

SELECCIÓN DE CARACTERÍSTICAS ORIENTADA A SISTEMAS DE RECONOCIMIENTO DE GRANOS MADUROS DE CAFÉ.

Feature selection oriented to recognition systems of ripe red coffee beans.

RESUMEN

En este trabajo se presenta una comparación de técnicas para la selección y extracción de características en el entrenamiento de sistemas de reconocimiento de frutos de café (verde y maduro). Las técnicas de selección evaluadas usan criterios *filtro* y *wrapper* con algoritmos búsqueda heurística. Las funciones de costo son análisis multivariado de varianza y clasificador bayesiano sobre distribuciones gaussianas, respectivamente. Se utiliza el análisis de componentes principales como técnica de extracción. Adicionalmente se realiza el preprocesamiento estadístico de los datos, el cual es necesario para dar validez a las suposiciones de las técnicas.

PALABRAS CLAVES: Búsqueda heurística, identificación frutos de café, MANOVA, PCA, reconocimiento de patrones, selección de características.

ABSTRACT

This work presents a comparison among feature selection and feature extraction techniques in training of recognition systems of ripe (green and red) coffee bean. The evaluated feature selection techniques use filter and wrapper criteria with heuristic search algorithms. The cost functions are multivariate analysis of variance and bayes classifier over gaussian distributions, respectively. We use principal component analysis as a feature extraction technique. Furthermore, we carry out a statistical data preprocessing, which is necessary for checking the assumptions of the employed methods.

KEYWORDS: Feature selection, heuristic search, MANOVA, pattern recognition, PCA, ripe coffee bean identification.

1. INTRODUCCIÓN

Las técnicas de aprendizaje de máquina han sido empleadas de manera satisfactoria, en un amplio rango de aplicaciones como: el control de procesos, análisis de datos, toma de decisiones, entre otras. El reconocimiento de patrones es una de las tareas principales consideradas dentro del aprendizaje de máquina y está directamente relacionado con: asignar etiquetas de pertenencia a una clase de observaciones de algún fenómeno, basándose en el análisis de un conjunto de datos previamente dado. Típicamente, un sistema de reconocimiento de patrones puede ser descrito como se presenta en la Figura 1. Inicialmente se tienen mediciones de un conjunto de objetos, posteriormente se aplica una serie de adaptaciones conocida como representación o caracterización de los objetos. Dichas representaciones o conjuntos de características pueden estar sujetas a posteriores transformaciones antes de ser introducidas en la etapa de clasificación (identificación), en la que se asignan las etiquetas o se establecen agrupaciones de los datos que conduzcan a la toma de decisiones. Sin embargo, el desarrollo de las etapas no es unidireccional,

GENARO DAZA S. *

Ingeniero Electrónico, M.Sc.
Universidad Nacional de Colombia
sede Manizales
gdazas@unal.edu.co

LUIS G. SÁNCHEZ G. *

Ingeniero Electrónico.
Universidad Nacional de Colombia
sede Manizales
lgsanchezg@unal.edu.co

JULIO F. SUÁREZ C.

Estadístico, M.Sc.
Profesor.
Universidad Nacional de Colombia
sede Manizales
jfsuarezc@unal.edu.co

***GRUPO DE CONTROL Y PROCESAMIENTO DIGITAL DE SEÑALES (GCYPDS).**

debido a que cada una de ellas requiere de retroalimentación para mejorar el desempeño de todo el sistema.

