internas como son textura y color, así como en sus propiedades externas que corresponden a la forma y el tamaño. Para caracterizar

las propiedades internas, se tomó una matriz cuadrada del centro de la imagen, de tamaño 50x50 píxeles, que corresponde a la porción central de cada fruto. De la misma forma, las propiedades externas, se miden a partir del contorno del fruto. Para obtener el contorno, primero se aplica el filtro mediana con una ventana cuadrada de tamaño 7 píxeles para la reducción del ruido. Luego, se convierten las imágenes RGB al espacio de color HSI. La componente de luminancia se umbraliza para separar el fondo del fruto. De esta forma, se obtiene una imagen resultante que contiene dos objetos el fondo de color blanco y el fruto de color negro. Finalmente, se aplica un filtro pasa altas tipo Sobel, para la obtención del contorno.

### 3. 1 Características

Se aplicaron diferentes algoritmos para obtener las medidas de color, textura y forma de cada uno de los frutos de café. No se obtuvieron medidas del tamaño debido a que en la adquisición no se hizo referencias de distancias, así como no se tiene un estudio óptico de los lentes de la cámara utilizada.

#### 3.1.1. Características de color

Los frutos de café de la variedad Colombia cereza roja, presentan un cambio de color durante todo el proceso de maduración, de hecho, la especificación cereza roja, indica que el fruto presenta un color rojo intenso homogéneo en toda la epidermis cuando se encuentra plenamente maduro. En las primeras etapas se presentan diferentes tonalidades de verde, le siguen tonos naranjas y rosados hasta tomar el color rojo de la madurez, y cuando están sobremaduros el color llega a ser violeta oscuro.

Para la caracterización del color se usaron seis espacios de representación del color: RGB, HSI, YIQ, YCbCr, Colores oponentes, y Ohta. Cada uno de estos espacios describe el color usando tres componentes. Por lo tanto, para cada imagen de fruto de café se obtuvo un vector de tamaño 18, donde cada posición muestra la media estadística de cada componente para los diferentes espacios.

#### 3.1.2. Características de textura

Las características de textura aplicadas a imágenes de café se centran en el análisis de la suavidad y homogeneidad en la distribución de la intensidad de los colores, así como el análisis de la rugosidad de la epidermis en las etapas finales cuando los frutos están demasiado maduros o secos. En el proceso de maduración del café se tienen etapas donde un color se presenta de manera homogénea sobre toda la epidermis, mientras que en otras etapas hay una variación suave o brusca de diferentes colores. El método de

extracción de características de textura que se usó en esta investigación, es el análisis estadístico de la matriz de coocurrencia. La matriz de coocurrencia tiene información de la distribución de las intensidades de los píxeles dentro de la imagen y sus relaciones con el vecindario.

La construcción de la matriz de coocurrencia depende de tres criterios: la condición, la dirección y la distancia. Condición: es la regla que debe cumplir un determinado píxel para poder ser cuantificado, usualmente esta condición es tomada como la ubicación del píxel de determinada intensidad que se encuentra en la dirección del vecino. Dirección: es la orientación espacial en la cual se evalúa la condición, como casos típicos se utilizan direcciones como 0 grados, 45 grados, 90 grados y 135 grados, (Gonzalez, & Woods, 2002). Distancia: es el número de píxeles que hay entre el par de píxeles evaluados, usualmente se utilizan 1, 3 y 5 píxeles de distancia. Se obtuvo una matriz de coocurrencia para cada una de las direcciones 0, 45, 90 y 135 grados, todos a una distancia fija de 1 píxel. La matriz final es la suma de las cuatro matrices de diferentes direcciones dividida en cuatro. Con base en la matriz de coocurrencia se obtienen los descriptores de textura. Se extrajeron 11 descriptores: Entropía, Uniformidad, Contraste, Correlación, Momento de diferencia de orden k, Momento inverso de diferencia de orden k, Máxima probabilidad, Homogeneidad local, Varianza, Directividad y Tendencia de Cluster. Estas características se evaluaron en cada una de las componentes R, G y B, obteniéndose así 33 características.

