



UNIVERSIDAD
DE PIURA

REPOSITORIO INSTITUCIONAL
PIRHUA

EVALUACIÓN DE LA CALIDAD DE LOS GRANOS DE CACAO POR IMÁGENES HIPERESPECTRALES BASADAS EN TÉCNICAS DE LÓGICA DIFUSA

Juan Soto-Bohórquez

Piura, junio de 2016

FACULTAD DE INGENIERÍA

Máster en Ingeniería Mecánico – Eléctrica con mención en Sistemas
Energéticos y Mantenimiento

Soto, J. (2016). *Evaluación de la calidad de los granos de cacao por imágenes hiperespectrales basadas en técnicas de lógica difusa* (Tesis de Máster en Ingeniería Mecánico – Eléctrica con mención en Sistemas Energéticos y Mantenimiento). Universidad de Piura. Facultad de Ingeniería. Piura, Perú.



Esta obra está bajo una [licencia](#)
[Creative Commons Atribución-](#)
[NoComercial-SinDerivadas 2.5 Perú](#)

[Repositorio institucional PIRHUA – Universidad de Piura](#)

UNIVERSIDAD DE PIURA
FACULTAD DE INGENIERÍA



**“Evaluación de la calidad de los granos de cacao por imágenes hiperespectrales
basadas en técnicas de lógica difusa”**

**Tesis para optar el Grado de Master
en Ingeniería Mecánico – Eléctrica con
Mención en Sistemas Energéticos y Mantenimiento**

Juan Carlos Soto Bohórquez

Piura, junio 2016

A mi amada esposa por acompañarme
durante todo este arduo camino.
A mi madre quien me apoya
todo el tiempo.

Prólogo

El Perú es uno de los países productores de cacao criollo y orgánico. El cual es reconocido por su gran calidad, considerado fino y de aroma; por ello se emplea en la fabricación de los chocolates más finos.

Los granos de cacao provienen de la vaina de un árbol conocido como Theobroma cacao. Convencionalmente, existen 3 variedades: Criollo, Trinitario y Forastero.

Actualmente el control de calidad de los granos de cacao se realiza de manera artesanal, controlando la calidad del producto manualmente, realizado por personas basadas en su experiencia; esta labor no genera resultados satisfactorios, afectando la calidad y la economía del mercado.

Un sistema automatizado basado en algoritmos de procesamiento de imágenes, visión artificial, permitirá clasificar el cacao garantizando su calidad, permitiendo: objetividad, poco contacto físico y una clasificación más específica.

Esta tesis es financiada por el Fondo para la Innovación, Ciencia y Tecnología (FINCyT) en el proyecto: “Sistema de procesamiento digital de imágenes para clasificación de productos agroindustriales” Contrato: N° 214-FINCyT-IA-2013.

Resultados de la investigación en el laboratorio de sistemas automáticos de control, permitió la publicación del artículo:

Cocoa bean quality assessment by using hyperspectral images and fuzzy logic techniques QCAV-2015.

Publicado en 12th International Conference on Quality Control by Artificial Vision – 2015. University of Burgundy, Le Creusot – Francia. DOI: 10.1117/12.2182598. SPIE Proceeding Vol. 9534. Autores: Juan Carlos Soto Bohórquez, Flavio Prieto, Guillermo Granda, William Ipanaqué Alama.
(<http://spie.org/Publications/Proceedings/Paper/10.1117/12.2182598>).

Así también debido a la actividad de investigación y desarrollo (I+D) en el laboratorio de Sistemas automáticos de control, y gracias a las competencias adquiridas, permitieron participar en estos otros artículos:

Hyperspectral analysis based anthocyanin index (ARI2) during cocoa bean fermentation process. Jessica Reyes, Juan Soto, William Ipanaqué. IEEE *Xplorer* Digital Library – Pag. 169 – 172. DOI: [10.1109/APCASE.2015.37](https://doi.org/10.1109/APCASE.2015.37). Asia-Pacific Conference on Computer Aided System Engineering (APCASE), 2015. Universidad San Francisco de Quito, Quito – Ecuador.

Modelación, Simulación y Control de un Módulo Experimental Para Control De Ángulos de Orientación de un Quadcopter. Julio Salinas, Marlon Villegas, Renzo Bocanegra, Juan Soto, Ernesto Paiva, William Ipanaqué. XXII Congreso internacional de ingeniería eléctrica, electrónica, computación y ramas afines, (IEEE INTERCON), 2015, Universidad Continental, Huancayo – Perú.

Modelado y Control PID en cascada de un Cuadricóptero para seguimiento de trayectorias. Ernesto Paiva Peredo, Juan Carlos Soto Bohorquez, Julio Salinas Castro, William Ipanaqué Alama. Congreso Chileno de Ingeniería Eléctrica, Electrónica, Tecnologías de la Información y Comunicaciones, (IEEE CHILECON), 2015, Universidad Central de Chile, Santiago de Chile – Chile.

Control con ganancia programable en proceso no lineal con sistema embebido FPGA. Juan Carlos Soto Bohórquez, William Ipanaqué. XV CLCA Congreso Latinoamericano de Control Automático, 2012, Lima – Perú, Pontificia Universidad Católica del Perú (PERU) – Red Peruana de Control Automático (REPCA).

Así también ha permitido la presentación de solicitud de patente y derecho de autor ante INDECOPI:

Derecho de autor: Certificado de Registro de Programas de ordenador (Software) ante INDECOPI. Nro. Partida Registral 0017-2013. Asiento 01. "Software para control de humedad en harina de pescado. SCADA. LADDER PLC".

Patente de Invención: INDECOPI N° Expediente 000826 -2013/DIN. “Sistema para medir la humedad de una masa de una línea de proceso”.

Agradezco a todas las aquellas personas que de una u otra forma me ayudaron en el desarrollo de esta tesis. De manera especial al Dr. Ing. William Ipanaqué Alama, mi asesor, por su constante dedicación y a mis compañeros del Departamento de Electrónica y Automática de la Universidad de Piura.

Gracias a mi madre por ser ejemplo, modelo, sacrificio y enseñarme a soñar en grande, para mí la mejor herencia. Así también a mi esposa Carina por darme los días más felices de mi vida y apoyo constante.

Resumen

Esta tesis es presentada en modalidad de artículo científico y propone el uso de imágenes hiperespectrales con el objetivo de tener un análisis rápido y preciso en la evaluación de la calidad de los granos de cacao, se toma en cuenta las grietas y los cambios del color interno.

Se muestra una aplicación del uso de técnicas de visión artificial para clasificar el cacao, garantizando su calidad y permita: objetividad, poco contacto físico y una clasificación más específica.

El desarrollo de la investigación está financiado por el FINCyT, con el proyecto “Sistema de procesamiento digital de imágenes para clasificación de productos agroindustriales”, contrato: N° 214-FINCyT-IA-2013.

Resultados fueron presentados en 12th International Conference on Quality Control by Artificial Vision – 2015. Llevado a cabo en University of Burgundy, Le Creusot – Francia.

La dimensión del cubo hiperespectral fue reducida por medio del Análisis de Componentes Principales (PCA). El primer componente principal del PCA brinda suficiente información para diferenciar claramente las grietas del grano, las cuales tienen una correlación con el grado de fermentación. Las características extraídas fueron procesadas a través de un bloque difuso capaz de describir la calidad. Se estimaron 3 funciones de membresía para la salida: malo, parcial y bueno, con funciones de tipo triangular y trapezoidal. Se planteó un total de 12 reglas; además el método de la bisectriz fue considerado para la defusificación.

Índice

Introducción	1
Capítulo 1 Introducción	5
1.1. Procesamiento de imágenes, una visión al estado del arte	5
1.1.1. Revisión de la literatura	5
1.2. Terminología de procesamiento de imágenes	13
1.2.1. Procesamiento	13
Capítulo 2 Características morfológicas existentes	19
2.1. Análisis de características externas	19
2.1.1. Definición de características externas	19
2.2. Análisis de características internas	22
2.2.1. Definición de imagen hiperespectral	22
2.2.2. Definición de espectro	24
2.3. Calibración del sistema RESONON PIKA IIG	25
2.4. Procedimiento de Toma de muestra	27
2.5. Clasificador basado en Fuzzy Logic	31
2.6. Resultados	36
Capítulo 3 Artículo Publicado	53
Conclusiones	61
Referencias	63

Introducción

Esta tesis propone el uso de imágenes hiperespectrales con el objetivo de tener un análisis rápido y preciso en la evaluación de la calidad de los granos de cacao tomando en cuenta las grietas y los cambios del color interno. [1], [2], [3], [4]

Un dispositivo electrónico permite a productores de cacao determinar la calidad de los granos. Estudio realizado en [5] muestra el estudio en detección de defectos internos en los pepinos. Investigaciones en el campo de procesamiento de visión artificial ha llegado a determinar características internas de este producto. El estudio tuvo como objetivo determinar un conjunto de longitudes de onda óptimas que podrían ser útiles para la detección de defectos internos en línea.

Se desarrolló un método no destructivo [6] para determinar la calidad de manzanas, discriminando el nivel de madurez de las manzanas en diferentes condiciones de almacenamiento a lo largo de la vida útil. Se utilizaron métodos de segmentación, pre procesamiento y el método de componentes principales.

Estudios relacionados con sistemas de visión artificial relacionados con técnicas de aprendizaje [7] se han aplicado cada vez más para la evaluación de calidad de los alimentos mediante visión por computador en los últimos años. Las investigaciones mostraron resultados de características, tales como tamaño, forma, color y textura.

Se desarrolló un método no destructivo para evaluar y mapear los índices de calidad [8]. Investigaciones que abarcan el desarrollo de dispositivos empleados en colorimetría se pueden encontrar en [9]. Estudio de análisis no destructivos para la calidad y determinación directa de enterobacterias en filetes de pollo, lo podemos encontrar en [10], se establecieron modelos de regresión de mínimos cuadrados parciales (PLSR).

Investigaciones en [11] determinó la calidad de los granos de cacao por medio de imágenes digitales. Las principales características que se obtuvieron fueron forma y color de los granos. Se utilizaron estructuras de Redes Neuronales Artificiales (RNA), para clasificar el grano en cuatro tamaños: granos enteros, granos quebrados, fracciones de grano, y granos con piel dañada.

Se desarrolló un sistema de visión artificial [12] para la clasificación de frutos de café en once categorías dependiendo de su estado de madurez. Para la descripción de la forma, color y textura de cada fruto de café se extrajeron 208 características,

reduciéndolas a 9 con base en los resultados de métodos de selección de características, univariado y multivariado.

Esta tesis, basada en la modalidad de artículo científico, muestra una aplicación del uso de técnicas de visión artificial, el desarrollo permitirá clasificar el cacao garantizando su calidad, permitiendo: objetividad, poco contacto físico y una clasificación más específica.

La dimensión del cubo hiperespectral fue reducida por medio del Análisis de las Componentes Principales (PCA), así mismo se extrajo el índice Atmospherically Resistant Vegetation Index (ARVI). La imagen formada por la componente principal PC1 del PCA brinda suficiente información para diferenciar claramente las grietas del grano, las cuales tienen una correlación con el grado de fermentación. Las características extraídas fueron procesadas a través de un bloque difuso capaz de describir la calidad. Se estimaron 3 funciones de membresía para la salida: malo, parcial y bueno usando funciones de tipo triangular y trapezoidal. Se planteó un total de 12 reglas; además el método de la bisectriz fue considerado para la defusificación.

Resultado de la investigación, desarrollado en el laboratorio de Sistemas Automáticos de Control, permitió la publicación del artículo:

Cocoa bean quality assessment by using hyperspectral images and fuzzy logic techniques.

Publicado en 12th International Conference on Quality Control by Artificial Vision – 2015. University of Burgundy, Le Creusot – Francia. DOI: 10.1117/12.2182598. SPIE Proceeding Vol. 9534. Autores: Juan Soto, Flavio Prieto, Guillermo Granda, William Ipanaqué. (<http://spie.org/Publications/Proceedings/Paper/10.1117/12.2182598>).

Se obtuvieron una muestra de granos de cacao los cuales fueron clasificados con personal especialista de la asociación. Con ellos se analizó cualitativamente cada grano de cacao y se etiquetaron para su reconocimiento en laboratorio. Se muestra los resultados de clasificación, el algoritmo le asocia un valor cuantitativo al grano de cacao, es decir, asocia un grado de bueno, parcial o malo. El clasificador basado en Lógica Difusa ha dado buenos resultados. La salida del bloque difuso tiene un rango entre 0 y 1 que va desde mal fermentado hasta bien fermentado respectivamente, esta muestra una marcada diferencia entre los granos mal fermentados con el resto, asignándoles un valor de 0.12. En cuanto a los granos parcialmente fermentados oscilan entre los valores 0.4 a 0.57. Por otro lado, el mínimo valor que presentan los granos bien fermentados es de 0.73, y el máximo valor de estos bordea el valor de 0.9. De esta manera se puede considerar a la salida del bloque difuso como un índice capaz de describir la calidad del cacao en tres niveles de fermentación.

Las características externas que nos indican subjetivamente la buena calidad de los granos de cacao son: el tamaño, color, forma. Internamente se presenta ausencia del cotiledón. Esta última característica se observa mediante la presencia de una coloración marrón o marrón rojiza en el interior, y estrías profundas.

Debido a la actividad de investigación, desarrollo e innovación (I+D+i) en el laboratorio de Sistemas automáticos de control, y gracias a las competencias adquiridas, permitieron participar en otros artículos:

Hyperspectral analysis based anthocyanin index (ARI2) during cocoa bean fermentation process. Jessica Reyes, Juan Soto, William Ipanaqué. IEEE *Xplorer* Digital Library – Pag. 169 – 172. DOI: [10.1109/APCASE.2015.37](https://doi.org/10.1109/APCASE.2015.37). Asia-Pacific Conference on Computer Aided System Engineering (APCASE), 2015. Universidad San Francisco de Quito, Quito – Ecuador.

Modelación, Simulación y Control de un Módulo Experimental Para Control De Ángulos de Orientación de un Quadcopter. Julio Salinas, Marlon Villegas, Renzo Bocanegra, Juan Soto, Ernesto Paiva, William Ipanaqué. XXII Congreso internacional de ingeniería eléctrica, electrónica, computación y ramas afines, (IEEE INTERCON), 2015, Universidad Continental, Huancayo – Perú.

Modelado y Control PID en cascada de un Cuadricóptero para seguimiento de trayectorias. Ernesto Paiva Peredo, Juan Carlos Soto Bohorquez, Julio Salinas Castro, William Ipanaqué Alama. Congreso Chileno de Ingeniería Eléctrica, Electrónica, Tecnologías de la Información y Comunicaciones, (IEEE CHILECON), 2015, Universidad Central de Chile, Santiago de Chile – Chile.

Control con ganancia programable en proceso no lineal con sistema embebido FPGA. Juan Carlos Soto Bohórquez, William Ipanaqué. XV CLCA Congreso Latinoamericano de Control Automático, 2012, Lima – Perú, Pontificia Universidad Católica del Perú (PERU) – Red Peruana de Control Automático (REPCA).

Así también ha permitido la presentación de solicitud de patente y derecho de autor ante INDECOPI:

Derecho de autor: Certificado de Registro de Programas de ordenador (Software) ante INDECOPI. Nro. Partida Registral 0017-2013. Asiento 01. "Software para control de humedad en harina de pescado. SCADA. LADDER PLC".

Patente de Invención: INDECOPI No. Expediente 000826 -2013/DIN. "Sistema para medir la humedad de una masa de una línea de proceso".

Capítulo 1

Introducción

1.1. Procesamiento de imágenes, una visión al estado del arte

1.1.1. Revisión de la literatura

En [13] se desarrolló un método no destructivo para evaluar y mapear los índices de calidad en el pimentón. Se estudiaron tres cultivos durante la etapa de maduración con imágenes dentro del rango visible y cercano al infrarrojo.

Los parámetros de sólidos solubles, cantidad de clorofila, carotenoide y contenido de ácido ascórbico fueron determinados y correlacionados con datos espectrales. La información espectral de los hipercubos fue analizada por procedimientos quimiométricos. Para el desarrollo del modelo se usó regresión de los mínimos cuadrados parciales, así mismo se realizaron comparaciones entre el análisis de regresión de mínimos cuadrados parciales del espectro de reflectancia, y los espectros de preprocesado tales como la primera derivada de la reflectancia.

La calidad de los pimentones es un concepto complejo basado en múltiples características: color (relacionado con clorofila y contenido de caroteno), sólidos solubles, materia seca y contenido de vitamina C.

Muchos métodos mecánicos y líneas de clasificación manuales están basados en la apariencia externa y falta la habilidad para analizar la calidad interna del producto.

El sistema de visión comercial basado en la medición del color es la línea de clasificación más usada porque es rápida y económica, pero el parámetro de color por sí solo no proporciona una segura determinación de la maduración.

Las imágenes en el rango espectral fueron usadas para la medición de la maduración del fruto, contenido de sólidos solubles y para detectar defectos internos de los pepinillos.

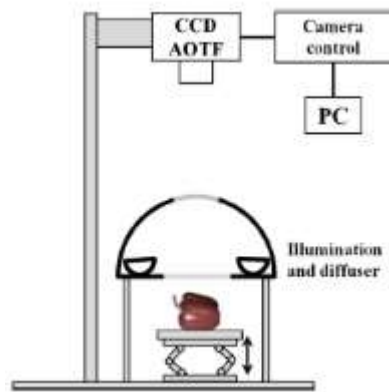


Figura 1. Esquema de sistema implementado
Fuente: J. Gómez, et al., (2007).

Otro estudio de análisis no destructivos para la determinación de la calidad en los alimentos, es el trabajo basado en la determinación directa de enterobacterias en filetes de pollo, lo podemos encontrar en [14].