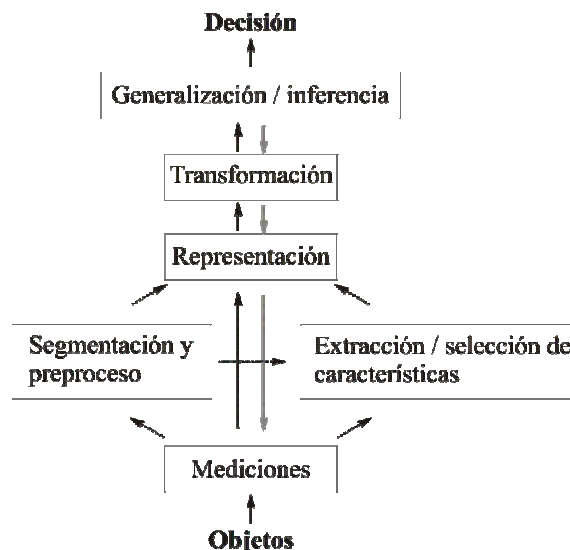


Figura 1. Sistema de reconocimiento de patrones.

El *entrenamiento* de estos sistemas, como parte básica del reconocimiento automático de patrones en procesos industriales, mejora si se realiza la *extracción / selección* de características de forma adecuada. En la identificación de patrones, es usual que los procedimientos de entrenamiento, posean un alto número de características, es decir, una alta dimensión del espacio inicial de entrenamiento, lo cual conduce a un tiempo de procesamiento significativo, y a la disminución en el rendimiento del sistema de clasificación; por ende, decae la eficiencia de los procesos productivos y se incrementan los costos asociados a éste. En este sentido, la selección efectiva de características debe ser tomada en cuenta, en orden de distinguir aquellas variables que ofrezcan mayor capacidad discriminante sin afectar la precisión en la clasificación.

En la selección efectiva de características, existen dos elementos importantes: una función de evaluación que puede ser tipo *filtro* o *wrapper* y un procedimiento iterativo de búsqueda completa o heurística [1].

Las funciones de costo tipo filtro son aquellas que no consideran el algoritmo de clasificación como generador de la medida que determina la selección de las características para el proceso de reconocimiento de patrones, es decir, filtran características irrelevantes antes que ocurra la etapa de clasificación. A menudo, las funciones de costo tipo filtro se basan en métodos estadísticos que emplean técnicas de análisis univariado o multivariado. Por otra parte, las funciones tipo wrapper emplean como función de evaluación el propio algoritmo de clasificación para eliminar las características menos influyentes [2].

En cuanto al procedimiento de búsqueda completa, éste asegura hallar el subconjunto óptimo de características, de acuerdo a la función de evaluación empleada, después de realizar un procedimiento exhaustivo de evaluación sobre todo el espacio de entrenamiento, mientras que la búsqueda heurística se basa en reglas empíricas que están orientadas a reducir la complejidad computacional, evitando disminuir el rendimiento del sistema. Los métodos heurísticos de búsqueda requieren de una condición de parada para prevenir que la búsqueda de subconjuntos de características se vuelva exhaustiva.

Los procedimientos de extracción de características son formas alternativas que buscan mejorar el rendimiento del sistema de clasificación a partir del entrenamiento.

Las técnicas de análisis multivariado también han sido aplicadas en este contexto, particularmente, el *análisis de componentes principales* (PCA – *Principal Component Analysis*), tiene por objetivo principal reducir la dimensión de un conjunto de características, conservando la mayor cantidad de información que sea posible. Esto se logra mediante la transformación a un nuevo conjunto de variables las cuales son no correlacionadas y se

ordenan de modo tal que unas pocas (las primeras) retengan la mayor cantidad de variación presente en el conjunto original de características [3].

Una aplicación de interés local para los sistemas industriales de reconocimiento de patrones es el proceso de identificación de frutos de café según su etapa de maduración. Actualmente la recolección de café es realizada de forma manual dando como resultado un café poco homogéneo a la hora de iniciar el proceso de beneficio. La poca selectividad del café según su etapa de maduración; granos verdes y maduros, entre otros, disminuye su calidad. El estado de maduración en que se encuentra el café produce sabores y aromas diferentes en el producto final. Esta disminución en la calidad tiene como consecuencia la reducción del precio en el mercado mundial. La identificación manual de frutos puede ser reemplazada por sistemas automatizados de reconocimiento de patrones empleando visión artificial, dado que estos sistemas se basan en el análisis de imágenes digitales de diferentes características del grano de café [4].