#### 3.1.3. Características de forma

La forma de los frutos de café variedad Colombia es elipsoidal-esférica, independiente de su estado de madurez. La forma elipsoidal se define con un centro y tres ejes de simetría, además todas las secciones planas de un elipsoide son elipses o círculos. En las primeras etapas la forma va variando de elipsoidal alargada, donde un eje de simetría es mucho mayor que los otros dos, a circular, donde los tres ejes son similares. En plena madurez, la forma es esférica es decir que los ejes de simetría son iguales. A medida que el fruto se sobremadura, la pulpa se va secando y la forma tiende otra vez a ser elipsoidal alargada. En una imagen de un fruto de café sólo se tiene información de una de las tres secciones de la elipsoide, con base en esta vista se obtienen dos de los tres ejes de simetría. En la Tabla 1, se presenta la lista de las 157 características de forma obtenidas. La descripción de estas características es explicada por [8].

**Tabla 1.** Características de forma

No. de Característica	CARACTERÍSTICA
1 a 36	Amplitud de los coeficientes de Fourier de la firma
37 a 72	Fase de los coeficientes de Fourier de la firma
73 a 82	Momentos centrales de la firma
83	Rectangularidad
84	Circularidad
85 a 120	Amplitud de los 36 descriptores de Fourier
121 a 156	Fase de los 36 descriptores de Fourier
157	Relación entre anchura y altura

**3.2. Selección de características**

La aplicación de los algoritmos de extracción de características permitieron formar un conjunto de 208 características que corresponden a medidas de forma, textura y color de un fruto de café. Los valores de las características se normalizaron para garantizar que todos los datos se mantengan dentro de un mismo rango de valores evitando que las que tienen valores más grandes tengan mayor influencia en la clasificación que aquellas con valores pequeños. Se aplicó una normalización no lineal dada por la ecuación (1), donde  $\tilde{x}_k$  corresponde a la media de los valores de la k-ésima característica y  $\sigma$  es la respectiva varianza. El i-ésimo valor normalizado  $\hat{x}_k$  se relaciona de forma no lineal con la constante r, pero en este caso no se investigó su incidencia sobre la normalización y se le dio el valor de 1. En la Ecuación 1, y es una variable auxiliar, utilizada para hallar el valor normalizado  $\hat{x}_k$ , en la segunda parte de la ecuación.

$$y = \frac{x_k - \tilde{x}_k}{r\sigma_k} \tag{1}$$

$$\hat{x}_k = \frac{1}{1 + \exp(-y)}$$

Después de normalizar el conjunto total de características extraídas se realiza un proceso de selección que permite determinar las características más adecuadas para la clasificación. El objetivo de la selección es reducir el número de características para simplificar tanto la implementación como la ejecución del clasificador. El conjunto final debe ser el más pequeño posible pero al mismo tiempo debe conservar al máximo la información discriminante. Muchas de las

características extraídas no ofrecen información que permita diferenciar entre clases, y otras tienen la misma información que otras. Se utilizaron dos métodos para la selección de las características uno que es la aplicación de la relación discriminante de Fisher como método univariado y el otro corresponde a un método multivariado. La explicación de dichos métodos se puede encontrar con detalle en (Sandoval & Prieto, 2007). El conjunto usado en la clasificación es el resultado de combinar los dos procesos de selección. Del total de 208 características se logró obtener un conjunto de 9 características que corresponden a 4 de textura, 3 de color y 2 de forma. En la tabla 2, se presentan las características seleccionadas, ordenadas según su prioridad en la discriminación.

**Tabla 2.** Las nueve características seleccionadas con los dos métodos utilizados.

ETIQUETA	CARACTERÍSTICA	TIPO
1 27	Momento de diferencia de orden k componente B	Textura
2 5	Momento de diferencia de orden k componente R	Textura
3 8	Homogeneidad local componente R	Textura
4 83	Amplitud de los 36 descriptores de Fourier 2	Forma
5 17	Momento inverso de diferencia de orden k componente G	Textura
6 167	Media componente RG de Colores Oponentes	Color
7 82	Amplitud de los 36 descriptores de Fourier 1	Forma
8 155	Media componente R de RGB	Color
9 171	Media componente I <sub>1</sub> de OHTA	Color

**4. Clasificación**

La técnica de clasificación implementada es estadística debido a la naturaleza de las muestras.