Se establecieron modelos de regresión de mínimos cuadrados parciales (PLSR), usando longitudes de onda en el rango espectral de 930-1450 nm con un coeficiente de determinación mayor a 0.82.

Los filetes de pollo fueron cortados en piezas para formar muestras con espesor de aproximadamente 5mm. Dichas muestras fueron posteriormente escaneadas por un sistema de imágenes.

A continuación, se muestra el sistema de adquisición de imágenes.

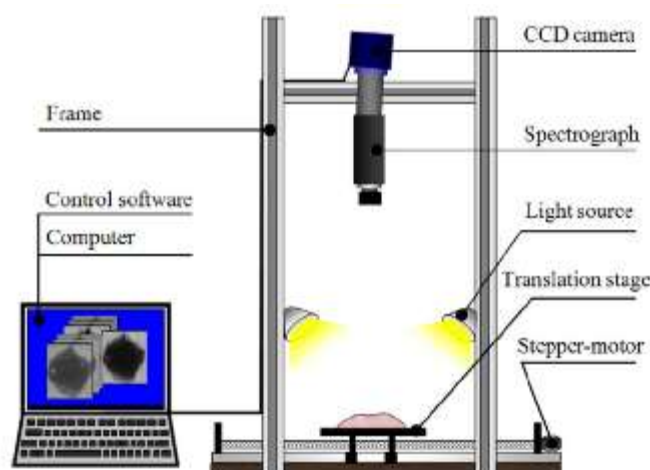


Figura 2. Esquema del sistema de visión artificial.
Fuente: Yao-Ze Feng, et al., (2013).

El sistema está formado por 6 componentes:

Un sistema de traslación accionado por un motor paso a paso.

Fuente de luz, conformado por 2 Lámparas halógenas de 500 W cada una.

Espectrógrafo (Inspector N17E).

Cámara CCD (Xeva 992, Xenics Infrared Solutions, Belgium).

Computadora con software de control para la adquisición de las imágenes.

Estructura de soporte.

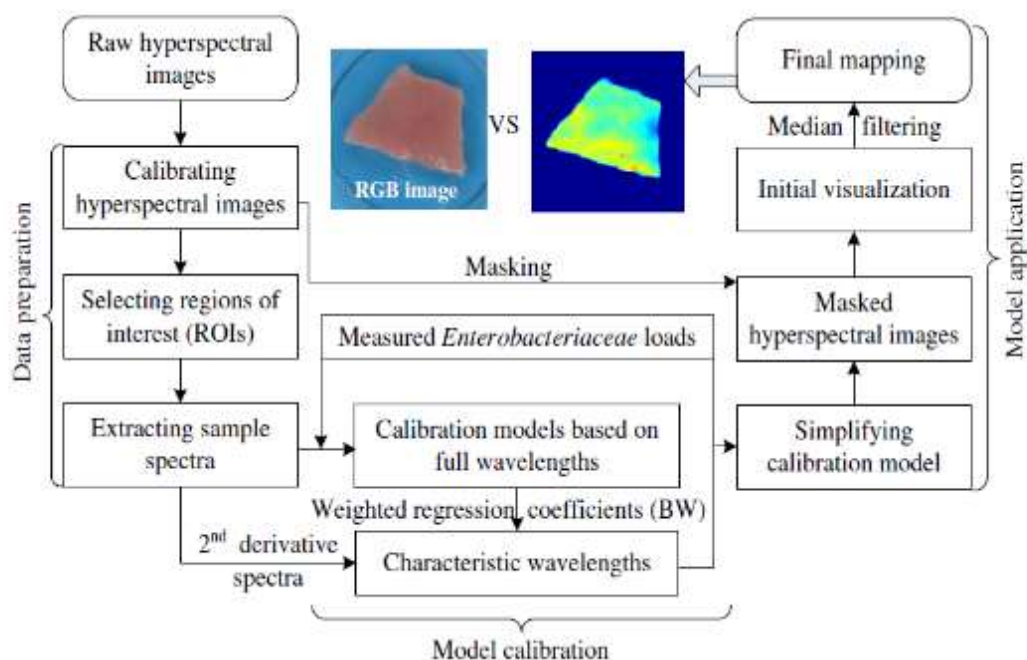


Figura 3. Modelo del sistema de procesamiento de imágenes.

Fuente: Yao-Ze Feng, et al., (2013).

Investigaciones desarrolladas para la evaluación y control estricto de la calidad en carne de porcino se tiene en [15], un método de detección rápida basadas en técnicas de imágenes hiperespectrales.

Investigaciones que abarcan el desarrollo de dispositivos empleados en colorimetría se pueden encontrar en [16]. Un dispositivo electrónico permite a productores de cacao determinar la calidad de los granos. Un sensor permite detectar la luz roja reflejada. Las diferencias en reflexión de luz roja (660 nm) podría ser cerca del 20% y muy fácil de medir. El procedimiento fue comprobado por un espectrómetro en 660nm de longitud de onda, y reveló un alto coeficiente de correlación permitiendo encontrar diferencias en la reflexión entre los granos de cacao secos y mal fermentados.

Dependiendo del escenario y procesos después de la cosecha, los granos de cacao adquieren colores de rojo y púrpura (cacao fresco) y marrón oscuro (fin del proceso de fermentación). Luego de esto se le puede aplicar una técnica de colorimetría para medir el color. Las mediciones de color son apropiadas para la estimación de la calidad del cacao. Se usó un colorímetro llamado triestímulo o cromametro los cuales determinan el

color en términos de xyz. Los cromametros son excelentes instrumentos para evaluar el color de objetos.

Granos de cacao secos fermentado y mal fermentado fueron tomados aleatoriamente.

Las mediciones de reflectancia fueron realizadas con un sensor que contiene un diodo de luz (LED) y un fotodiodo en una configuración óptica. En una caja de plástico negro se taladraron 2 agujeros de 5mm de diámetro. Un agujero es perpendicular a la superficie del objetivo (granos de cacao) en el extremo opuesto contiene un fotodiodo de silicio. El segundo agujero fue perforado con un ángulo de 30 grados con respecto al primer agujero que contiene el LED. Ambos agujeros cilíndricos se unen en la superficie donde está localizado el objetivo (granos de cacao). Ver figura 4.

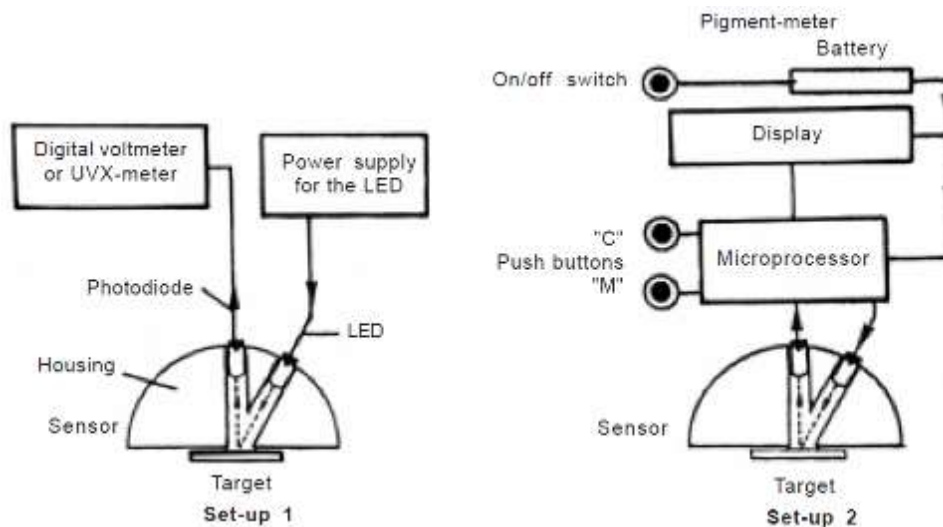


Figura 4. Esquema del sistema de adquisición de imágenes.

Fuente: C. Meulemas, et al., (2002).

Un estudio realizado [5] en la detección de defectos internos en los pepinos. Investigaciones en el campo de procesamiento de visión artificial ha llegado a determinar características internas de este producto.

El estudio tuvo como objetivo determinar un conjunto de longitudes de onda óptimas que podrían ser útiles para la detección de defectos internos en línea.

Se adquirieron imágenes de 300 pepinillos frescos y 280 pepinillos enteros de clase defectuoso y normal. Las bandas de frecuencia seleccionadas serán útiles para el desarrollo de un sistema de inspección clasificación en línea de pepinillos y encurtidos. Ver figura 5.

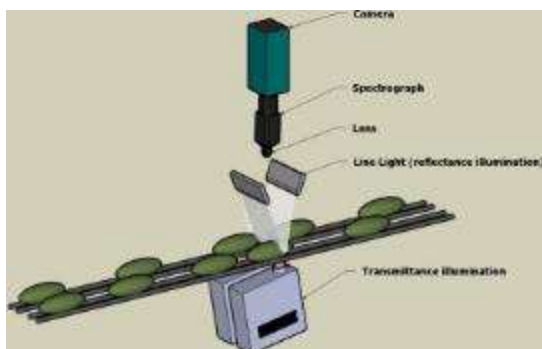


Figura 5. Esquema del sistema de procesamiento de imágenes.

Fuente: A. Diwan, et al., (2006).

Autores en [17] centran su investigación de visión artificial para evaluar la maduración del melocotón. Esta permite determinar el grado de maduración de la fruta. Los índices considerados ópticos (IND1 y IND2, IND3 y IAD, se basan en la combinación de longitudes de onda cercanas al pico de absorción de clorofila en 680nm. Figura 6.

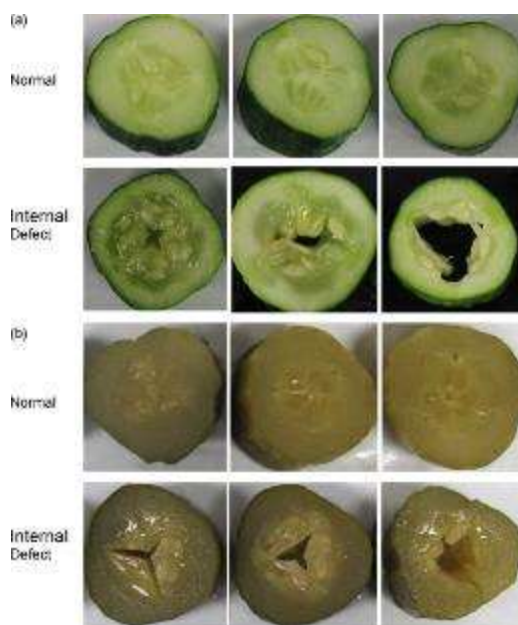


Figura 6. Defectos internos del pepino.

Fuente: L. R. Lleo, et al., (2011).

Se determinó la región eficaz para la detección de defectos internos en los pepinos y pepinillos, esta fue en el infrarrojo cercano entre 740-1000 nm. Bandas de frecuencia de 745, 805, 965 y 985 nm y una resolución de 20 nm para los pepinos logró una precisión del 94,7%. Para pepinillos enteros, las bandas de onda entre 745, 765, 885 y 965nm y una resolución espectral de 40 nm produjeron una precisión de 82,9%. Se demostró que una alta resolución espectral no es necesaria cuando se utiliza las bandas de frecuencia seleccionadas para la detección de defectos internos de los pepinos y pepinillos.

Estudios relacionados con sistemas de visión artificial relacionados con técnicas de aprendizaje [18] se han aplicado cada vez más para la evaluación de calidad de los alimentos mediante visión por computador en los últimos años.

Las investigaciones mostraron resultados de características, tales como tamaño, forma, color y textura. Estas características son datos objetivos utilizados para representar los productos alimenticios, que pueden ser utilizados para formar el conjunto de entrenamiento. Una vez que se ha obtenido el conjunto de entrenamiento, el algoritmo de aprendizaje extrae los datos necesarios para tomar una decisión, ver Figura 7.

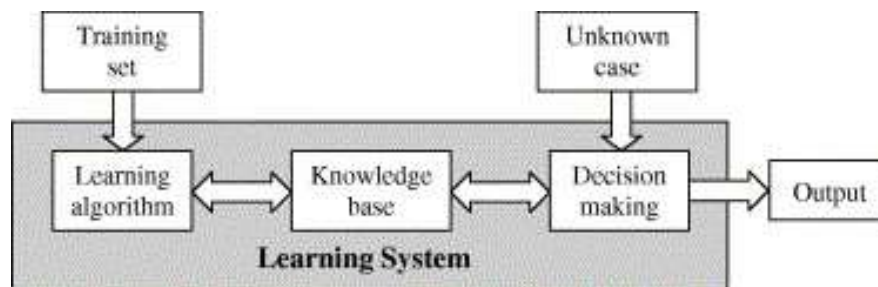


Figura 7. Muestra la configuración del sistema.

Fuente: Yao-Ze Feng, et al., (2013).

Investigaciones en Indonesia, donde se determinó la calidad de los granos de cacao [10], por medio de imágenes digitales. Las principales características que se obtuvieron fueron forma y color de los granos. Además se utilizaron dos estructuras de RNA (redes neuronales artificiales), para clasificar el grano en cuatro tamaños: granos enteros, granos quebrados, fracciones de grano, y granos con piel dañada.

El sistema que se utilizó para la adquisición y procesamiento de las imágenes se muestra en la Figura 8:



Figura 8. Sistema de procesamiento de imágenes para los granos de cacao.

Fuente: Kıvanç Kılıç, et al., (2007).

Las imágenes obtenidas de los granos de cacao fueron las siguientes:

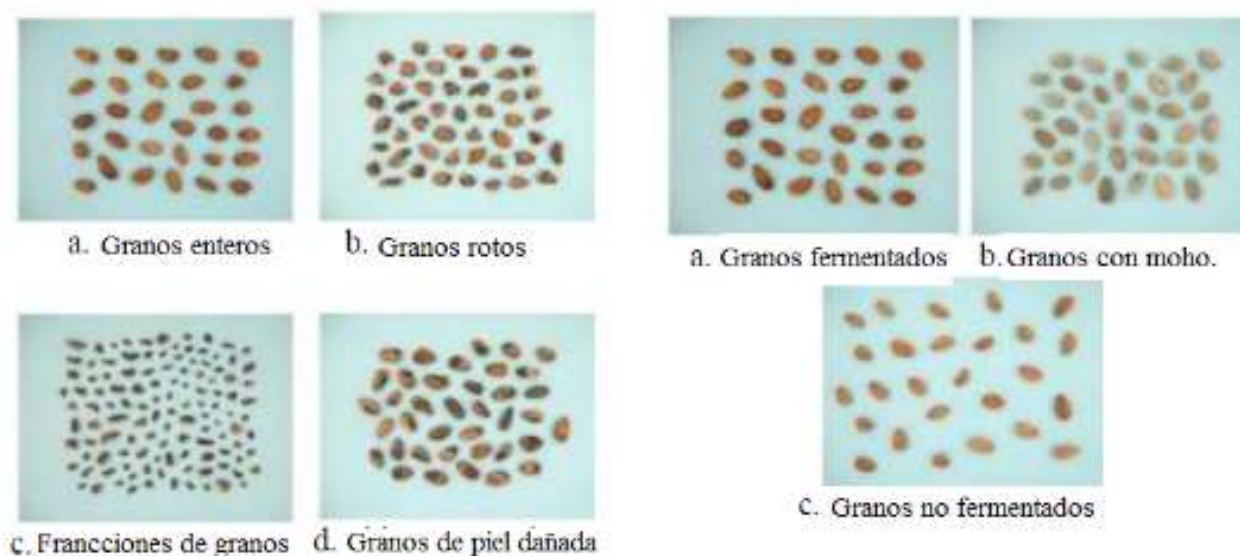


Figura 9. Imágenes de selección de los granos.

Fuente: Kivanç Kılıç, et al., (2007).

Se desarrolló un método no destructivo [19] para determinar la calidad de manzanas. Las imágenes fueron evaluadas discriminando el nivel de madurez de las manzanas en diferentes condiciones de almacenamiento a lo largo de la vida útil. Para el análisis de las imágenes se utilizaron métodos de segmentación, pre procesamiento y el método de componentes principales.

El color constituye la primera relación que se establece entre el producto y el consumidor, siendo esencial para la toma de decisiones de compra.

En el proceso de maduración, la degradación de la superficie de clorofila de la fruta se lleva a cabo mientras que los pigmentos de color, ambos carotenoides (amarillo) y antocianinas (rojo), están siendo sintetizados. Es decir, dos colores pueden aparecer: la primera, de color amarillo, es el fondo el color, y la segunda, de color rojo, es el color de la superficie. El color de manzanas de buena calidad varía de verde a tonos de color verde amarillento que muestran un rubor rosáceo ligero.

El proceso de segmentación de la imagen se muestra en la siguiente figura 10:

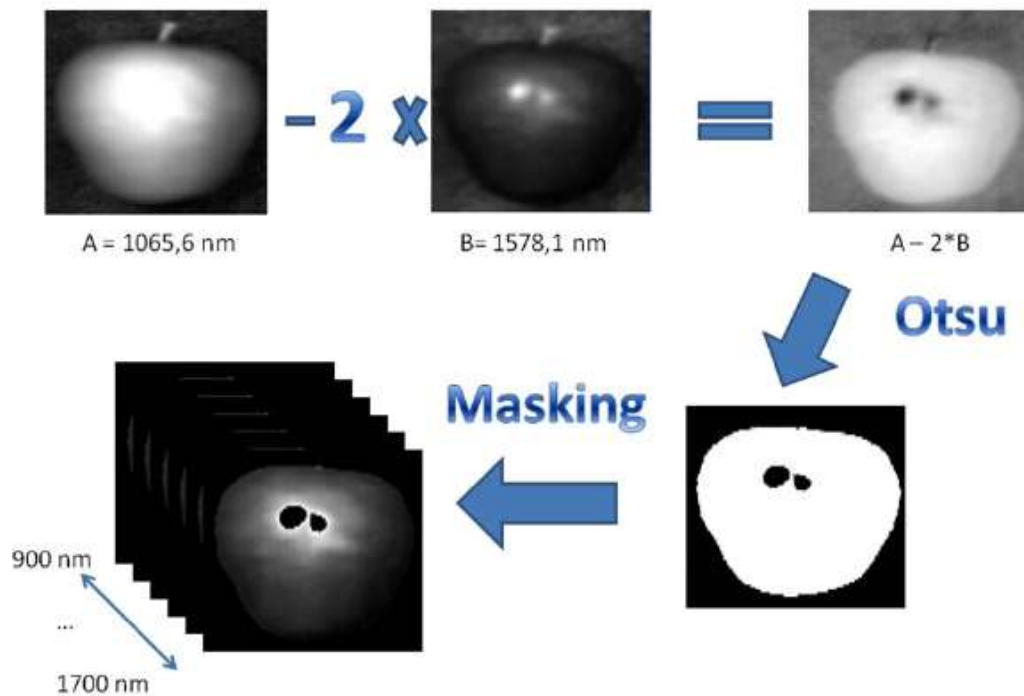


Figura 10. Análisis de segmentación de la imagen.