Por tanto, en este trabajo se presenta una comparación de técnicas para la selección y extracción de características en el entrenamiento de sistemas de reconocimiento de frutos de café (verde y maduro), empleando selección de características con base en funciones de costo filtro y wrapper, junto con algoritmos de búsqueda heurística, y extracción de características por medio de la técnica PCA.

Los algoritmos de extracción / selección se ejecutan sobre características correspondientes a los valores de intensidad de los píxeles en RGB de imágenes de granos de café. Las técnicas son comparadas con base en el error de validación de un clasificador de tipo bayesiano sobre distribuciones gaussianas. Además se analiza el desempeño del sistema de reconocimiento aplicando las diferentes técnicas de extracción / selección, por medio de curvas *ROC – Receiver Operating Characteristic* y *DET – Detection Error Trade-off*.

2. EXTRACCIÓN / SELECCIÓN DE CARACTERÍSTICAS

2.1. Algoritmo filtro con base en análisis multivariado de varianza

Dado un espacio inicial de características de dimensión p para el entrenamiento del sistema, se realiza la evaluación sucesiva de subconjuntos de variables empleando como función de evaluación el *análisis multivariado de varianza* (MANOVA – *Multivariate Analysis of Variance*). La generación de subconjuntos de búsqueda se realiza con base en la técnica de *selección secuencial hacia adelante* (SFS- *Sequential Forward Selection*). En MANOVA el criterio de evaluación

estadístico es la separabilidad entre clases, que se realiza mediante una prueba de hipótesis sobre la igualdad o desigualdad de los vectores de promedios entre las clases. Se asume que los datos están generados con base en el siguiente modelo:

$$\mathbf{x}_{ij} = \mathbf{m}_k + \varepsilon_{ij}, \mathbf{m}_k = \mathbf{m} + \boldsymbol{\alpha}_k \quad (1)$$

siendo j la observación y k la clase, \mathbf{m}_k es el vector de medias para cada clase y ε_{kj} es la respectiva perturbación del modelo, \mathbf{m} es la media global de las clases y $\boldsymbol{\alpha}_k$ es la perturbación sobre esta media global.

La comparación de los vectores de medias de las k clases para encontrar diferencias significativas, se realiza mediante la prueba hipótesis:

$$\begin{aligned} H_0 : \mathbf{m}_1 = \mathbf{m}_2 = \dots = \mathbf{m}_L \\ H_1 : \exists \text{ al menos un par } \mathbf{m}_k \neq \mathbf{m}_i; \forall k, i \in \{1, \dots, L\} \end{aligned} \quad (2)$$

donde L es el número de clases.

La estadística de Wilks es comúnmente usada al interior de MANOVA, para probar la hipótesis H_0 , que corresponde a la relación de verosimilitud dada por:

$$\Lambda = \frac{|\mathbf{E}|}{|\mathbf{E} + \mathbf{H}|} \quad (3)$$

la cual es conocida como Λ de Wilks. Siendo \mathbf{H} la matriz de hipótesis que puede entenderse como una medida de dispersión entre la media de las clases, mientras que la matriz de error \mathbf{E} se relaciona con la medida de dispersión entre las observaciones para cada clase. La hipótesis H_0 se rechaza si la dispersión entre los patrones es mayor que la dispersión de las observaciones dentro de los patrones, y así, $\Lambda \in [0,1]$ tiende a cero. Por otro lado, el Λ de Wilks puede ser similar a un estadístico F , pero de manera inversa. Un valor grande del estadístico F rechaza H_0 [5].