**4.1. Clasificador Bayesiano**

Este clasificador se fundamenta en la regla de Bayes del mínimo error (Duda, Hart, & Stork, 1998). Un objeto, con unas características determinadas, pertenece a una clase si la probabilidad de pertenecer a ésta clase es mayor que la probabilidad de pertenecer a cualquier otra clase, como se muestra en la Ecuación 3,

$$X \in \Omega_i \text{ si } P(\omega_i)p(X/\omega_i) > P(\omega_j)p(X/\omega_j) \tag{3}$$

donde  $\Omega$  es el espacio de características, que está dividido en regiones  $\Omega_i$ ,  $i = 1, 2, \dots, N$  donde N es el número de clases.  $P(\omega_i)$  es la probabilidad a priori por la cual un objeto con características X, pertenece a la clase  $\omega_i$  y  $p(X/\omega_i)$  es la función de probabilidad condicional de la clase  $\omega_i$  para X. En la práctica, las funciones de probabilidad no se conocen y por lo tanto se deben estimar. Para estimarlas, primero se asume la forma de la función de probabilidad, y luego se hallan

sus parámetros a partir del conjunto de entrenamiento. Se supone que las funciones de probabilidad de las clases están descritas por distribuciones normales o gaussianas como se muestra en la Ecuación 4,

$$p(\omega_i / X) = \frac{1}{(2\pi)^{\frac{d}{2}} \|C_i\|^{\frac{1}{2}}} e^{-\frac{1}{2}(X-\mu_i)^T (C_i)^{-1} (X-\mu_i)} \quad (4)$$

donde d es la dimensión del vector de características,  $\mu_i$  es el vector d-dimensional que contiene la media de las características en una clase  $w_i$  y  $C_i$  es la matriz de covarianza. Para obtener  $p(X/w_i)$ , sólo se necesita reemplazar los parámetros  $\mu_i$ ,  $C_i$ . Estos parámetros y la probabilidad a priori  $P(w_i)$ , se estiman a partir del conjunto de muestras. En la Ecuación 4. T indica que el vector es transpuesto.

**4.2. Evaluación**

Para estimar el error de clasificación se utilizó el método de validación cruzada (10-fold), (Moore, 2001). En la estimación del error se usa el conjunto de muestras disponible, el cual se divide en el conjunto de entrenamiento y el de prueba. El clasificador se diseña usando las muestras de entrenamiento y luego se evalúa obteniendo el error de clasificación para las muestras de prueba. Con base en el error obtenido se puede predecir el desempeño del clasificador ante nuevas muestras. Para obtener una medida confiable del desempeño, el conjunto de muestras debe ser lo suficientemente grande y, los conjuntos de entrenamiento y de prueba deben ser independientes.

**4.3. Análisis de los resultados**

La información de la prueba de clasificación se dispone en una matriz de confusión. La matriz de confusión es una matriz cuadrada cuyo orden es el número de clases. En las columnas se presentan las clases reales mientras que en las filas se presentan las clases asignadas por el clasificador; en la Tabla 3, se muestra una matriz de confusión para dos clases. La suma vertical muestra la distribución real de las clases, mientras que la suma horizontal muestra la distribución de las clases producida por el clasificador. El análisis de los clasificadores se hace con base en la matriz de confusión de donde se obtienen los indicadores de desempeño.

**Tabla 3. Matriz de confusión**

	Verdadero	Falso
Verdadero	Verdadero Positivo (VP)	Falso Positivo (FP)
Falso	Falso Negativo (FN)	Verdadero Negativo (VN)

El desempeño de la clasificación se puede verificar con los indicadores de Proporción de verdaderos positivos, con el cual se puede medir la efectividad del clasificador, y la proporción de verdaderos negativos, que nos muestra la especificidad del clasificador.

Proporción de verdaderos positivos:  $PVP = VP / (VP+FN)$ . (Efectividad).

Proporción de verdaderos negativos:  $PVN = VN / (VN+FP)$ . (Especificidad).