Fuente: C. Garrido-Novell, et al., (2012).

Se desarrolló un sistema de visión artificial [12] para la clasificación de frutos de café en once categorías dependiendo de su estado de madurez.

Para la descripción de la forma, color y textura de cada fruto de café se extrajeron 208 características. La reducción del conjunto de características de 208 a 9 se hizo con base en los resultados de dos métodos de selección de características, univariado y multivariado. Las características seleccionadas corresponden a 4 características de textura, 3 de color y 2 de forma. Este conjunto final de características se evaluó en dos técnicas de clasificación: Bayesiano y redes neuronales. Con el clasificador Bayesiano se obtuvo un error de clasificación del 5,43% y requirió un tiempo de clasificación de 5,5 ms, mientras que usando redes neuronales el error de clasificación fue de 7,46 %, pero disminuyó el tiempo de clasificación a 0,8 ms.

Las características de textura aportan mayor información discriminante que las de color, demostrando que para la separación lo mejor no es sólo el valor del color sino también su distribución en la superficie del fruto.

Las características de forma seleccionadas contienen la información de baja frecuencia con la cual se reconoce el cambio en la redondez. La mejor característica de color corresponde a una componente del modelo de colores oponentes mostrando una mejor representación del color en otros modelos diferentes a los comúnmente utilizados RGB y HSI. [20], [21], [22].

Una investigación realizada en la detección de defectos en granos de avena se aprecia en [23]. Se realizaron investigaciones por medio de cámaras, para la adquisición de la imagen a analizar. Mediante el método de componentes principales se obtuvo la detección de sus características.

El objetivo principal de esta investigación fue determinar las características mediante el modelo estadístico de análisis de componentes principales. Los resultados mostraron que es posible reconocer con precisión granos de avena individuales por imágenes. El estudio también demostró que da buenos resultados de clasificación para ciertas longitudes de onda. En la siguiente figura se muestra la imagen a analizar segmentada.

Un estudio desarrollado en la evaluación de carne de ganado vacuno [11]. El sistema de visión artificial fue empleado para la determinación de medidas colorimétricas de la carne. Para la determinación de estas medidas se realizó un análisis de similitud. Las características colorimétricas en la carne son muy importantes porque es una medida de su calidad. El color de la carne es una de sus características organolépticas más importantes.

Las características organolépticas influyen en la aceptabilidad del producto y juega un papel muy importante en la decisión de la compra de la carne [24], [25], [26], [27], [28]. Las medidas colorimétricas también permiten la detección de ciertas anomalías o defectos que los alimentos pueden presentar.

Investigación realizada para clasificar granos, con respecto a su tamaño y la cuantificación del color de las muestras [29]. El sistema consiste en un hardware y un software. El hardware se desarrolló para capturar una imagen estándar de las muestras. Con respecto al software se realizó algoritmos en Matlab para la segmentación, la operación morfológica y la cuantificación del color de las muestras.

Un estudio se llevó a cabo para el desarrollo de un sistema de visión artificial por computadora para la detección de patatas irregulares en tiempo real [30]. Los algoritmos se programaron y desarrollaron específicamente para la adquisición y procesamiento de imágenes, control de todo el proceso, el ahorro de la clasificación de resultados y monitorear el progreso de todas las operaciones. Se formuló una base de datos de imágenes con patatas de diferentes tamaños, y además se extrajeron características como perímetro, centroide, área, momento de inercia, longitud y el ancho. Dichas características son muy importantes para determinar la calidad del producto.

1.2. Terminología de procesamiento de imágenes

A continuación se presenta una relación de terminas usados en el campo de procesamiento de imágenes.

1.2.1. Procesamiento

El procesamiento digital de imágenes se define como el conjunto de técnicas que son aplicadas a las imágenes digitales con el fin de facilitar la búsqueda de información o mejorar su calidad. Se puede dividir en dos puntos principales:

El procesamiento propiamente dicho, el cual está referido a la realización de transformaciones y a la restauración y mejoramiento de las imágenes.

El análisis, que consiste en la extracción de propiedades y características de las imágenes, así como la clasificación e identificación y el reconocimiento de patrones.

1. Pixel

Es la unidad más sencilla que compone a una imagen digital. Si se amplía una imagen digital, se pueden ver los pixeles que la conforman. Estos son los cuadritos de color que en conjunto conforman la imagen, son representados como los elementos de una matriz.

2. Vecindad

Es la relación que existe entre un pixel con los pixeles más cercanos a él. Existen dos tipos de vecindad:

La vecindad 4-vecinos, que se constituye de los pixeles que se encuentran arriba, abajo, a la derecha y a la izquierda del pixel seleccionado.

La vecindad 8-vecinos, la cual está constituida por los pixeles que conforman la vecindad 4-vecinos más los cuatro pixeles que se encuentran en las diagonales del pixel en cuestión.

3. Conectividad

Es un concepto muy común en la detección de regiones u objetos contenidos en una imagen determinada. Se define como una situación de adyacencia y vecindad. Basados en estos, se puede dividir en:

Conectividad 4, sucede cuando el punto más cercanos entre dos o más objetos se encuentran en relación de 4-vecinos, de manera que conforman un solo objeto.

Conectividad 8, sucede cuando el punto más cercanos entre dos o más objetos se encuentran en relación de 8-vecinos, de manera que conforman un solo objeto.

4. Distancia entre pixeles

Es una medida muy usada en el procesamiento de imágenes. Sus aplicaciones van desde la similitud hasta la medición de objetos en una imagen. Existen muchos tipos de medidas para encontrar la distancia entre pixeles, no obstante las más usadas son:

Distancia euclidiana: se define como la longitud de la hipotenusa del triángulo rectángulo con catetos distancia horizontal y distancia vertical entre los pixeles.

$$D_e(p_1, p_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

Distancia city-block: está conformada por la suma de la distancia horizontal y vertical entre ambos pixeles, la cual está definida como:

$$D_{C-B}(p_1, p_2) = |x_1 - x_2| + |y_1 - y_2|$$

Distancia chessboard: es la máxima distancia entre los recorridos vertical y horizontal que hay entre dos pixeles. Se define como:

$$D_{CH}(p_1, p_2) = \max(|x_1 - x_2|, |y_1 - y_2|)$$

De estas tres distancias, la distancia city-block es la que tiene mayor valor mientras que la distancia chessboard tiene el menor valor.

5. Imágenes y matrices

En Matlab una imagen es considerada una función bidimensional descrita en forma matricial:

$$I = \begin{pmatrix} I(1,1) & I(2,1) \cdots & I(M,1) \\ I(1,2) & I(2,2) \cdots & I(M,2) \\ \vdots & \vdots & \vdots \\ I(1,N) & I(2,N) \cdots & I(M,N) \end{pmatrix}$$

Donde M es el número de filas y N el número de columnas, y MxN es la dimensión de la imagen o número de pixeles. A partir de esto Matlab puede indexar datos matricialmente cambiando las posiciones de los elementos de la matriz.

6. Tipos de datos

Generalmente se utilizan números enteros para indexar las coordenadas de una imagen. Sin embargo, los valores de los pixeles no están restringidos a valores enteros del intervalo [0,255], sino que pueden tomar distintos valores positivos, negativos o decimales, incluso números complejos.

Tipo double (o números de coma flotante), por lo que muchas operaciones realizadas sobre la imagen tendrán este formato. Este tipo de datos contiene pixeles cuyos valores son decimales, ya se negativos o positivos.

El tipo de dato uint8 es el que normalmente se presenta cuando cargamos una imagen a Matlab. Presenta pixeles con números enteros que van de 0 a 255.

Otro tipo de dato popular es el logical, el cual permite representar de manera compacta imágenes binarias, por lo que los valores que pueden tomar los pixeles son 0 ó 1.

Un tipo de datos muy utilizado en Simulink es el single, el cual permite representar un número real con una precisión simple, con un intervalo de representación [10-38,1038].

7. Tipos de imágenes

El toolbox de procesamiento de imágenes de MatLab permite la manipulación de cuatro tipos de imágenes:

Imágenes de intensidad o escala de grises: es una matriz cuyos valores han sido escalados para representar un intervalo determinado. Si se trata de una imagen uint8 entonces los datos que la conforman se encuentran en el intervalo [0,255], si la imagen es del tipo double entonces los datos que la conforman se encuentran en el intervalo [0,1].

Imágenes binarias: una imagen binaria del tipo logical se representa en MatLab como un arreglo que sólo contiene unos y ceros. Estos ceros y unos son especiales, ya que no representan una cantidad, sino que se pueden considerar como banderas que indican un estado de verdadero (1) o falso (0).

Imágenes indexadas

Imágenes de color RGB: se basan en la combinación de los colores rojo (R), verde (G) y azul (B). Este modelo es un formato de color aditivo, esto significa que la combinación de colores se basa en la adición de los componentes individuales considerando como base el negro.

8. Indexado de vectores y matrices

El indexado se puede definir como la discriminación y búsqueda de elementos o píxeles contenidos en una imagen. Muchas operaciones de procesamiento de imágenes se basan en esto por lo que las operaciones de indexado deben ser consideradas importantes.

Indexado de vectores: Si se tiene un vector 1xN, puede ser indexado utilizando sólo un índice numérico, de esta forma si se considera un vector llamado “v”, entonces al introducir v(n) nos dará el elemento con la posición n. Si se desea indexar un intervalo utilizamos v(n1:k:n2), que nos da un vector que va desde el elemento n1 al elemento n2, con un espaciamiento k entre elementos; si el espaciamiento es uno, k se puede omitir.

Indexado de matrices: Para el indexado de matrices son necesarios dos números, uno para elegir los renglones y otro para las columnas.

9. Histogramas

Son distribuciones que describen la frecuencia con que se presentan los valores de intensidad o píxeles de la imagen. El histograma no proporciona información acerca del origen de los píxeles que lo conforman, es decir se pierde la información espacial de los píxeles en la imagen, por lo que resulta imposible reconstruir una imagen a partir de su

histograma. Sin embargo, el histograma muestra características muy importantes de una imagen, como por ejemplo el contraste y el rango dinámico, los cuales son problemas debidos a la captación de la imagen que es necesario solucionar para que se puedan analizar con mejor claridad propiedades de la imagen en bloques posteriores de procesamiento.

10. Filtro

Los filtros pueden ser considerados como un conjunto de operaciones en donde el nuevo pixel no depende únicamente del pixel original, sino que también depende de otros pixeles que están en una determinada vecindad en relación con él. Este conjunto de operaciones no afecta la geometría de la imagen resultante, de tal forma que la relación entre la imagen resultante y la imagen original sigue siendo igual que en las operaciones de pixel 1 a 1.

Filtros de suavizado: cuando en una imagen hay un cambio abrupto de intensidad o si la intensidad se mantiene constante, se realiza un filtro de suavizado de manera que el pixel adopte un promedio de intensidad con sus vecinos.

Filtros de diferencia: Si un filtro contiene como parte de sus coeficientes números negativos, su operación puede interpretarse como la diferencia de dos diferentes sumas. Una suma de todas las combinaciones lineales debidas a los coeficientes positivos del filtro menos la suma de todas las combinaciones lineales debidas a los coeficientes negativos, dentro de la región definida por el filtro $R(x,y)$.

11. Ruido

Se considera como ruido a la variación aleatoria del brillo o color de las imágenes digitales, el cual es producido por el dispositivo de entrada (que puede ser una cámara digital). Es difícil obtener imágenes con estas características, porque por lo general se introducen en una imagen en MatLab y de esta forma poder comprobar la efectividad de los filtros.

12. Bordes y contornos

Los bordes pueden ser considerados como puntos en una imagen en los cuales la intensidad en una determinada dirección cambia drásticamente. El valor del borde en una imagen depende del cambio presentado en la intensidad. Este cambio normalmente es calculado a partir de la derivada.

13. Esquinas de una imagen

Son puntos prominentes contenidos en una imagen que se caracterizan por presentar también un alto valor del gradiente pero a diferencia de los bordes este alto valor del gradiente no solo se manifiesta en una dirección sino en varias. Se puede imaginar a las esquinas como puntos en la imagen que colindan con diferentes bordes a la vez.

Capítulo 2

Descripción y resultados

2.1. Análisis de características externas

En el Perú, la clasificación y control de calidad de granos de cacao se realiza mayormente de forma manual, sólo algunos productores tienen un sistema mecánico de clasificación por tamaño y forma.

2.1.1. Definición de características externas

2.1.1.1. Color

El análisis de la característica del color externo del grano de cacao no ha dado resultados satisfactorios para la determinación de la calidad. Debido que el grano de cacao cuenta con una capa producto de la baba que ha pasado por el proceso de fermentación y secado. Esta baba hace imposible la determinación del color. Por otro lado esta capa o cáscara son en su mayoría del mismo tono. Por tal motivo el parámetro de color se correlacionará con algunas características internas, sin necesidad de destruir el grano.

En la imagen a continuación se presenta un grano partido, en donde se ha adquirido la imagen del interior del grano. Se pueden apreciar las diferentes estrías y un vacío en el centro, ausencia del cotiledón. Se ha podido calcular el porcentaje de área en donde el color rojizo es más predominante, sin embargo la presencia de pocas estrías indica que el grano no necesariamente es de buena calidad.



Figura 11. Análisis de color del grano.
Fuente: Elaboración Propia.

2.1.1.2. Forma

La forma común del grano de cacao es semejante a una elipse, es por ello que aquellos granos de cacao que no se asemejan a una elipse tienden a ser granos descartados. Estos granos han sido afectados en los procesos de fermentación, debido a una inadecuada remoción, y posiblemente en el secado, por un inadecuado volteo. Tiempos de secado al sol inadecuados puede darle al grano un golpe térmico brusco y dañarlo.

Se determinan el valor de los radios y calcula R/r , siendo esta una medida cuantitativa de la forma elíptica del grano.



Figura 12. Análisis de forma del grano.
Fuente: Elaboración Propia.

2.1.1.3. Tamaño

Algunos productores poseen un sistema de clasificación mecánico según el tamaño del grano, por medio de un agujero se seleccionan los que no pueden pasar a través de este agujero y los que si pasan (su diámetro es menor que el agujero) son descartados. El sistema utilizado es por zaranda, haciendo una clasificación solo por tamaños, más no por calidad.

Se ha detectado, en el uso de la zaranda, granos de cacao clasificados de tamaño grande tienen otras características de forma achatada; mirando el interior del grano el cotiledón y/o las estrías no han formado correctamente, dando como resultado un cacao de mala calidad.

Es necesario no solo mirar el tamaño, sino también la forma. En la Figura 12 muestra el cálculo del área A_g , el cual nos indica el tamaño del grano. Los algoritmos han sido desarrollados en el IDE Matlab, se cuenta también con un GUI interfaz con el usuario.

2.1.1.4. Clasificación de grano

El consumo de cacao va en aumento en el mundo y con mayor razón en los países en vías de desarrollo como el nuestro, es por esto que se hace evidente la necesidad de mejorar la selección y clasificación del grano de cacao, para que disminuyan las pérdidas por pudrición y mejorar la presentación del producto final.

Una mala clasificación nos da como resultado:

Bajos precios.

Producto es rechazado.

Dificultad para ingresar en mercados extranjeros.

Mayor almacenamiento con probabilidades de malograrse.

Clasificación según color

Como se mencionó anteriormente la clasificación del color externo no presenta resultado satisfactorio. El color externo no es el mismo del color interno.

Clasificación según forma

La clasificación de la forma del cacao tiene un impacto en el presente proyecto. La forma que tiene el producto de la pérdida del cotiledón y la formación de las estrías haciendo que el grano se hinche.

No se aceptan granos muy deformes, es decir que presenten grandes protuberancias. Su forma debe asemejar a una elipse. Esta clasificación es clave para su correlación con otros parámetros adquiridos y presentados más adelante.

El parámetro que nos va a indicar la forma de la elipse es el producto R/r . El algoritmo calculará los radios mayor y menor de la elipse, R y r respectivamente y posteriormente con el valor de R/r tendremos una medida cuantitativa de la forma elíptica del grano.

El algoritmo deducir si hay granos pegados y debe determinar la cantidad de ellos.

El resultado de esta clasificación es en apto y no apto.

Clasificación según tamaño

Existen diferentes calibres, esto de acuerdo al mercado que se desea abastecer, actualmente no existe una clasificación normada.

En la Asociación de Productores ASPROBO realizan una clasificación por 3 tamaños. Del mismo modo, el algoritmo repetirá la clasificación: Tipo1 es de tamaño grande, tipo2 tamaño intermedio y tipo 3 el más pequeño descarte.

La Tabla1 describe las clases en que han sido agrupados respecto al calibre del grano.