El siguiente es el algoritmo heurístico *SFS* basado en MANOVA, para el caso particular se evalúa la probabilidad acumulada sobre la distribución F , que permite comparar conjuntos de variables de dimensión diferente:

Algoritmo 1. MANOVA progresivo

- 1: Calcular el estadístico F (transformación del Λ de Wilks) para los subconjuntos de una sola característica. A partir de estos valores se escoge la característica que reporte el mayor valor de estadístico F y se calcula su respectivo valor de probabilidad acumulado sobre la distribución F .
- 2: Construir subconjuntos de dos dimensiones, combinando la característica previamente escogida en el paso 1 con cada una de las características restantes (aplicando *SFS*). Cada uno de estos subconjuntos se evalúa a través de la prueba de Wilks y se actualiza su respectivo estadístico F .
- 3: Seleccionar el subconjunto bidimensional con el mayor estadístico F . Calcular su respectivo valor de probabilidad acumulada. Dicho valor debe exceder el valor calculado en

el paso 1 para que el subconjunto pueda ser analizado, de otra manera, la búsqueda concluye.

- 4: Construir subconjuntos de características agregando una característica al subconjunto actual (*SFS*). Estos nuevos grupos de análisis corresponden al subconjunto seleccionado en el paso 3 y cada una de las características restantes.
 - 5: Retornar al paso 3 y actualizar el subconjunto utilizando el mismo criterio. Continuar con el paso 4 repitiendo el proceso de actualización una y otra vez. El algoritmo se detiene como se menciona el paso 3 cuando las características agregadas no incrementan el valor de probabilidad acumulada. Así el tamaño del subconjunto final de características es p' donde $p' \leq p$. De esta forma, se pueden seleccionar las características que de manera conjunta son más discriminantes.
-

Cabe anotar, si durante la ejecución de Algoritmo 1 se llega a presentar que en los cálculos aparece $\Lambda \rightarrow 0/0$, se asume la dependencia lineal en el subconjunto evaluado, luego la probabilidad acumulada es cero y se desecha la característica incrementada al subconjunto en análisis.

2.2. Algoritmo wrapper con base en clasificador bayesiano sobre distribuciones gaussianas

La selección de características por medio de un esquema wrapper puede realizarse empleando la estrategia heurística de búsqueda *SFS* y utilizando como función de evaluación el porcentaje de acierto de validación en clasificador bayesiano. El objetivo es encontrar un subconjunto de características que maximice el criterio de evaluación.

2.3. Análisis de componentes principales – PCA

El análisis de componentes principales ha sido la tendencia dominante para el análisis de datos en un gran número de aplicaciones. Su atractivo recae en la simplicidad y capacidad de reducción de dimensión, minimizando el error cuadrático de reconstrucción que se obtiene a partir de una combinación lineal de variables latentes conocidas como componentes principales. Los parámetros del modelo pueden ser calculados directamente de la matriz de datos centralizada \mathbf{X} bien sea por descomposición en valores singulares o la diagonalización de la matriz de covarianza (positiva semidefinida) [3]. Sea \mathbf{x}_i el i -ésimo vector de observación columna de longitud p , siendo $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)^T$, la matriz de rotación \mathbf{U} con la que se calculan las p' componentes principales denotadas por \mathbf{z} que resumen \mathbf{x} .

$$\mathbf{z} = \mathbf{U}^T \mathbf{x} \quad (4)$$

\mathbf{U} puede ser calculada a partir de los primeros p' valores propios de $\mathbf{X}^T \mathbf{X}$, esto es,

$$\mathbf{X}^T \mathbf{X} \mathbf{U} = \mathbf{U} \Lambda \quad (5)$$

Para este trabajo en particular, la dimensión del subespacio de representación que contiene p' componentes principales es determinada con base en el error de entrenamiento (o aparente) calculado para un clasificador bayesiano sobre distribuciones gaussianas.

3. MARCO EXPERIMENTAL

3.1. Base de datos

Se evalúa una base de datos que contiene 4000 imágenes de granos de café, de las cuales 2000 imágenes son de granos verdes (Figura 2) y las 2000 restantes son de granos maduros (Figura 3). Las imágenes fueron tomadas en condiciones de iluminación variable, pero con fondo controlado, forzando que cada escena sólo contenga un grano de café. Las características extraídas a partir de las imágenes son los valores de intensidad en los tres canales de color del modelo RGB cuantizados a 8 bits por canal, para un área del grano de 20×20 píxeles, tal como se muestra en la Figura 4 y Figura 5, para frutos verdes y maduros respectivamente. En consecuencia el conjunto inicial de características posee 1200 variables.