Curvas ROC. El espacio ROC (Receiver Operating Characteristic) define un sistema de coordenadas usadas para visualizar el desempeño del clasificador. Las curvas ROC presentan el compromiso entre efectividad y especificidad del clasificador, un aumento en la sensibilidad está acompañado por un decremento en la especificidad. Es decir, las curvas ROC muestran la relación entre las muestras clasificadas adecuadamente (PVP, Proporción de Verdaderos Positivos) y las muestras que no pertenecen a la clase pero se clasificaron como si lo fueran (PFP Proporción de Falsos Positivos). En el espacio ROC la PFP se dibuja como variable independiente y la PVP como variable dependiente. Cada clasificador es representado por el punto (PFP, PVP).

Para dibujar la curva ROC se construye la línea convexa formada por los puntos (PFP, PVP) de los clasificadores que se estén evaluando ( $PFP=1-PVN$ ), junto con los puntos de los clasificadores triviales (0,0) y (1,1). La curva más cercana a los bordes izquierdo y superior en el espacio ROC, es la prueba más acertada porque significa que hay mayor acierto. La curva que más se acerque a la diagonal de 45 grados en el espacio ROC, es la prueba menos acertada. El mejor sistema de entrenamiento es el que produce un conjunto de clasificadores que maximice el área bajo la curva.

**5. Resultados**

Para evaluar el desempeño del sistema de clasificación se realizó validación cruzada para un conjunto de 800 imágenes, que corresponden a 100 imágenes por etapa de maduración.

**5.1 Resultados con el Clasificador Bayesiano**

En la Tabla 4 se muestra la matriz de confusión para el mejor clasificador bayesiano. La proporción de verdaderos positivos, es decir el número de frutos de una etapa de maduración que fueron clasificados como de la etapa correspondiente, según la matriz de confusión es de 96.88%. Esto significa que hay un error de clasificación del 3.12%. Aunque los resultados son buenos, el proceso de clasificación exige un desempeño del 98%. El número de frutos de café que pertenecen a una clase y fueron clasificados como de

otra, corresponde a la proporción de falsos positivos, para nuestro caso esa proporción es del 2.5%. El mejor clasificador tendría un índice cercano a 0.

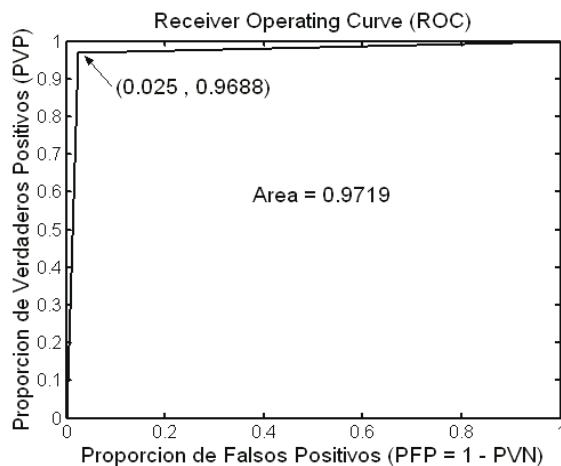
**Tabla 4.** Matriz de confusión. Clasificador Bayesiano

	1	2	3	4	5	6	7	8	Total
1	100	0	0	0	0	0	0	0	100
2	0	100	0	0	0	0	0	0	100
3	0	0	100	0	0	0	0	0	100
4	0	0	0	100	0	0	0	0	100
5	0	0	0	0	100	5	0	0	105
6	0	0	0	0	0	90	10	0	100
7	0	0	0	0	0	5	90	5	100
8	0	0	0	0	0	0	0	95	95
Total	100	100	100	100	100	100	100	100	800

**5.2. Curvas ROC**

En la Figura 2, se presenta la curva ROC del clasificador bayesiano implementado. En una gráfica ROC el mejor clasificador es aquel que presenta un área bajo la curva igual a 1 mientras que el peor 0.5. En éste caso el área bajo la curva es de 0.9719 que corresponde a un valor cercano al ideal, mostrando un buen compromiso entre efectividad y especificidad del clasificador implementado. Una relación adecuada entre efectividad y especificidad, permite evaluar la generalidad del clasificador, es decir, que clasifique adecuadamente tanto las muestras usadas en el entrenamiento y prueba así como las muestras nuevas o muestras con algún tipo de error.