CLASE	CALIBRE
Tipo 1	2 – 3 cm
Tipo 2	1.5 – 2 cm
Tipo 3	- 1.5 cm

2.2. Análisis de características internas

La determinación del parámetro interno usada para determinar la calidad del cacao, es la formación de estrías. Como se mencionó en el proceso de fermentación del cacao sufre la pérdida del cotiledón y la formación de estas estrías, sufriendo un hinchamiento del grano.

La característica interna utilizada en este trabajo es el área de estrías internas formadas. Estas pruebas se realizaron con la captura de imágenes del grano partido, cortado a la mitad por una guillotina. Se adquirieron las imágenes, llamadas cubos hiperespectrales, para luego ser procesadas en el software Matlab. Las imágenes hiperespectrales constituyen una fuente de información mucho más amplia que una imagen convencional.

2.2.1. Definición de imagen hiperespectral

Una imagen hiperespectral está constituida por un conjunto de imágenes monocromáticas (variando en número en función de la resolución del sistema utilizado), que se unen formando un hipercubo, en el que dos dimensiones constituyen las dimensiones espaciales y la tercera componente la espectral (figura 13). Estas imágenes presentan mucha información redundante

Una imagen hiperespectral es una imagen en la que cada punto no viene descrito por un sólo valor de intensidad (como en una imagen "en blanco y negro") o por tres componentes de color (como en una imagen RGB de la pantalla del ordenador), sino por un vector de valores espectrales que se corresponden con la contribución de la luz detectada en ese punto a estrechas bandas del espectro, típicamente de entre 1 y 5 nm de ancho (el espectro visible abarca desde 400nm hasta 700 nm), en muchas ocasiones no están limitados estrictamente al rango visible sino también en el infrarrojo y el ultravioleta (no son raros los sensores que cubren desde los 300 nm hasta los 2500 nm).

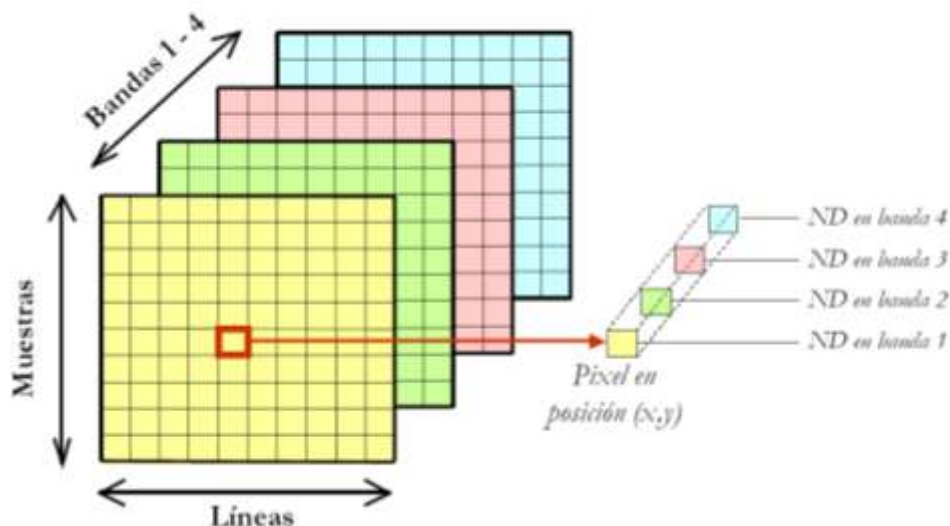


Figura 13. Concepto de una imagen hiperespectral.
Fuente: Elaboración Propia.

Una imagen hiperespectral es un registro de información, un conjunto de imágenes, basado en un mapa espacial de la variación espectral de la toma de imagen. Además estos cubos son base de datos tridimensionales compuestos por dos dimensiones espaciales (plano xy) y una dimensión espectral (plano λ). La dimensión espacial hace referencia a la morfología de la imagen adquirida, mientras la espectral corresponde a la composición de la misma.

El hipercubo está formado por imágenes contiguas, siendo cada una de ellas una representación de la intensidad y distribución espacial de la toma realizada (imagen) a una longitud de onda determinada.

De esta manera, cada píxel del hipercubo abarca el espectro respectivo según su posición espacial, a lo largo del rango de longitud de onda del cubo, que se convierte como una "huella digital" para caracterizar la composición del mismo. La figura 4 muestra la huella digital de un grano de cacao en la etapa de clasificación.

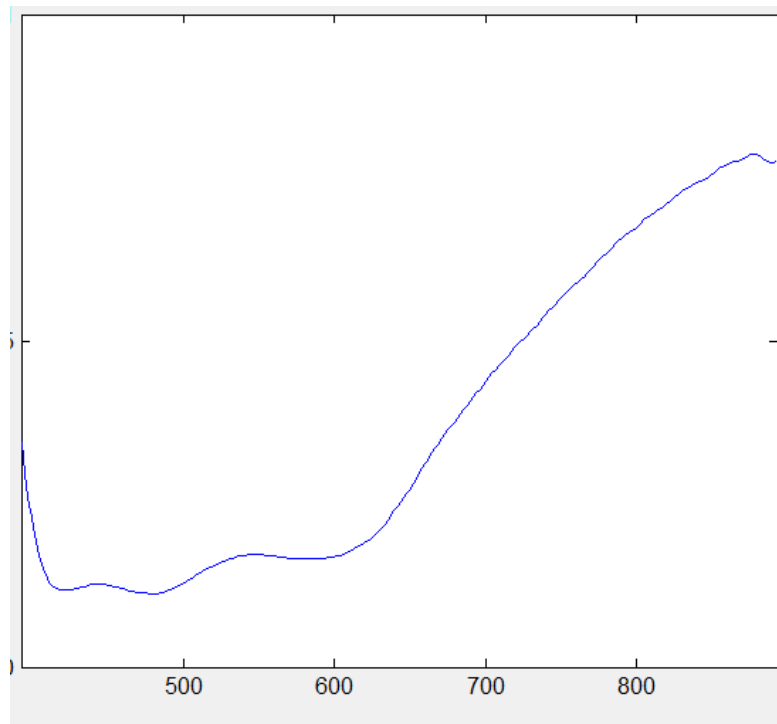


Figura 14. Firma digital o espectro de una región, foto adquirida de un cacao en una longitud de onda entre 400nm y 900nm.

Fuente: Elaboración Propia.

2.2.2. Definición de espectro

Cuando se hace pasar la luz a través de un prisma óptico se produce el efecto llamado dispersión que consiste en la separación de las distintas longitudes de onda que forman el rayo incidente.

La luz blanca al descomponerla forma lo que llamamos un espectro continuo, que contiene el conjunto de colores que corresponde a la gama de longitudes de onda que la integran.

El espectro visible para el ojo humano está en el rango de los 380nm de longitud de onda para el color violeta hasta los 780nm para el color rojo. Fuera de estos límites, el ojo no percibe ninguna clase de radiación. La figura siguiente muestra lo expuesto.

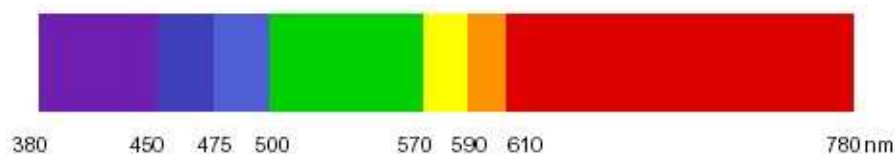


Figura 15. Espectro visible de la luz

Fuente: Elaboración Propia.

2.3. Calibración del sistema RESONON PIKA IIG.

La siguiente información describe como configurar el sistema para buscar la reflectancia adecuada de trabajo.

Las herramientas de calibración se encuentran en la barra de herramientas que proporciona el programa Spectronon Pro. Figura 16.



Figura 16. Barra de herramientas Spectronon Pro.
Fuente: Resonon Inc., (s/f). Elaboración: Propia.

Paso 1: Focalizar el lente

Se da click en el botón F localizada en la barra de calibración. Esto revelará una imagen del espectro capturado por la cámara. Uno de los ejes de esta imagen representa la (posición) eje espacial de su objeto, y el otro es el eje espectral (longitud de onda). Figura 17.

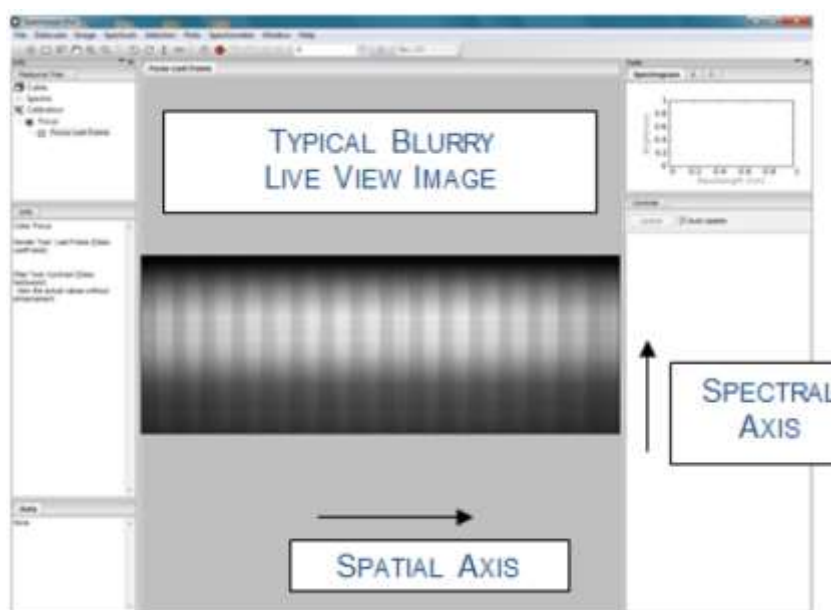


Figura 17. Imagen resultado de focalizar el lente.
Fuente: Resonon Inc., (s/f). Elaboración: Propia.

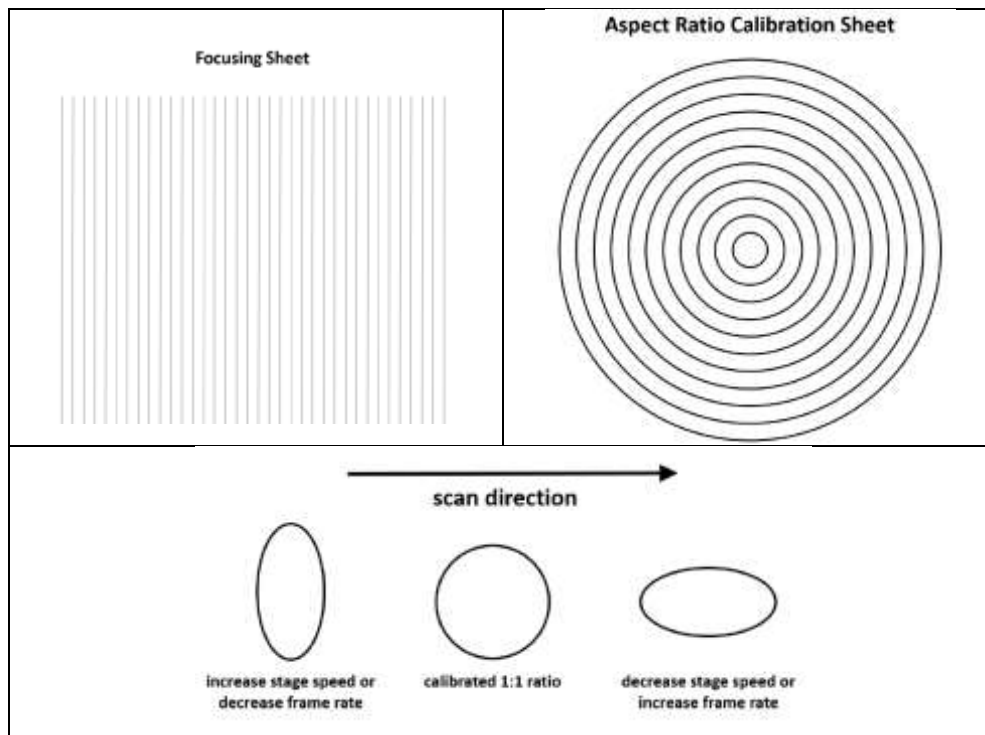


Figura 18. Plantillas para calibrar el lente.
Fuente: Resonon Inc., (s/f). Elaboración: Propia.

Paso 2: Remover la corriente oscura

La corriente oscura es la corriente eléctrica residual que fluye en un dispositivo fotoeléctrico cuando no hay iluminación incidente.

El programa de Spectronon Pro hace que sea fácil de quitar el ruido promedio de la corriente oscura en la toma del cubo hiperespectral. Para esto es necesario hacer click en el botón Dark Current (D) que se encuentra en la barras de herramientas, con esta herramienta se le darán instrucciones para bloquear toda la luz que entra a su espectrómetro de imagen por el bloqueo de la lente.



Figura 19. Ventana de aceptación Spectronon Pro.
Fuente: Resonon Inc., (s/f). Elaboración: Propia.

Para que la acción se pueda realizar la lente del objetivo debe estar bloqueado, luego se hace acepta las instrucciones. SpectrononPro recoge múltiples planos oscuros y usa estas medidas para restar el ruido ó corriente oscura.

Paso 3: Ajuste de reflectancia de referencia

La medición de la reflectancia absoluta de un objeto requiere de calibración para dar cuenta de los efectos de iluminación. El material de referencia debe ser uniforme. Ejemplos de materiales de referencia incluyen láminas de teflón blanco. Ver Figura 20.



Figura 20. Panel de referencia.

Fuente: Resonon Inc., (s/f). Elaboración: Propia.

Una vez que el material de referencia está en su lugar se realiza una corta exploración del material de referencia. La reflectancia será escalada a su material de referencia y compensa las variaciones espaciales de iluminación.

Paso 4: Ajustar el número de líneas a escanear.

Para escanear una imagen, se escribe el número de líneas que desea escanear. Un número de líneas de inicialización es a menudo de 200. Figura 21. Una cascada de imágenes aparecerá en el panel de imágenes de SpectrononPro.

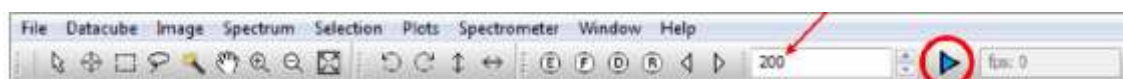


Figura 21. Barra de Herramientas Spetronon Pro.

Fuente: Resonon Inc., (s/f). Elaboración: Propia.

2.4. Procedimiento de Toma de muestra

La prueba de corte fue realizada por los técnicos de la asociación y separados en 3 grupos (bien fermentados, parcialmente fermentados y mal fermentados). La figura 22 muestra el lote de granos de cacao utilizados en el estudio, estos son tomados a la culminación del proceso de secado.



Figura 22. Granos de cacao en el final del proceso de secado.
Fuente: Elaboración Propia.

Asimismo, los productores suelen separar los granos en 2 calibres. Los granos del calibre 1 son más grandes que los del calibre 2. Esta clasificación es mecánica y se realiza por medio de una zaranda. A continuación en la Fig. 23 se muestra la imagen de la zaranda.



Figura 23. Sistema Zaranda par clasificación de cacao.
Fuente: Elaboración Propia.

A continuación en la Fig. 24 se muestra la imagen de la salida de la zaranda, se aprecia dos tamaños de clasificación.



Figura 24. Sistema Zaranda.
Fuente: Elaboración Propia.

Generalmente la prueba de corte se realiza antes del secado para determinar el día que los granos deben ser retirados de la fermentación. No obstante, nuestro objetivo es estudiar la calidad final de los granos. Por este motivo, esta prueba se realizó después del secado. Cada uno de los granos cortados fue etiquetado por el grado de fermentación y calibre al que pertenece. Un total de 30 muestras fueron estudiadas. Se tomaron 3 capturas de cada grano en diferentes posiciones. Las imágenes fueron adquiridas desde una cámara hiperspectral Resonon Inc., modelo Pika IIg, U.S.A.



Figura 25. Sistema Resonon Pro usado para el análisis de los parámetros internos de los granos de cacao.
Fuente: Elaboración Propia.

La iluminación se ha dado usando lámparas halógenas (Figura 26). Cada captura tiene una dimensión espacial de 200x640 píxeles y consta de 240 canales espectrales en el rango de 400nm a 900nm.



Figura 26. Fuente de iluminación del sistema Resonon Pro.
Fuente: Elaboración Propia.

Los cubos fueron capturados por medio del software Spectronon (Resonon, U.S.A.) y luego exportados a MATLAB v13a (MathWorks, U.S.A.).

Se extrajeron 3 dimensiones del cubo a analizar: canal R en 860nm, canal G en 650nm y canal B en 550nm para formar una imagen RGB (Figura 27 y Figura 27E); se realizó la operación de segmentación obteniendo una región de interés (ROI), delimitándola por medio de una figura rectangular.

Las dimensiones de esta ROI permitieron formar un sub-cubo hiperespectral de 240 planos para cada grano, siendo el tamaño del vector ROI diferente. Los 240 vectores de cada grano fueron sometidos a un análisis de componentes principales (PCA), donde la primera componente (PC1) contenía cerca del 95 % de la información del grano, la segunda componente (PC2) el 2 % y la tercera componente (PC3) el 0.1 % de la información. La nueva imagen formada por PC1 (Fig. 27B y Fig. 27F) brinda suficiente información para diferenciar claramente las grietas del grano, las cuales tienen una correlación negativa con esta componente.