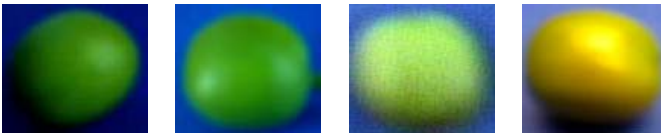


Figura 2. Granos verdes.

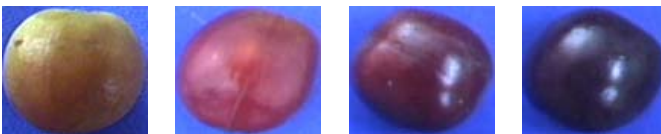


Figura 3. Granos maduros.

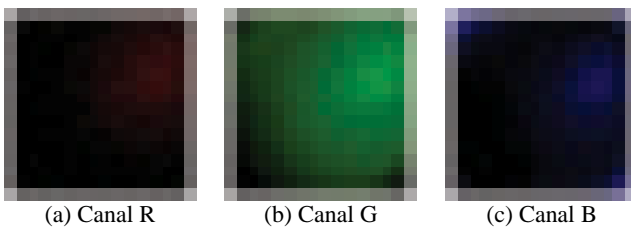


Figura 4. Características grano verde.

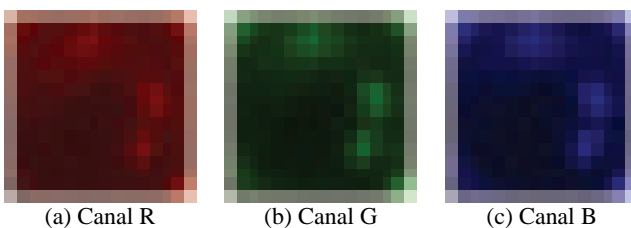


Figura 5. Características grano maduro.

Para comprobar que los resultados obtenidos del entrenamiento del sistema de identificación de frutos de café se pueden generalizar sobre nuevas observaciones,

se construyen dos conjuntos de observaciones disyuntos (conjunto de entrenamiento y conjunto de clasificación). El conjunto de entrenamiento consta de 1000 observaciones independientes e idénticamente distribuidas por clase. Este conjunto sólo es empleado en las etapas de preproceso y de extracción / selección de variables. El conjunto de clasificación contiene las 2000 observaciones restantes. Este conjunto, es utilizado para entrenar y validar el clasificador final por medio de una estrategia de validación cruzada *leave-M-out*.

3.2. Preprocesamiento de los datos

El preprocesamiento de los datos tiene como objetivo la disminución de la influencia, y en lo posible, la eliminación de los errores ocasionados por fallas esporádicas debidas a malos ajustes en etapas anteriores a la generación de características. El preprocesamiento de los datos sirve también como control de la homogeneidad de las propiedades estadísticas inherentes a las características del fenómeno aleatorio en análisis [6].

En particular, el preproceso de los datos consiste inicialmente en eliminar las características cuyo valor de kurtosis no se encuentran en un intervalo adecuado, que refleje la gaussianidad de la variable de análisis. El intervalo adecuado seleccionado es $2 \leq \text{kurtosis} \leq 4$. Si hay pocos datos atípicos (*outliers*) muy alejados del resto, la variabilidad de las desviaciones será grande, por lo que el coeficiente de kurtosis será alto. Además, cuando el valor de kurtosis es bajo (tiende a uno), es posible que la característica refleje la mezcla de dos o más poblaciones. Por otra parte, coeficientes de kurtosis alrededor de 3 corresponden a distribuciones normales [7].