**Figura 2.** Curva ROC del clasificador



**6. Conclusiones**

En este trabajo se presenta un sistema de clasificación bayesiana que usa características físicas extraídas a partir de una imagen para clasificar café según la etapa de maduración en la que se encuentra. Se clasifican los frutos de café en ocho etapas de maduración diferentes, que cubren todo el proceso de maduración

desde las primeras etapas de verdes hasta los frutos sobremaduros y secos. Las características que se usaron corresponden a 4 características de textura, 3 de color y 2 de forma, que se obtuvieron mediante la implementación de algoritmos de visión artificial. La técnica de clasificación que se usa es de tipo estadístico maximizando la información dada por las muestras. Los resultados obtenidos muestran un desempeño del 96.88%. El clasificador muestra una buena capacidad de generalización, que corresponde a el compromiso entre efectividad y especificidad, como se observa en la curva ROC que genera un área del 0.9719 muy cercano al 1 ideal. Para alcanzar el desempeño que requiere la aplicación que es del 98%, se necesita profundizar en el estudio tanto de los algoritmos usados en la extracción de las características, así como mejorar el algoritmo de clasificación, ya sea cambiando parámetros o implementando otra técnica más compleja que genere un mejor rendimiento.

Las 9 características utilizadas en el clasificador fueron seleccionadas de un conjunto inicial de 208 características, usando dos métodos de selección. Las características que mostraron mejor desempeño corresponden a las de textura, indicando que para la diferenciación de las etapas de maduración no sólo es importante el valor del color, como en la clasificación manual, sino que se hace relevante su distribución sobre la superficie. Las características de forma permiten distinguir principalmente entre las etapas iniciales y finales de las etapas intermedias, debido a la redondez que van adquiriendo en la maduración, y que es plena en las etapas centrales.

Las medidas de textura se hicieron usando la matriz de concurrencia, dicha matriz se construyó usando algunos parámetros con sus valores por defecto. Con el fin de aprovechar la alta capacidad discriminante que mostraron las características de textura con los métodos de selección de características, se genera una necesidad de estudiar los parámetros de la matriz de concurrencia como son: la condición, distancia y orientación.

**Bibliografía**

Wallis, J., Montoya, E., Vélez, J. & Oliveros, C. (2004). Calidad y eficacia de dos métodos no selectivos de recolección manual de café (*Coffea arabica*). *Cenicafé*, 55(1): 45 – 51.

López, H., Oliveros, C. & Ramirez, C. (2006). Disminución del costo unitario de la cosecha de café con el empleo de un método de recolección manual asistido. *Cenicafé*, 57(4): 262 – 273.

Marín, S., Arcila, J., Montoya, E. & Oliveros, C. (2003). Cambios físicos y químicos durante la maduración

- del fruto de café (*Coffea arabica* L. var Colombia). *Cenicafé*, 54(3): 208 – 225.
- Puerta, G. (1998). Calidad en taza de las variedades de *Coffea arabica* L. cultivadas en Colombia. *Cenicafé*. 49(4): 265 – 278.
- Isaza, L., Montoya, E., Vélez, J. & Oliveros, C. (2006). Evaluación de la concentración de los frutos maduros de café empleando técnicas no selectivas de recolección manual. *Cenicafé*, 57(4): 274 – 287.
- Sandoval, Z. & Prieto F. (2007). Caracterización de café cereza empleando técnicas de visión artificial. *Rev.Fac. Nal.Agr.Medellín*. Vol.60,No.2. p.4105- 4127.
- Betancur, J., Prieto, F. & Osorio, G. (2006). Segmentación de Frutos de Café mediante métodos de Crecimiento de Regiones. *Revista Facultad Nacional de Agronomía, Universidad Nacional de Colombia Sede Medellín, Medellín, Colombia*. Vol. 59, No. 1, Páginas 3311-3333.
- Gonzalez, R. & Woods, R. (2002). Tratamiento Digital de imágenes. Addison-Wesley/díaz de Santos. Segunda Edición.
- Duda, R.; Hart, P. & Stork, D. (1998). *Pattern Classification*. John Wiley and Sons. USA.
- Moore, A. (2001). Cross-validation for detecting and preventing overfitting. School of Computer Science Carnegie Mellon University.