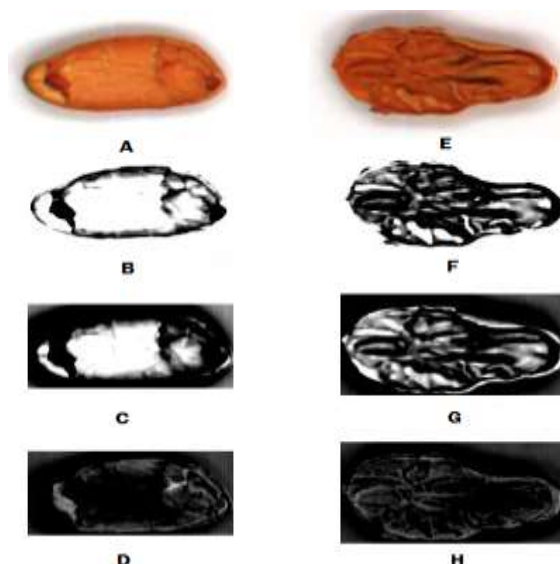


Figura 27. (A) Imagen NIR grano sin fermentar, (B) componente PC1, (C) componente PC2. (D) componente PC3. (E) imagen NIR grano bien fermentado, (F) componente PC1. (G) componente PC2. (H) componente PC3.
Fuente: Elaboración Propia.

La característica relacionada con las grietas o estrías y la ausencia del cotiledón se relaciona con la calidad del grano de cacao, como se mencionó la formación de estas es debido al buen proceso de fermentación y secado.

Para extraer la característica relacionada con las grietas, se segmentó la zona central grano (lugar donde se encuentra el cotiledón), se calculó el área y la relación que tiene con el área total del grano. Para la característica relacionada con el color interno, se segmentaron las regiones no agrietadas y se determinó su espectro medio.

Un nuevo análisis PCA se realizó por separado para estudiar estos espectros medios. Los resultados indicaron que las dimensiones 123 hasta la 189 exponían una mejor diferencia entre los 3 niveles de fermentación, siendo la dimensión 154 la de mayor peso. De esta manera, se consideró esta región del espectro medio, para formar un índice que cuantificara el color interno del grano.

Al finalizar esta etapa, se cuenta con 2 índices que son las entradas de la siguiente fase, estos son: el índice que describe la relación entre el área central agrietada y el área total; y el índice que describe el color interno.

2.5. Clasificador basado en Fuzzy Logic

La lógica difusa ha llegado a obtener una gran fama debido a la gran variedad de aplicaciones que tiene, que van desde el control de procesos industriales complejos, hasta el diseño de una serie de dispositivos artificiales que presentan deducción automática, como serían artefactos electrónicos ya sea para uso doméstico o de entretenimiento; así como también de sistemas de diagnóstico.

Se considera que el concepto de lógica difusa apareció en 1965, de forma general, en la Universidad de California en Berkeley, introducido por Lotfi A. Zadeh.

La lógica clásica implica enunciados totalmente falso o verdadero, los cuales han servido para modelar gran parte del razonamiento natural; pero en el razonamiento humano, no se es tan determinista con los valores de verdad. Por ejemplo, cuando se hace una calificación como: el cielo es azul, este enunciado nos llevaría a preguntarnos qué tan azul, aunque esto no implique cuantificar en una escala dicho valor. La lógica difusa se adapta mejor al mundo real donde vivimos, e incluso puede comprender y operar con expresiones que utilizamos como por ejemplo:

“Hace mucho calor”, “no es muy alto”, etc. Ésta adaptación a nuestro lenguaje se basa en la comprensión de nuestros cuantificadores (mucho, poco, regular, etc.)

La lógica difusa generalmente trata de crear aproximaciones matemáticas para la solución de ciertos problemas a partir de datos imprecisos, por lo que se emplean mucho en aplicaciones electrónicas o computacionales. La lógica difusa tiene que ver con el uso de valores de verdad no-deterministas, los cuales tienen una connotación de incertidumbre. A sus enunciados se les asocia con valores de verdad que son grados de veracidad o falsedad, mucho más amplios que los simplemente verdadero o falso. En general la teoría difusa provee un mecanismo para la construcción de representaciones lingüísticas; es decir, de estructuras de inferencia que son apropiadas en nuestro lenguaje y razonamiento diario.

Conjuntos difusos

Son conjuntos que contienen elementos de manera parcial, que a diferencia de la lógica clásica contiene totalmente un elemento. En un conjunto difuso, se relaciona el grado de pertenencia de cada elemento del universo a este conjunto. Este grado de pertenencia resulta ser un número entre 0 y 1. Cuanta más cerca se encuentre el grado de pertenencia a 1, el elemento pertenecerá más al conjunto y en cuanto más cercano a 0, el elemento menos pertenecerá al conjunto.

Operaciones en conjuntos difusos

Subconjunto

$$A \subseteq B \Leftrightarrow \mu_A(x) \leq \mu_B(x), \quad \forall x \in U$$

Complemento

$$\underline{A} = U - A \Leftrightarrow \mu_{\underline{A}}(x) = 1 - \mu_A(x)$$

Unión

$$C = A \cup B \Leftrightarrow \mu_C(x) = \max(\mu_A(x), \mu_B(x)) = \mu_A(x) \vee \mu_B(x)$$

Intersección

$$C = A \cap B \Leftrightarrow \mu_C(x) = \min(\mu_A(x), \mu_B(x)) = \mu_A(x) \wedge \mu_B(x)$$

Función de pertenencia

Los grados de pertenencia de un elemento cualquiera a un conjunto difuso se les conocen como funciones de pertenencia.

Sea U un universo de discurso, por decir $U = R^n$, un conjunto difuso A se define con la función de pertenencia $\mu_A(x)$:

$U \rightarrow [0,1]$ donde $\mu_A(x)$, representa el grado en que el elemento $x \in U$ pertenece al conjunto difuso A

Se recomienda generalmente el uso de funciones simples, que busquen simplificar los cálculos matemáticos pero que no pierdan exactitud, ya que precisamente tratan de definir un conjunto difuso. Las funciones de pertenencia más comunes son:

Triangular: Definido por los límites inferior (a) y superior (b), y el valor modal (m), tal que $a < m < b$

$$\mu_A(x) = \begin{cases} 0 & \text{si } x < a \\ \frac{(x-a)}{(m-a)} & \text{si } x \in (a, m] \\ \frac{(b-x)}{(b-m)} & \text{si } x \in (m, b) \\ 0 & \text{si } x \geq b \end{cases}$$

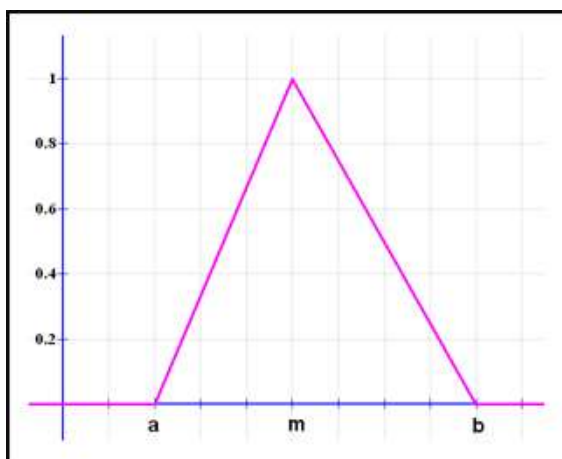


Figura 28. Función de pertenencia triangular.
Fuente: Elaboración Propia.

Trapezoidal: Definida por un límite inferior (a), un límite superior (d), un límite soporte inferior (b), y un límite de soporte superior c , tal que $a < b < c < d$

$$\mu_A(x) = \begin{cases} 0 & \text{si } (x < a) \text{ ó } (x > d) \\ \frac{(x-a)}{(b-a)} & \text{si } a \leq x \leq b \\ 1 & \text{si } b \leq x \leq c \\ \frac{(d-x)}{(d-c)} & \text{si } c \leq x \leq d \end{cases}$$

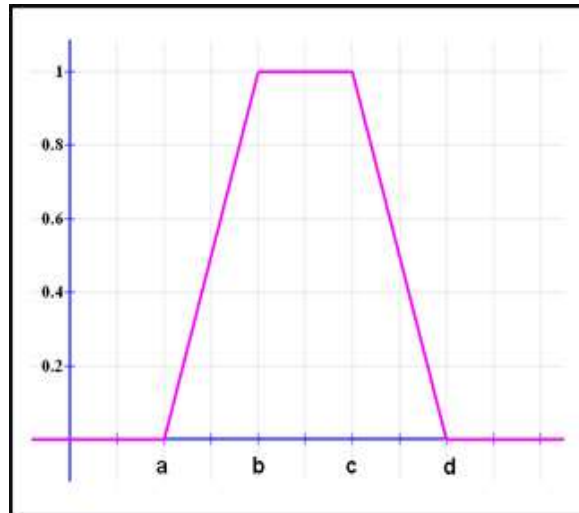


Figura 29. Función de pertenencia trapezoidal.
Fuente: Elaboración Propia.

Se conocen dentro de este tipo de función, dos casos particulares:

Función R: De parámetros $a = b = -\infty$

$$\mu_A(x) \begin{cases} 0 & \text{si } x > d \\ \frac{(d-x)}{(d-c)} & \text{si } c \leq x \leq b \\ 1 & \text{si } x < c \end{cases}$$

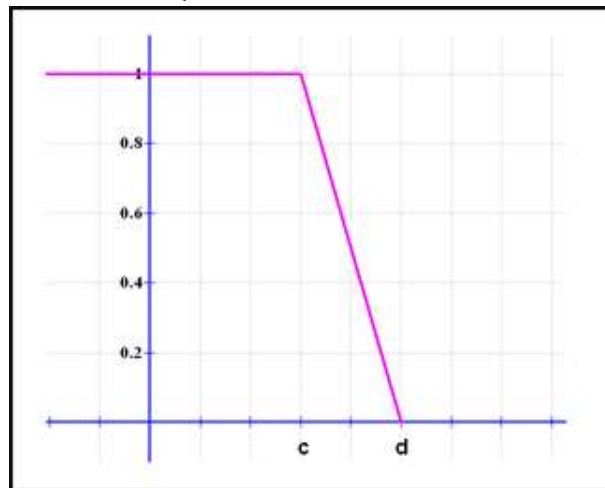


Figura 30. Función de pertenencia trapezoidal tipo R.
Fuente: Elaboración Propia.

Función L: De parámetros $c = d = +\infty$

$$\mu_A(x) \begin{cases} 0 & \text{si } x < a \\ \frac{(x-a)}{(b-a)} & \text{si } a \leq x \leq b \\ 1 & \text{si } x > b \end{cases}$$

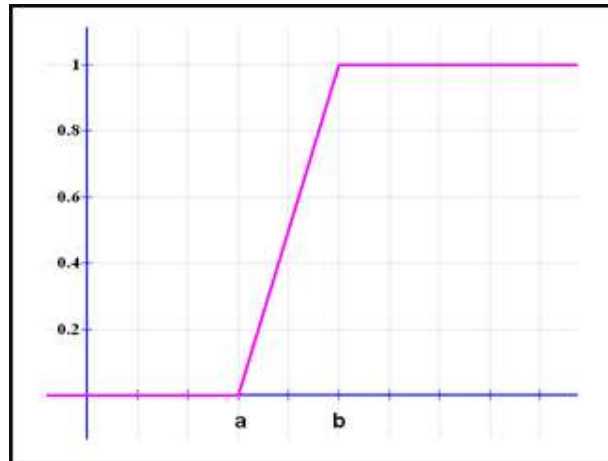


Figura 31. Función de pertenencia trapezoidal tipo L.
Fuente: Elaboración Propia.

Gaussiana: Definida por un valor medio (m) y una desviación estándar ($k > 0$). Donde se cumple que cuanto menor sea K , más estrecha será la campana de Gauss.

$$\mu_A(x) = e^{\frac{-(x-m)^2}{2k^2}}$$

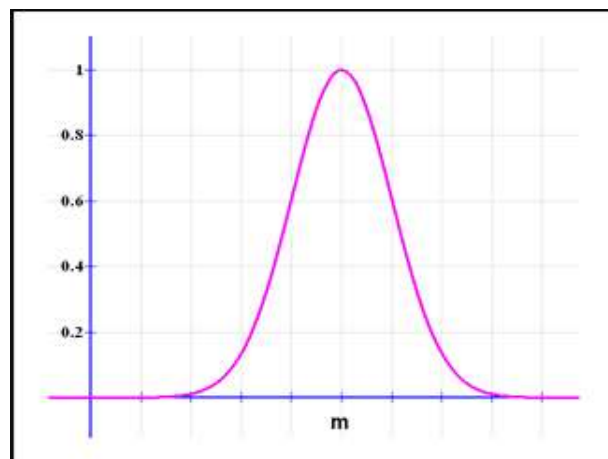


Figura 32. Función de pertenencia gaussiana.
Fuente: Elaboración Propia.

En este trabajo se presenta el uso de esta técnica basada en lógica difusa aplicándolo como un bloque clasificador, cuyas dos entradas y salida son: el porcentaje de área de grietas, el color y la calidad respectivamente. A continuación se explica cada una de estas variables.

El índice relacionado con las grietas del grano fue descrito en el rango de $[0,1]$ en 4 funciones de membresía: bajo, medio, alto y elevado; cada una del tipo trapezoidal.

El índice relacionado con el color interno fue descrito en el mismo rango y tipo. Se consideraron 3 funciones de membresía: bajo, medio y alto.

La salida, relacionada con la calidad del grano, fue descrita en el mismo rango. Se estimaron 3 funciones de membresía: malo, parcial y bueno. La función “parcial” fue descrita del tipo triangular. Las otras dos funciones del tipo trapezoidal. Además, un total de 12 reglas fueron planteadas (Tabla 2). Por último, el método de la bisectriz fue considerado para la defusificación.

Table 2. Reglas de lógica difusa

	Index 1	Index 2	Quality
1	LOW	LOW	UNFERMENTED
2	LOW	MEDIUM	PARTLY
3	LOW	HIGH	UNFERMENTED
4	MEDIUM	LOW	FERMENTED
5	MEDIUM	MEDIUM	PARTLY
6	MEDIUM	HIGH	UNFERMENTED
7	HIGH	LOW	FERMENTED
8	HIGH	MEDIUM	PARTLY
9	HIGH	HIGH	UNFERMENTED
10	HIGHER	LOW	FERMENTED
11	HIGHER	MEDIUM	FERMENTED
12	HIGHER	HIGH	UNFERMENTED

2.6. Resultados

A continuación se muestra el resultado de la clasificación para la determinación de la calidad interna del grano. En la figura 33 se muestra la firma espectral del grano y el resultado de esta clasificación; en este caso el grano está bien fermentado. La técnica de clasificación por lógica difusa permite darle un valor al grado de calidad del grano.

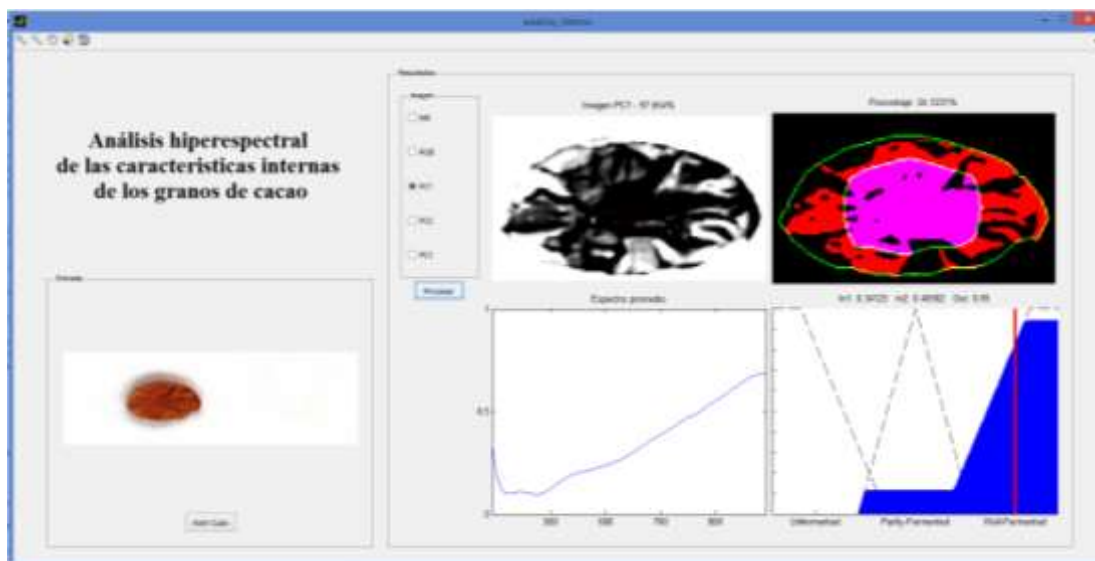


Figura 33. Resultado de la evaluación del algoritmo de determinación de calidad del grano, el clasificador está basado en técnicas de Lógica Difusa.

Fuente: Elaboración Propia.

En la figura 34 se muestra el toolbox de la herramienta de lógica difusa. Este toolbox permite ingresar las funciones de membresía, las reglas de fusificación y defusificación. Se han dado dos entradas al sistema, color y el porcentaje de grietas, como resultado indica si el grano ha sido bien fermentado, parcialmente fermentado o mal fermentado.