3.3. Extracción / Selección de características

El conjunto de variables preprocesadas obtenido a partir del procedimiento anteriormente descrito se utiliza como entrada para cada una de las técnicas de extracción / selección, selección filtro con MANOVA, selección wrapper y extracción PCA (presentadas en la Sección 2). Los conjuntos de variables resultantes de cada una de las técnicas aplicadas son utilizados en el clasificador final.

3.4. Clasificación y evaluación

Como algoritmo de decisión entre frutos maduros y verdes se emplea un clasificador bayesiano sobre distribuciones gaussianas que corresponde a una función cuadrática. Con el objetivo de evaluar el desempeño de cada uno de los conjuntos de variables extraídas / seleccionadas, se comparan medidas de discriminancia y confiabilidad derivadas del clasificador ya mencionado. En particular, como medida de discriminancia se emplea la tasa de aciertos en validación y como medida de confiabilidad se considera el intervalo de confianza para

la tasa de aciertos de validación. Para la estimación de los errores de validación se emplea la estrategia de validación cruzada *leave-M-out*, la cual consiste en generar L conjuntos, que corresponden a particiones aleatorias del conjunto de N observaciones en pares de entrenamiento-validación donde se retienen M observaciones para validar (se entrena con $N - M$ observaciones). En este tipo de evaluación los errores de validación no necesariamente corresponden a conjuntos disyuntos [8]. Para este trabajo $N = 2000$ (1000 por clase), $M = 600$ (300 por clase) y $L = 100$.

También, se analiza la curva ROC la cual caracteriza la sensibilidad en la detección de una clase en particular (granos maduros) y el porcentaje de observaciones que se clasifican como pertenecientes a la clase de interés sin serlo (granos verdes clasificados como maduros) [9]. Además, se considera la curva DET en la cual estudia el desempeño de clasificación con respecto a las observaciones que fueron mal clasificadas en ambas clases (traslape de las clases), sin enfocar el análisis sobre una clase en particular [10].

4. RESULTADOS Y DISCUSIÓN

Inicialmente, en la etapa de extracción de características se construyó un conjunto de 1200 variables, conjunto que fue reducido en la etapa de preprocesamiento de datos en un 44.08% aproximadamente, obteniéndose un nuevo subconjunto conformado por 671 características, cada una de ellas con distribución normal. Sin embargo, el subconjunto actual de variables es aún muy grande y aplicar un clasificador simple (tipo Bayesiano) no es posible; por tanto, se hace necesario la etapa de extracción / selección de variables. La Tabla 1 resume los resultados de la reducción de dimensión obtenidos por cada una de las estrategias presentadas en la Sección 2 y las respectivas tasas de acierto e intervalos de confianza para $\alpha = 0.05$ sobre la distribución t .

Técnica extracción / selección	Número de variables o componentes	Tasa de acierto promedio	Intervalo de confianza
Filtro (MANOVA)	3	92.95 %	92.78 % a 93.12 %
Wrapper	9	97.56 %	97.45 % a 97.67 %
PCA	25	98.5 %	98.41 % a 98.6 %

Tabla 1. Resultados de reducción, discriminacia y confiabilidad.

De los resultados presentados en la Tabla 1 es posible observar el compromiso entre el error de validación y la dimensión del espacio seleccionado siendo el conjunto con más variables el de mayor tasa de acierto en la validación. Los intervalos de confianza presentados establecen diferencias significativas entre los porcentajes de acierto promedio para los tres esquemas de extracción

/ selección desarrollados en este trabajo. Al indagar sobre las variables escogidas se encontró que los tres canales del RGB aportan información sobre la madurez del grano. En la Figura 6 se presenta la distribución de las observaciones sobre el espacio de las 3 características seleccionadas al utilizar la técnica de filtro MANOVA, como variables seleccionadas se obtuvo un píxel cercano al centro del área de trabajo de la imagen de 20×20 píxeles en cada uno de los canales R, G y B. Las observaciones más alejadas en cada una de las 2 agrupaciones de clase (ver Figura 6) corresponden a frutos en los cuales se presenta saturación del sensor a causa de la iluminación no controlada.