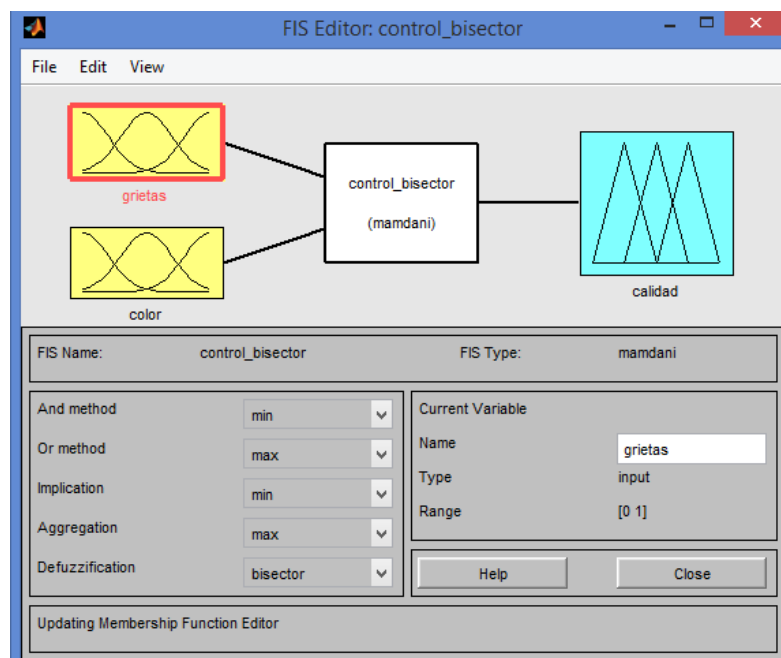


Figura 34. Sistema Toolbox de Lógica Difusa – Matlab.
Fuente: Elaboración Propia.

La Figura 35 muestra las funciones de membresía de la variable grietas.

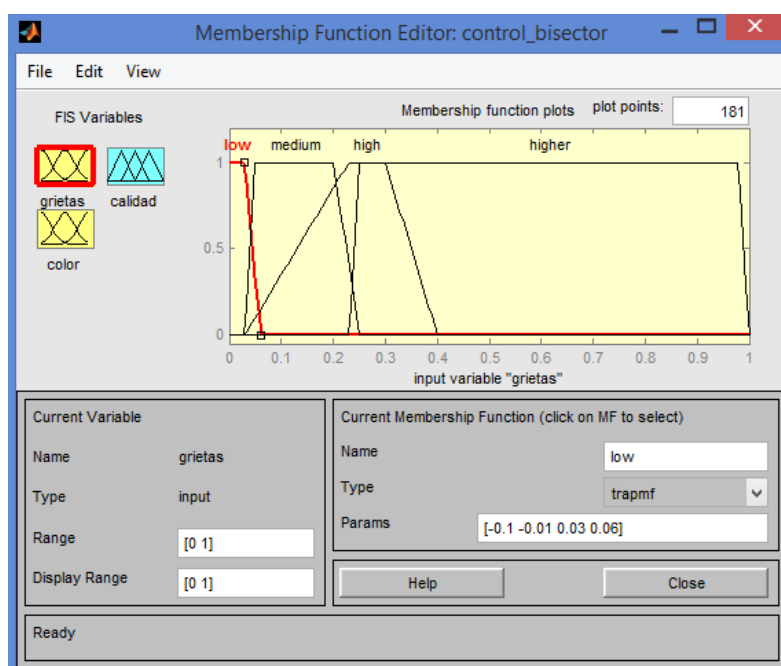


Figura 35. Funciones de membresía de la variable de entrada grietas.
Fuente: Elaboración Propia.

La Figura 36 muestra las funciones de membresía de la variable color.

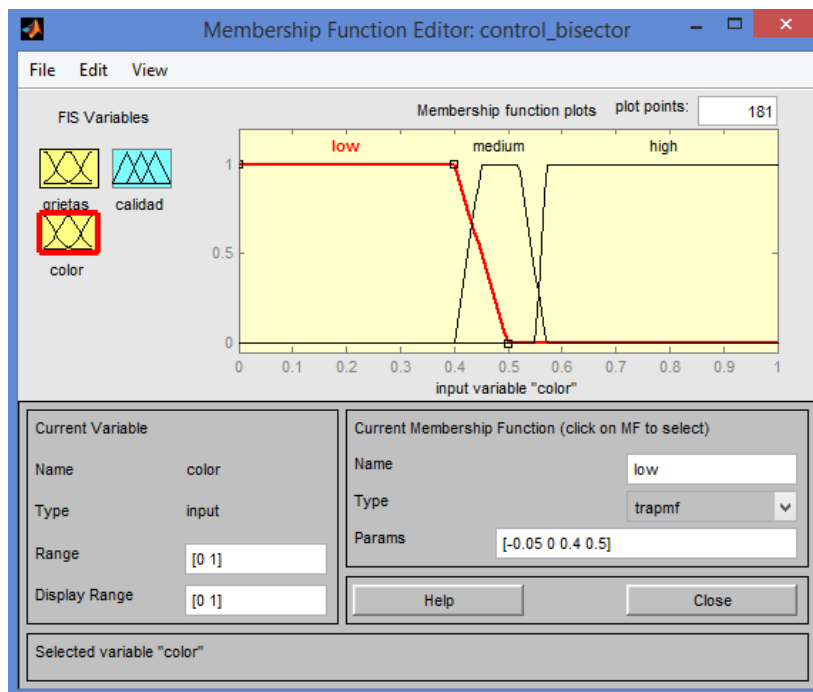


Figura 36. Funciones de membresía de la variable de entrada color.

Fuente: Elaboración Propia.

La figura 37 muestra las funciones de membresía de la variable de salida calidad. Se le ha dado Malo, Parcial y Bueno.

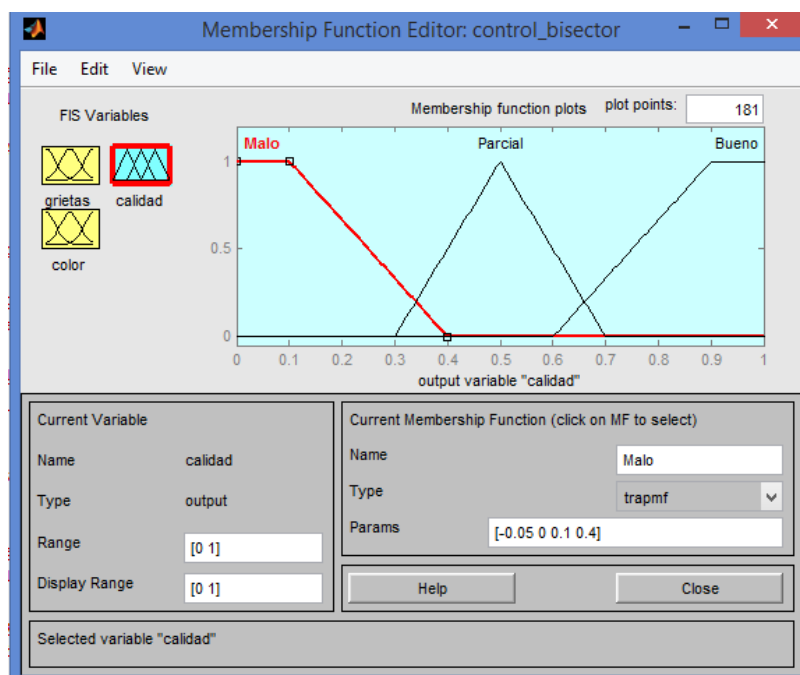


Figura 37. Funciones de membresía de la variable de salida calidad.

Fuente: Elaboración Propia.

A continuación se presentan los resultados obtenidos del algoritmo el cual analiza las características internas del cacao.

Se obtuvieron una muestra de granos de cacao los cuales fueron clasificados con personal especialista de la asociación. Con ellos se analizó cualitativamente cada grano de cacao y se etiquetaron para su reconocimiento en laboratorio.

La Figura 38, muestra el cacao #1, el cual se aprecia que es un cacao que no ha tenido una buena fermentación; esto debido a que las estrías no están bien formadas, así también no se aprecia la pérdida del cotiledón en la parte central del grano.



Figura 38. Imagen del grano de cacao #1 patrón.
Fuente: Elaboración Propia.

En la figura 39 muestra el resultado del algoritmo, en donde se clasifica al grano, la clasificación visual va de acorde con los resultados del algoritmo y asocia un valor cuantitativo al grado de calidad.

A continuación mostramos los resultados de clasificación, se aprecia que el algoritmo le asocia un valor cuantitativo al grano de cacao, es decir, asocia un grado de bueno, parcial o malo. El clasificador usado ha sido el de Lógica Difusa, otro clasificador como el PCA no dio buenos resultados.

A continuación se presentan el resultado de los 10 primeros granos de cacao que se tienen en nuestra base de datos, estas imágenes servirán de patrones para la correlación.

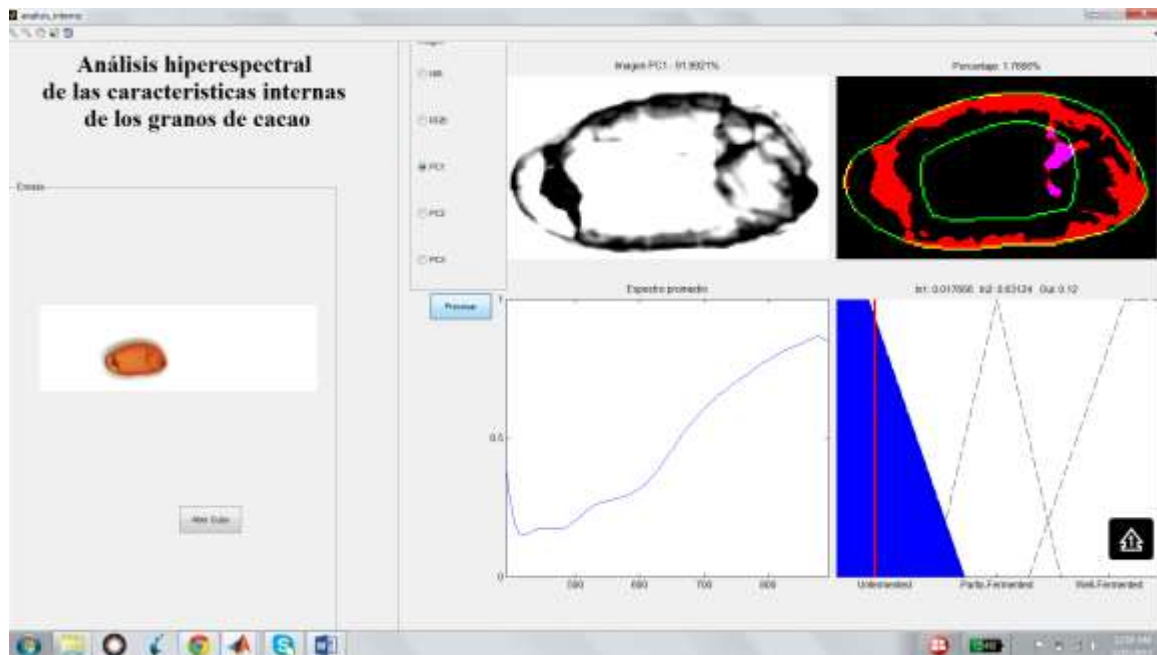


Figura 39. Imagen resultado del grano de cacao #1 patrón y su clasificación MALO.

Fuente: Elaboración Propia.

La figura 40, muestra el cacao #2, el cual se aprecia que es un cacao que SI ha tenido una buena fermentación; esto debido a que las estrías que están bien formadas, y se aprecia la pérdida del cotiledón en la parte central del grano.



Figura 40. Imagen del grano de cacao #2 patrón.

Fuente: Elaboración Propia.

En la figura 41 muestra el resultado del algoritmo, en donde se clasifica al grano, la clasificación visual va de acorde con los resultados del algoritmo y asocia un valor cuantitativo al grado de calidad, en este caso BUENO.

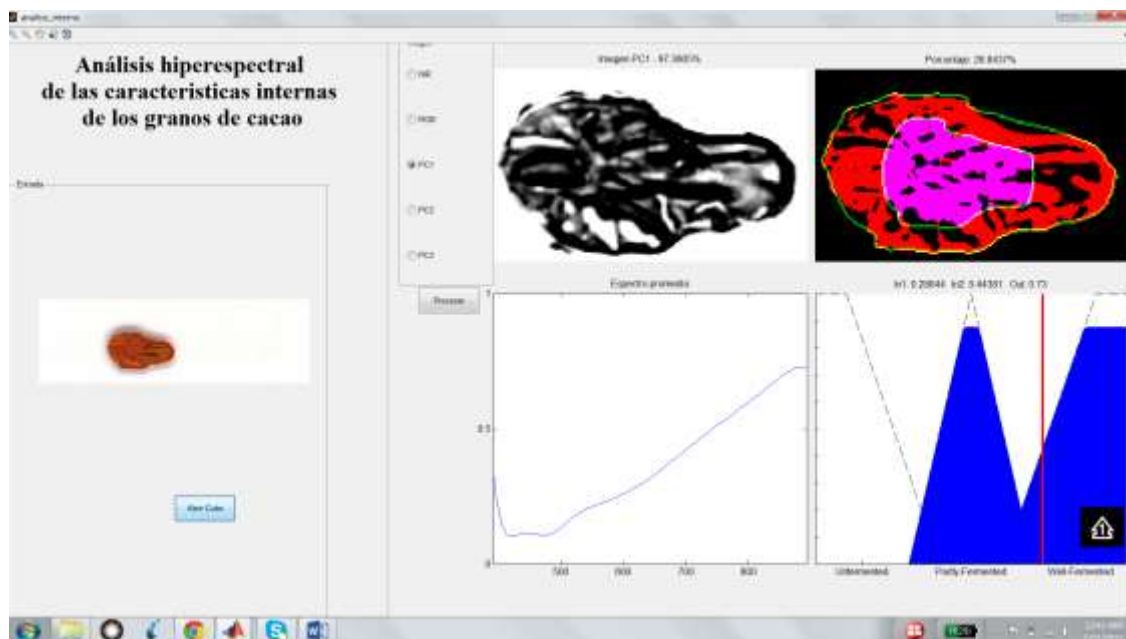


Figura 41. Imagen resultado del grano de cacao #1 patrón y su clasificación BUENO.
Fuente: Elaboración Propia.

La figura 42, muestra el cacao #3, el cual se aprecia que es un cacao que SI ha tenido una buena fermentación; esto debido a que las estrías que están bien formadas, y se aprecia la pérdida del cotiledón en la parte central del grano.



Figura 42. Imagen del grano de cacao #3 patrón.
Fuente: Elaboración Propia.

En la figura 43 muestra el resultado del algoritmo, en donde se clasifica al grano, la clasificación visual va de acorde con los resultados del algoritmo y asocia un valor cuantitativo al grado de calidad, en este caso BUENO.

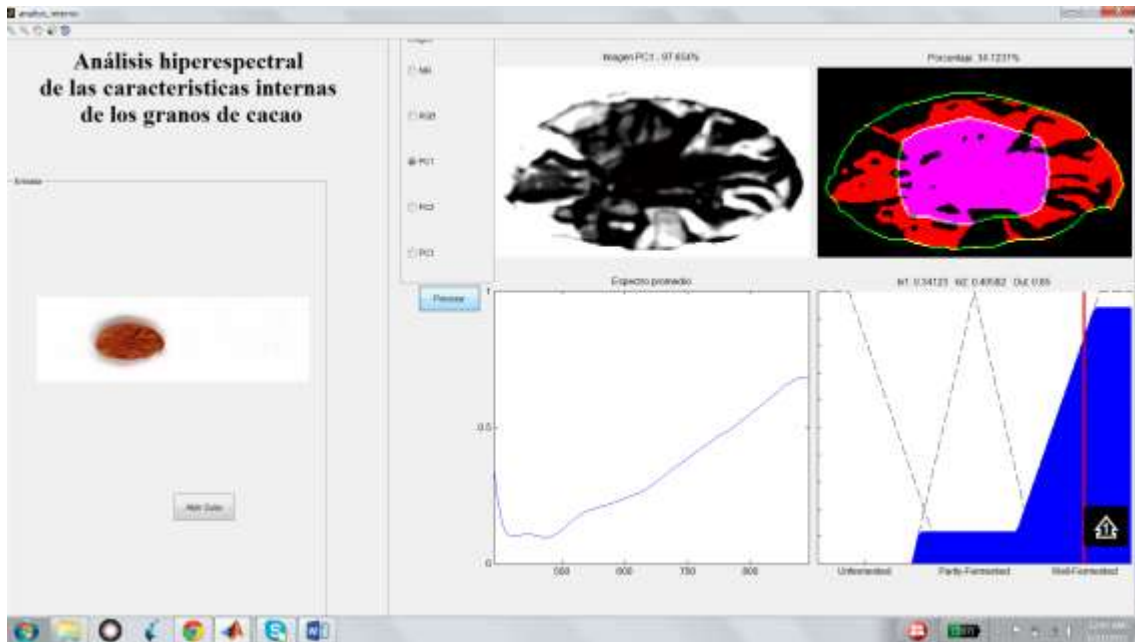


Figura 43. Imagen resultado del grano de cacao #1 patrón y su clasificación BUENO.
Fuente: Elaboración Propia.

La figura 44, muestra el cacao #4, el cual se aprecia que es un cacao que SI ha tenido una buena fermentación; esto debido a que las estrías que están bien formadas, y se aprecia la pérdida del cotiledón en la parte central del grano.



Figura 44. Imagen del grano de cacao #4 patrón.
Fuente: Elaboración Propia.

En la figura 45 muestra el resultado del algoritmo, en donde se clasifica al grano, la clasificación visual va de acorde con los resultados del algoritmo y asocia un valor cuantitativo al grado de calidad, en este caso BUENO.

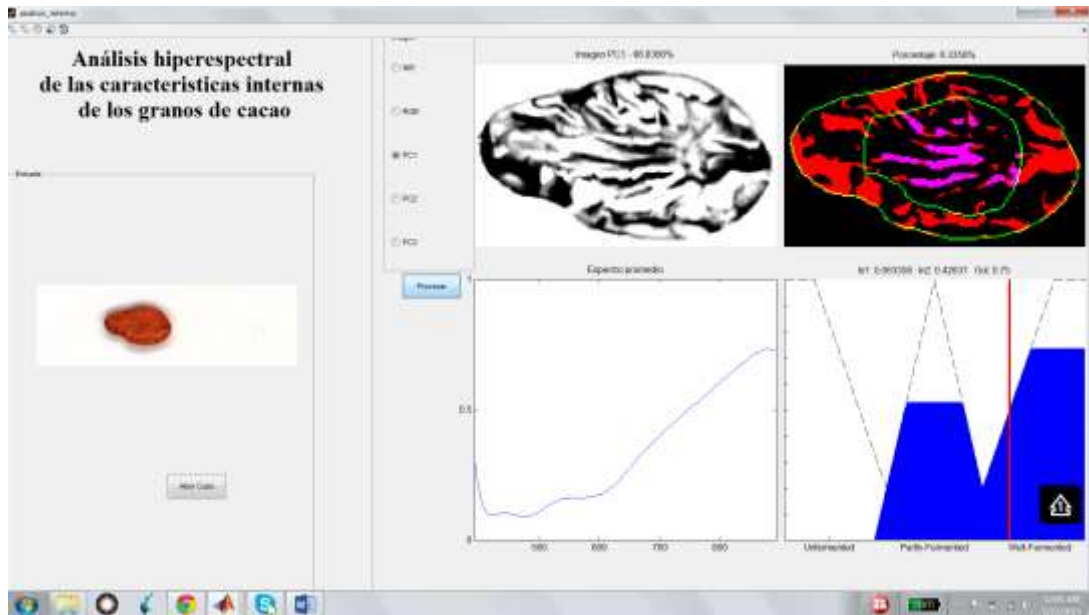


Figura 45. Imagen resultado del grano de cacao #1 patrón y su clasificación BUENO.
Fuente: Elaboración Propia.

La figura 46, muestra el cacao #5, el cual se aprecia que es un cacao que ha tenido una fermentación que ha sido parcial; esto debido a que las estrías que están del todo bien formadas, y se aprecia la pérdida del cotiledón no está del todo ausente en la parte central del grano.



Figura 46. Imagen del grano de cacao #5 patrón.
Fuente: Elaboración Propia.

En la figura 47 muestra el resultado del algoritmo, en donde se clasifica al grano, la clasificación visual va de acorde con los resultados del algoritmo y asocia un valor cuantitativo al grado de calidad, en este caso PARCIAL.

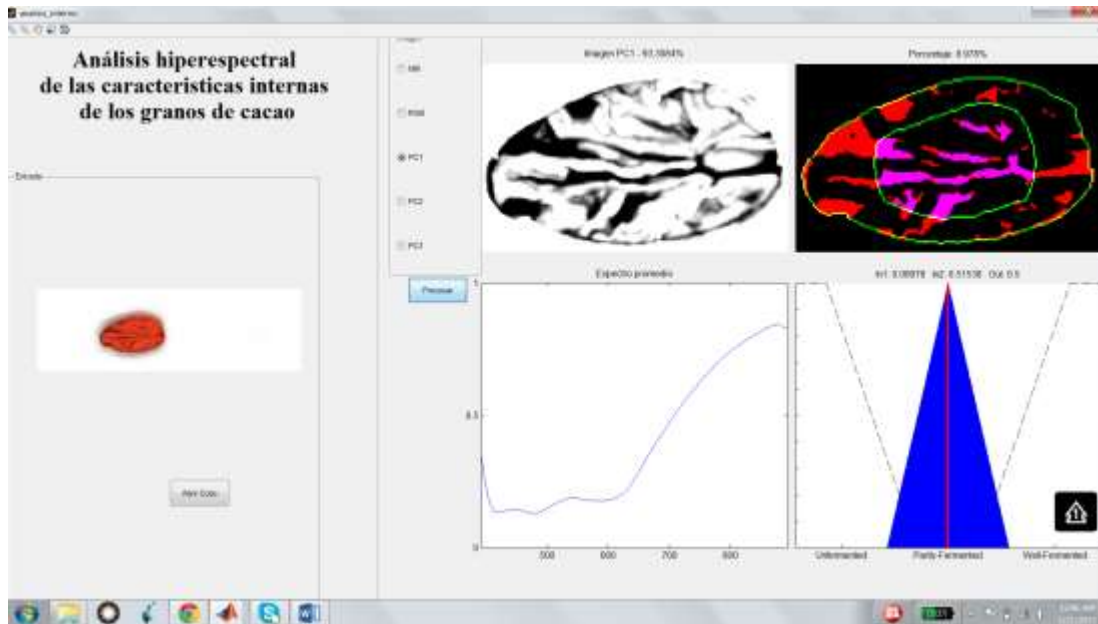


Figura 47. Imagen resultado del grano de cacao #1 patrón y su clasificación PARCIAL.
Fuente: Elaboración Propia.

La figura 48, muestra el cacao #6, el cual se aprecia que es un cacao que ha tenido una fermentación que ha sido parcial; esto debido a que las estrías que están del todo bien formadas, y se aprecia la pérdida del cotiledón no está del todo ausente en la parte central del grano.



Figura 48. Imagen del grano de cacao #6 patrón.
Fuente: Elaboración Propia.

En la figura 49 muestra el resultado del algoritmo, en donde se clasifica al grano, la clasificación visual va de acorde con los resultados del algoritmo y asocia un valor cuantitativo al grado de calidad, en este caso PARCIAL.

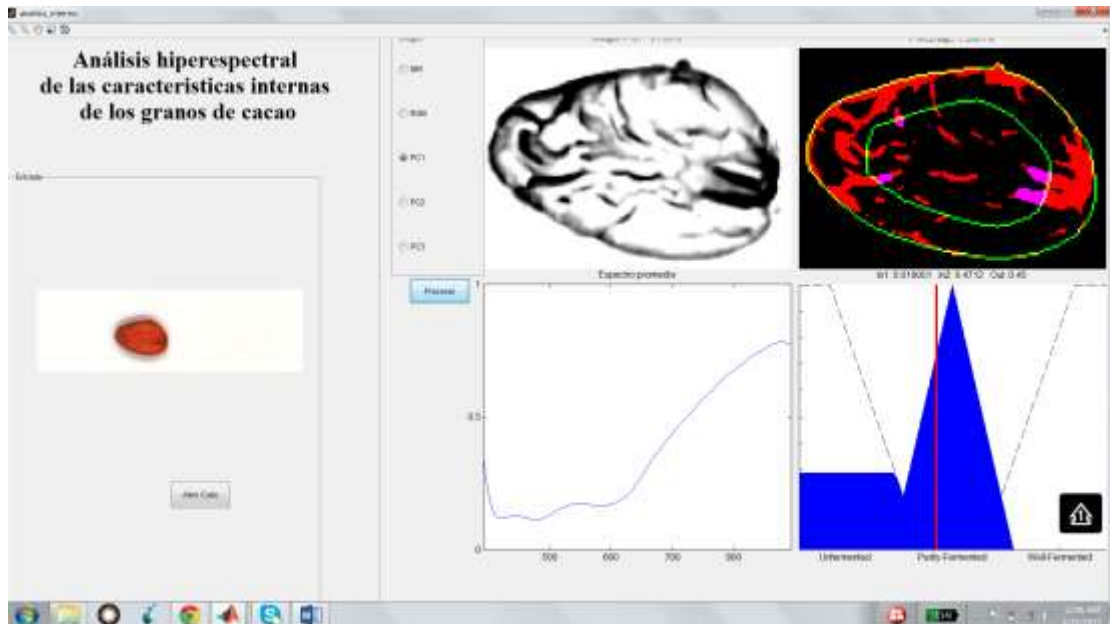


Figura 49. Imagen resultado del grano de cacao #1 patrón y su clasificación PARCIAL.
Fuente: Elaboración Propia.

La figura 50, muestra el cacao #7, el cual se aprecia que es un cacao que no ha tenido una buena fermentación; esto debido a que las estrías no están bien formadas, así también no se aprecia la pérdida del cotiledón en la parte central del grano.



Figura 50. Imagen del grano de cacao #7 patrón.
Fuente: Elaboración Propia.

En la figura 51 muestra el resultado del algoritmo, en donde se clasifica al grano, la clasificación visual va de acorde con los resultados del algoritmo y asocia un valor cuantitativo al grado de calidad, en este caso MALO.

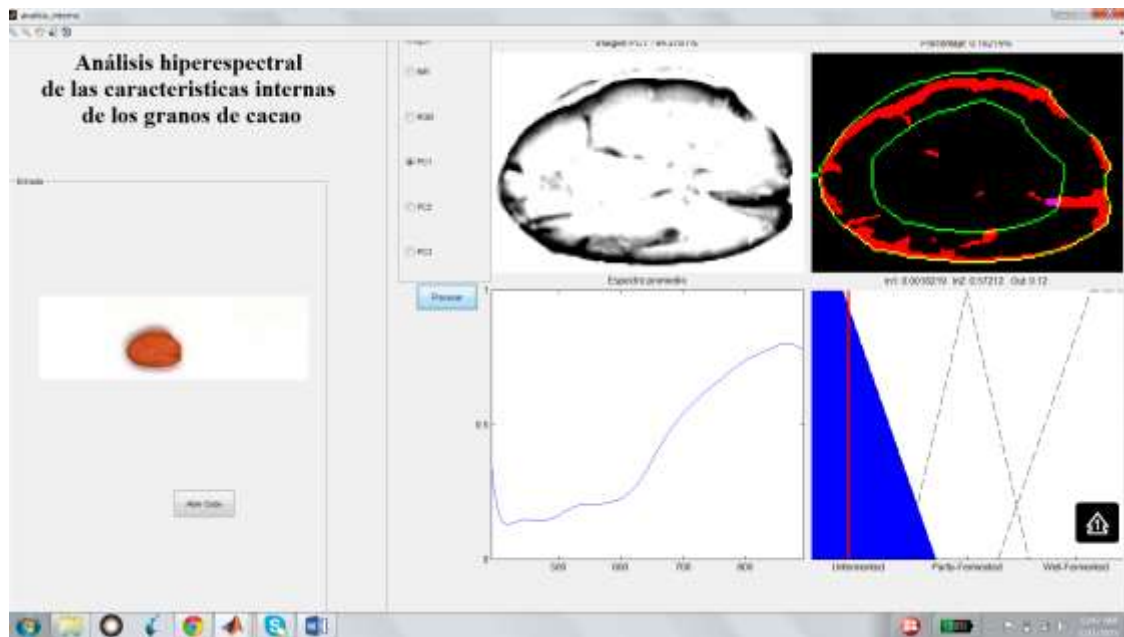


Figura 51. Imagen resultado del grano de cacao #1 patrón y su clasificación MALO.
Fuente: Elaboración Propia.

La figura 52, muestra el cacao #8, el cual se aprecia que es un cacao que no ha tenido una buena fermentación; esto debido a que las estrías no están bien formadas, así también no se aprecia la pérdida del cotiledón en la parte central del grano.



Figura 52. Imagen del grano de cacao #8 patrón.
Fuente: Elaboración Propia.

En la figura 53 muestra el resultado del algoritmo, en donde se clasifica al grano, la clasificación visual va de acorde con los resultados del algoritmo y asocia un valor cuantitativo al grado de calidad, en este caso MALO.

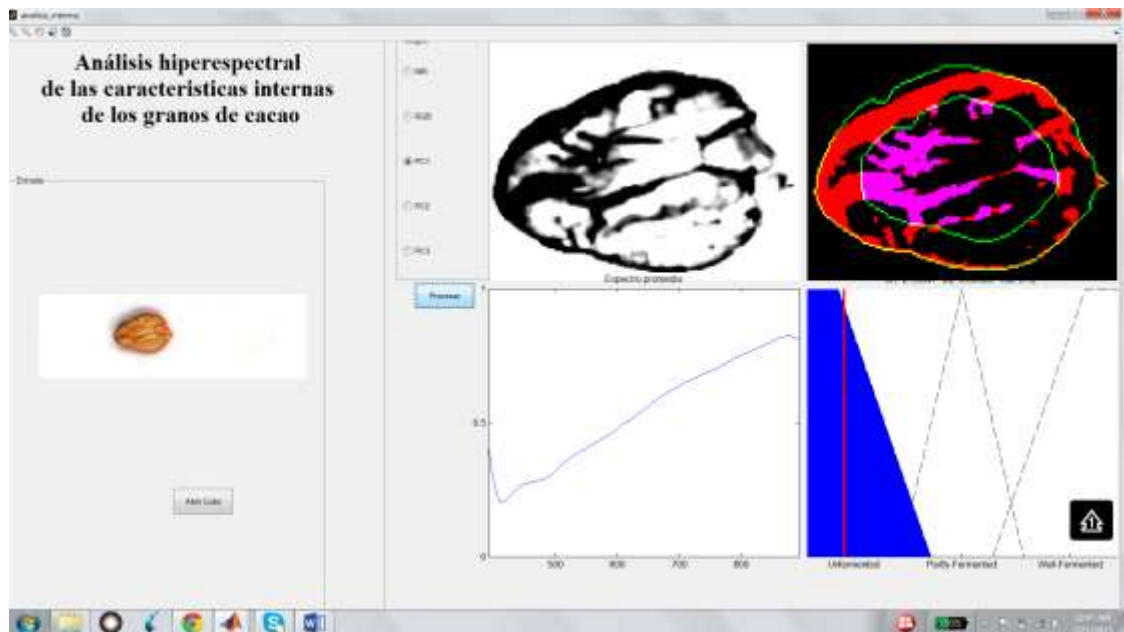


Figura 53. Imagen resultado del grano de cacao #1 patrón y su clasificación MALO.
Fuente: Elaboración Propia.

La figura 54, muestra el cacao #9, el cual se aprecia que es un cacao que ha tenido una fermentación que ha sido parcial; esto debido a que las estrías que están del todo bien formadas, y se aprecia la pérdida del cotiledón no está del todo ausente en la parte central del grano.



Figura 54. Imagen del grano de cacao #9 patrón.
Fuente: Elaboración Propia.

En la figura 55 muestra el resultado del algoritmo, en donde se clasifica al grano, la clasificación visual va de acorde con los resultados del algoritmo y asocia un valor cuantitativo al grado de calidad, en este caso PARCIAL.

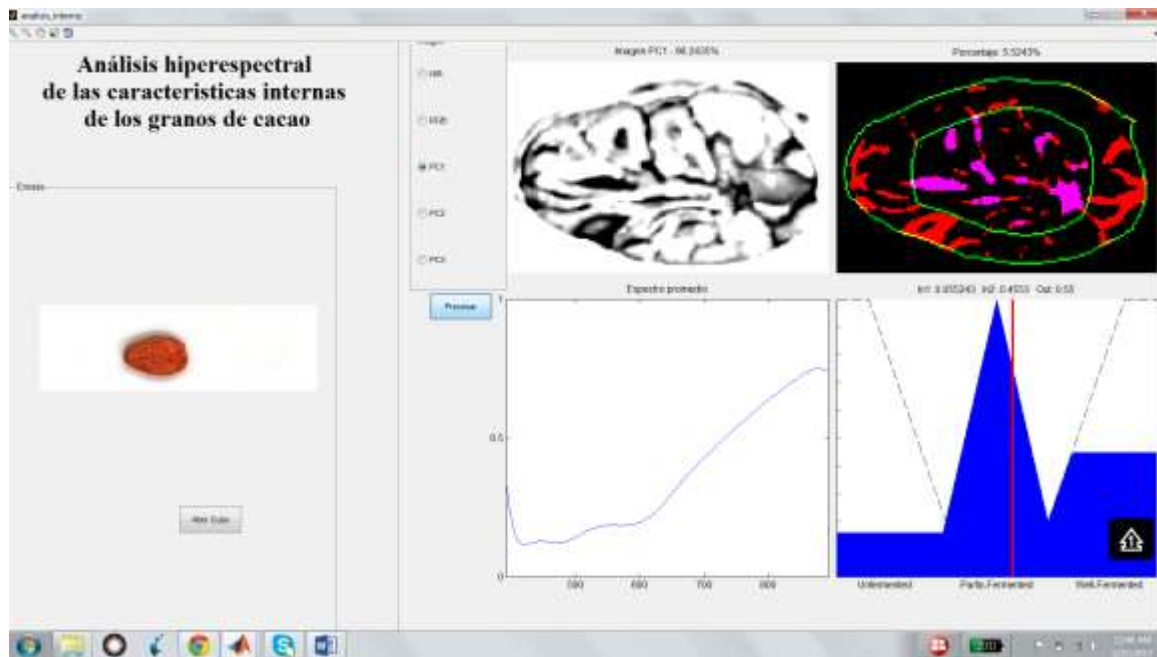


Figura 55. Imagen resultado del grano de cacao #1 patrón y su clasificación PARCIAL.
Fuente: Elaboración Propia.

La figura 56, muestra el cacao #10, el cual se aprecia que es un cacao que ha tenido una fermentación que ha sido parcial; esto debido a que las estrías que están del todo bien formadas, y se aprecia la pérdida del cotiledón no está del todo ausente en la parte central del grano.



Figura 56. Imagen del grano de cacao #10 patrón.
Fuente: Elaboración Propia.

En la figura 57 muestra el resultado del algoritmo, en donde se clasifica al grano, la clasificación visual va de acorde con los resultados del algoritmo y asocia un valor cuantitativo al grado de calidad, en este caso PARCIAL.