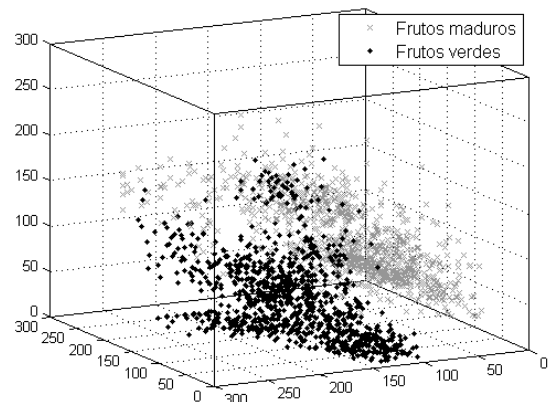


Figura 6. Espacio de características seleccionado utilizando la técnica tipo filtro (MANOVA) para el conjunto de observaciones de entrenamiento.

Por otra parte, las curvas de desempeño de las Figuras 7 y 8 resaltan la superioridad de los resultados obtenidos con la técnica PCA. Particularmente de la Figura 7 se puede ver como al utilizar PCA se logra una mayor sensibilidad en el sistema de identificación, por tanto hay una mayor capacidad de detectar correctamente los frutos maduros, manteniendo un muy bajo porcentaje de falsas aceptaciones. Por ejemplo si el sistema detecta un 95% de los granos maduros la cantidad de granos verdes clasificados como maduros es menor que al 0.2%. Dichos resultados son ratificados en la Figura 8, donde además, se observa a partir de las pendientes de las curvas que la extracción / selección de características realizada con PCA, produce un nuevo espacio de características en el cual las observaciones se ajustan de mejor manera al modelo de clasificación bayesiana sobre distribuciones gaussianas empleado.

La interpretación conjunta de las tres metodologías de extracción / selección sugiere la utilización de promedios en los píxeles de cada uno de los canales RGB como caracterización para una posible implementación de un sistema industrial de reconocimiento de frutos maduros. Lo que trae como ventaja una evidente simplificación en la solución del problema de identificación, posibilitando

el desarrollo de un hardware sencillo, económico, confiable y eficaz.

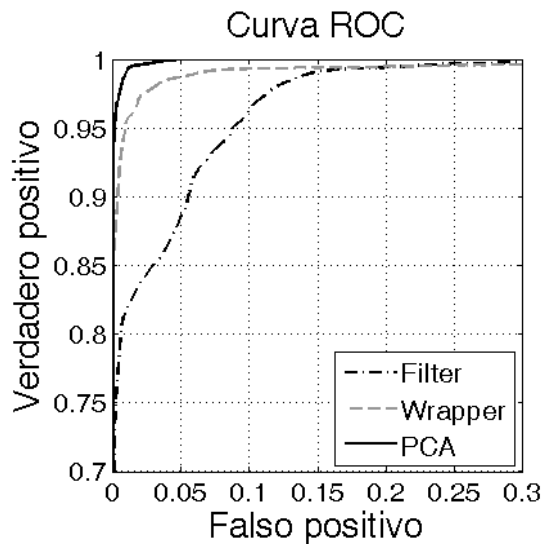


Figura 7. Curva ROC. Compara los resultados de clasificación obtenidos luego de emplear tres esquemas diferentes de extracción / selección de características.

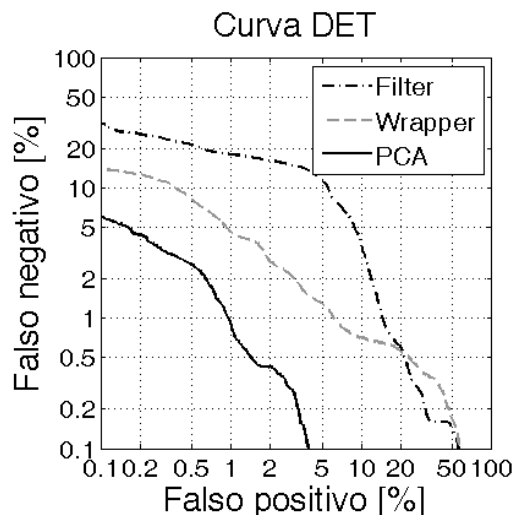


Figura 8. Curva DET. Compara los resultados de clasificación obtenidos luego de emplear tres esquemas diferentes de extracción / selección de características.

5. CONCLUSIONES

En este artículo se presentaron tres metodologías de extracción / selección de características (reducción de dimensión) las cuales permitieron simplificar el problema de identificación de frutos maduros de café, disminuyendo la cantidad de información a analizar a fin de obtener un adecuado desempeño del sistema. Entre las técnicas comparadas, PCA reportó el mejor desempeño y confiabilidad. Además, el análisis de los resultados de la etapa de extracción / selección sugiere la utilización de promedios de los píxeles de cada uno de los canales RGB, como caracterización para una posible implementación de un sistema industrial de

reconocimiento de frutos maduros, motivo por el cual se concluye que no es necesario un sofisticado sistema de adquisición de registros fotográficos y procesamiento de imágenes; un sistema de sensores que den idea de la intensidad de los colores R, G y B que refleja el fruto verde o maduro sería suficiente.

6. AGRADECIMIENTOS

Este trabajo se realiza en el marco del proyecto "Técnicas de computación de alto rendimiento en la interpretación automatizada de imágenes médicas y bioseñales", financiado por la Dirección de Investigaciones DIMA, de la Universidad Nacional de Colombia Sede Manizales.

7. BIBLIOGRAFÍA

- [1] A. L. Blum and P. Langley, "Selection of relevant features and examples in machine learning," *Artificial Intelligence*, vol. 97 (1), pp. 245-271, Dec. 1997.
- [2] A. R. Webb, *Statistical Pattern Recognition*, 2nd ed., Indianapolis, IN, USA: John Wiley & Sons Ltd., 2002, pp. 305-318.
- [3] I. T. Jolliffe, *Principal Component Analysis*, 2nd ed., New York, NY, USA: Springer, 2002, pp. 29-62.
- [4] J. E. Hernández, "Clasificación de frutos de café según su etapa de maduración basada en redes neuronales artificiales," Trabajo de grado, Dept. de Ing. Eléctrica, Electrónica y Computación, Universidad Nacional de Colombia, 2004.
- [5] A. C. Rencher, *Methods of Multivariate Analysis*, 2nd ed., Hoboken, NJ, USA: Wiley Interscience, 2002, pp. 156-247.
- [6] J. F. Suárez, G. Daza, L. G. Sánchez y G. Castellanos, "Selección efectiva de características en el entrenamiento de sistemas industriales de reconocimiento," en *2006 XVI simposio de estadística*, versión Electrónica CD.
- [7] D. Peña, *Análisis de Datos Multivariantes*, C. Fernández, Ed., Madrid, España: McGraw-Hill, 2002, pp. 120-132.
- [8] The MathWorks. (2007). *Bioinformatics Toolbox 2 User's guide*. The MathWorks, Inc., Natick, MA, USA. [Online]. Available: http://http://www.mathworks.com/access/helpdesk/help/pdf_doc/bioinfo/bioinfo_ug.pdf
- [9] R. O. Duda, P. E. Hart and D. G. Stork, *Pattern Classification*, 2nd ed., Hoboken, NJ, USA: Wiley Interscience, 2000, pp. 20-83.
- [10] N. Sáenz-Lechón, J. I. Godino-Llorente, V. Osma-Ruiz and P. Gómez-Vilda, "Methodological issues in the development of automatic systems for voice pathology detection," *Biomedical Signal Processing and Control*, vol. 1, pp. 120-128, 2006.