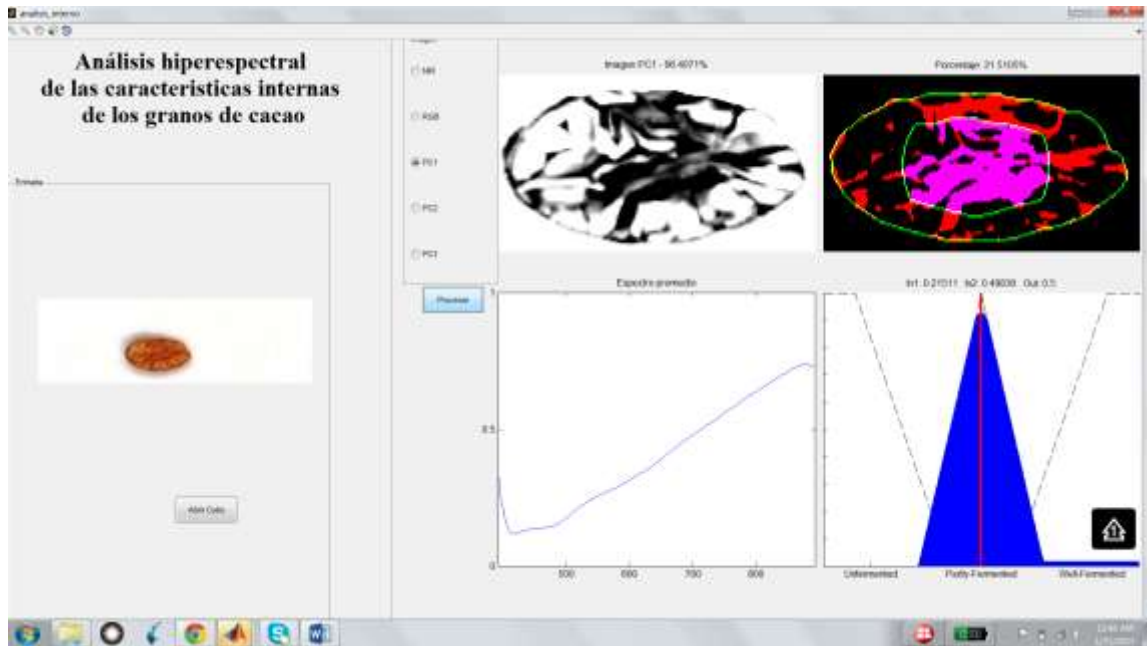


Figura 57. Imagen resultado del grano de cacao #1 patrón y su clasificación PARCIAL.
Fuente: Elaboración Propia.

Se muestra los resultados de la validación del código desarrollado. La Tabla 3 muestra los resultados de las pruebas realizadas con granos considerados en buen estado.

Table 3. Resultados con granos buenos.

	C.internas				C.externas			
	Porcentaje	R	G	B	Ancho	R/r	Área	Num.
a1bueno	0	80	36	28	145.15	1.52	25	0
a2bueno	0	112	46	20	135.47	1.78	25.33	0
a3bueno	0.41116	69	31	20	144.63	1.62	26.45	0
a4bueno	0	113	47	20	141.5	1.65	25.87	0
a5bueno	0.043074	84	36	22	146.81	1.64	27.47	0
a6bueno	0.0083938	84	37	23	141.53	1.5	23.37	0
a7bueno	0.081776	129	60	29	129.98	1.82	24.04	0
a8bueno	0.0082991	96	38	15	124.67	1.88	22.81	0

La Tabla 4 muestra los resultados de las pruebas realizadas con granos considerados de descarte.

Table 4. Resultados con granos de descarte.

	C.internas				C.externas			
	Porcentaje	R	G	B	Ancho	R/r	Área	Num.
b1descarte	0.23906	65	34	31	127.72	1.93	24.58	1
b2descarte	0	114	48	22	171.7	1.49	34.47	0
b3descarte	0.094838	102	43	25	125.91	1.74	21.52	0
b4descarte	0.13061	168	96	48	125.45	1.83	22.35	7
b5descarte	0	111	55	27	111.81	1.6	15.51	0
b6descarte	0.0053056	107	46	18	113.73	1.84	18.24	0
b7descarte	0	105	49	28	97.14	1.57	11.59	0
b8descarte	0	121	56	27	125.55	1.81	22.13	1
b9descarte	0	119	55	27	133.27	1.59	22	0
b10descarte	0	125	61	33	120.44	1.96	22.27	0
b11descarte	0	108	52	32	99.39	1.54	11.88	0
b12descarte	0	91	45	26	110.03	1.6	15.06	0

En el parámetro de porcentaje resulta cero pues es necesario que se analice el grano cortado por la mitad y así poder calcular el porcentaje de área hueca en el grano.

Los datos en color rojo, son aquellos que sobrepasan el valor límite especificado como es: Ancho del grano (134) y área del grano (22.5) con sus respectivas escalas.

En los resultados anteriormente mostrados en los dos distintos grupos (buen estado y de descarte) clasificados por agricultores, se quiso obtener una clasificación según el color (RGB) que predominaba en el grano pero no se logró ver una gran diferencia entre los dos grupos. Deducimos entonces que en el análisis externo del grano de cacao, el color predominante no es indicativo del tipo de grano (buen estado o de descarte).

La característica que mejores resultados nos da es el tamaño del grano, en este caso se han escogido dos parámetros que son el ancho del grano (diámetro) y el área del grano. Como valores límite se usó el valor de 134 en el ancho del grano y el valor de 22.5 en el área del grano, cada uno con su respectiva escala.

Analizando el ancho de grano obtenemos una validez del 75% aproximadamente y en el área del grano obtenemos una validez del 83 % aproximadamente. Estos resultados nos ayudan a realizar un primer descarte para luego continuar con el análisis de características internas.

Los resultados a continuación muestran las pruebas basadas en las características internas del cacao, analizando los 30 primeros granos de cacao. Se muestra las pruebas realizadas al sistema de adquisición de imágenes, a los algoritmos desarrollados.

La salida del bloque difuso, figura 58, muestra una marcada diferencia entre los granos mal fermentados con el resto, asignándoles un valor de 0.12. En cuanto a los granos parcialmente fermentados oscilan entre los valores 0.4 a 0.57. Por otro lado, el mínimo valor que presentan los granos bien fermentados es de 0.73; asimismo, el máximo valor de estos bordea el valor de 0.9.

De esta manera se puede considerar a la salida del bloque difuso como un índice capaz de describir la calidad del cacao en tres niveles de fermentación.

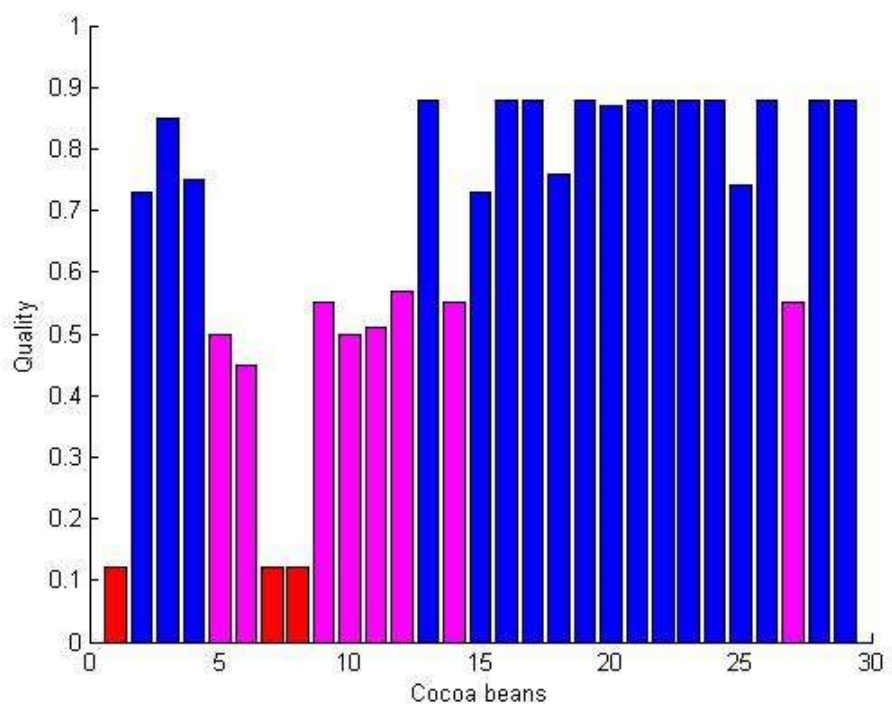


Figura 58. Respuesta del bloque difuso. Barras rojas: granos no fermentados. Barras púrpura: granos parcialmente fermentados. Barras azules: granos bien fermentados.

Fuente: Elaboración Propia.

Capítulo 3

Artículo Publicado

Cocoa bean quality assessment by using hyperspectral images and fuzzy logic techniques

*Juan Soto^{*a}, Guillermo Granda^a, Flavio Prieto^b, William Ipanaque^a, Jorge Machacuay^a*

^aUniversidad de Piura, Av. Ramón Mugica 131, Piura, PE; ^bUniversidad Nacional de Colombia, Av. Carrera 30 # 45, Bogotá, CO.

ABSTRACT

Nowadays, cocoa bean exportation from Piura-Peru is having a positive international market response due to their inherent high quality. Nevertheless, when using subjective techniques for quality assessment, such as the cut test, a wastefulness of grains is generated, additional to a restriction in the selection as well as improvement approaches in earlier stages for optimizing the quality. Thus, in an attempt to standardize the internal features analyzed by the cut test, for instance, crack formation and internal color changes during the fermentation, this research is submitted as an approach which aims to make use of hyperspectral images, with the purpose of having a quick and accurate analysis. Hyperspectral cube size was reduced by using Principal Component Analysis (PCA). The image generated by principal component PC1 provides enough information to clearly distinguish the internal cracks of the cocoa bean, since the zones where these cracks are, have a negative correlation with PC1. The features taken were processed through a fuzzy block, which is able to describe the cocoa bean quality. Three membership functions were defined in the output: unfermented, partly fermented and well fermented, by using trapezoidal-shaped and triangular-shaped functions. A total of twelve rules were propounded. Furthermore, the bisector method was chosen for the defuzzification. Begin the abstract two lines below author names and addresses.

Keywords: Cocoa beans, Food quality, PCA, Fuzzy Logic, Hyperspectral.

INTRODUCTION

Quality assurance of cocoa beans has attracted worldwide interest in order to control and prevent mixing of fermented with unfermented cocoa beans. Currently, this assessing of cocoa beans quality is done by a subjective method known as the cut test, which assesses the quality for only a small fraction of total grains. The cut test is widely used due to its simplicity, but it has some drawbacks such as color determination, since it is considered subjective and difficult to standardize. This method involves cutting a number of beans lengthwise and analyzing their internal color. A correctly fermented bean is characterized by its brown color and the presence of cracks, due to the cotyledon's death; a partly fermented bean by a certain violet hue, and an unfermented bean by its grey color and the lack of cracks. Likewise, an unfermented bean has a significant concentration of polyphenols, astringent taste and

unpleasant flavor, there is an inverse relationship between the flavor developed and the internal purple color. Furthermore, the appearance of cracks involves external physical changes such as the swelling of the bean. In conclusion, the cocoa bean quality is determined by the degree of fermentation.

Besides the cut test, there are other methods used for cocoa bean examination such as: sensory test and gas chromatography mass spectroscopy [1], high performance liquid chromatography, fermentation index [2]; even though they are reliable and accurate, they are often time consuming, expensive, destructive and require an elaborated sample preparation. [3]. Fourier transform near-infrared (FT-NIR) spectroscopy combined with Support Vector Machine (SVM) and synergy interval partial least square (Si-PLS) was attempted in this study for cocoa bean authentication [4].

Thus, in an attempt to standardize the internal features analyzed by the cut test, this research is submitted as an approach in which the use of hyperspectral images is proposed, with the aim of evaluating quality in an objective and quantitative way. The results involve a relation between these internal features and the grain's external shape, so that the developing of objective and non-invasive methods by machine vision can be carried out. These methods allow us to assess the quality without the need of cutting the beans as well as improving the contemporary fermentation and drying techniques.

METHODS AND MATERIALS

2.1 Fermentation and dried

Cocoa beans were collected by local farmers from Morropon and Piura, Perú, in October. In this season, the cocoa bean production is low. Later, the grains were taken to the local producers association for the fermentation and drying process.

For the fermentation procedure, the cocoa beans were placed into wooden boxes of 1m³ and covered by sacks and banana leaves, in order to conserve heat, for 6 days. Posterior to this, they were removed and sun dried for 5 days.

2.2 Image acquisition

The images were acquired by a hyperspectral camera, using halogen bulbs (Fig. 1). Each capture has 200x640 pixels and consists of 240 spectral channels, from 400 nm to 900 nm.

The cut test was performed by the experts of the local producers association. The grains analyzed were classified in three ways: unfermented, partly fermented and well fermented. Likewise, the producers are used to classify the grains depending on their size. The grains that belong to size 1 are bigger than the grains of size 2. This classification is mechanical, using a shale shaker.

Generally, this test is made before the drying process, in order to calculate the moment in which the cocoa beans have to be removed from fermentation. However, our aim is to analyze the final quality. For this reason, this test was carried out after sun dry.

The grains analyzed were labeled according to the degree of fermentation and size which they belonged to. A total of 30 samples were examined. Additionally, three captures of each grain placed in different positions were acquired.



Figure 1. Resonon hyperspectral system used for cocoa beans analysis.

2.3 Data analysis

The images were acquired by a hyperspectral camera, using halogen bulbs (Fig. 1). Each capture has 200x640 pixels and consists of 240 spectral channels, from 400 nm to 900 nm.

PCA

The hyperspectral cubes were captured by the Spectronon software and then exported and processed on Matlab.

Three channels were extracted from the hyperspectral cube to form an image (Fig. 2A y Fig. 2E) with the aim of segmenting the grain's area. The size of the smallest rectangle (Heightrectangle & Widthrectangle) that contains this area was used to form a new hyperspectral sub-cube, which contains only the information about the cocoa bean analyzed.

After this, each channel of the new hyperspectral sub-cube was resized to a column vector with length Heightrectangle x Widthrectangle. The 240 vectors were submitted for a principal component analysis (PCA), in which the first component (PC1) provides about the 95% of the information; the second component (PC2) just provides the 2% and the third component (PC3) the 0.1%.

The new image based on PC1 (Fig. 2B y Fig. 2F) provides enough information to clearly distinguish the internal cracks of the cocoa beans, since the zones, where these cracks are, have a negative correlation with PC1.

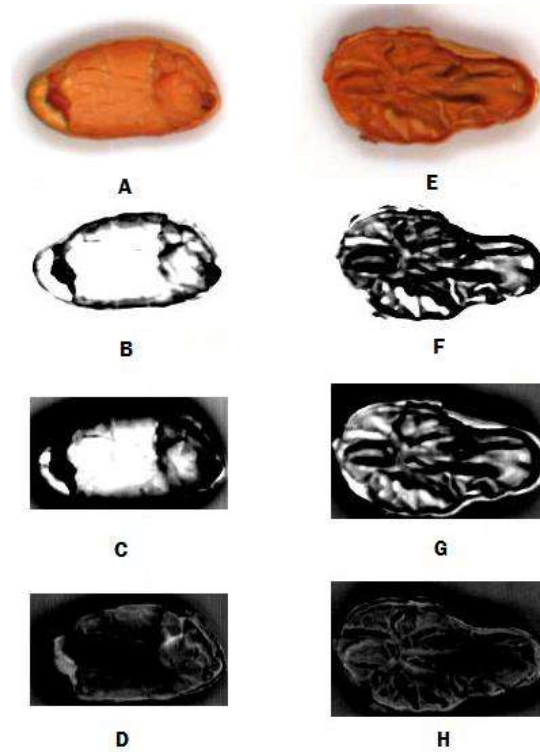


Figure 2. (A) Bean unfermented's Image NIR. (B) Bean unfermented's PC1. (C) Bean unfermented's PC2. (D) Bean unfermented's PC3. (E) Bean well fermented's Image NIR. (F) Bean well fermented's PC1. (G) Bean well fermented's PC2. (H) Bean well fermented's PC3. Image NIR: channel R set to 860nm, channel G set to 650nm and channel B set to 550nm.

In order to achieve the index related to the internal cracks, the grain's central zone was segmented (here is where the cotyledon is), after the area of this zone was calculated as well as the relation with the grain's total area.

To achieve the index related to the internal color, the uncracked zones were segmented. Subsequently, their average spectrum was determined. A new PCA analysis was carried out separately to study these spectrums. The results showed that the channels from the 123th to the 189th, exhibited the best values to distinguish the three degrees of fermentation. In particular, the channel 154th was the best. Hence, this channels range has allowed to form an index, which is able to quantify the grain's internal color.

Finally, two indexes are available, which are the next stage's inputs. These indexes are: the first index, which is able to represent the relation between the central cracked area and the total area; and the second index, which is able to represent the internal color.

Fuzzy Logic

The index related to cracks was described in the range [0,1] by four membership functions: Low, Medium, High and Higher; each function was trapezoidal-shaped. The index related to internal color was described in the same range and shape. Three membership functions were established: Low, Medium and High.

The output, which is related to grain quality, was described in the same range. Three membership functions were defined: unfermented, partly fermented and well fermented. The "partly fermented" function was triangular-shaped. The others were trapezoidal-shaped.

In addition, a total of 12 rules were propounded (Table 1). Furthermore, the bisector method was chosen for the defuzzification.

Table 1. Fuzzy Rules

	Index 1	Index 2	Quality
1	LOW	LOW	UNFERMENTED
2	LOW	MEDIUN	PARTLY
3	LOW	HIGH	UNFERMENTED
4	MEDIUM	LOW	FERMENTED
5	MEDIUM	MEDIUM	PARTLY
6	MEDIUM	HIGH	UNFERMENTED
7	HIGH	LOW	FERMENTED
8	HIGH	MEDIUM	PARTLY
9	HIGH	HIGH	UNFERMENTED
10	HIGHER	LOW	FERMENTED
11	HIGHER	MEDIUM	FERMENTED
12	HIGHER	HIGH	UNFERMENTED

RESULT

The platform MATLAB evaluates the quality of the cocoa bean, the classifier is based on fuzzy logic. A random sample of 200 cocoa beans are picked. Shows a cocoa bean of poor quality, this is a grain unfermented (Fig. 3). Shows a partly fermented grain (Fig. 4). Shows a well fermented grain (Fig.5).

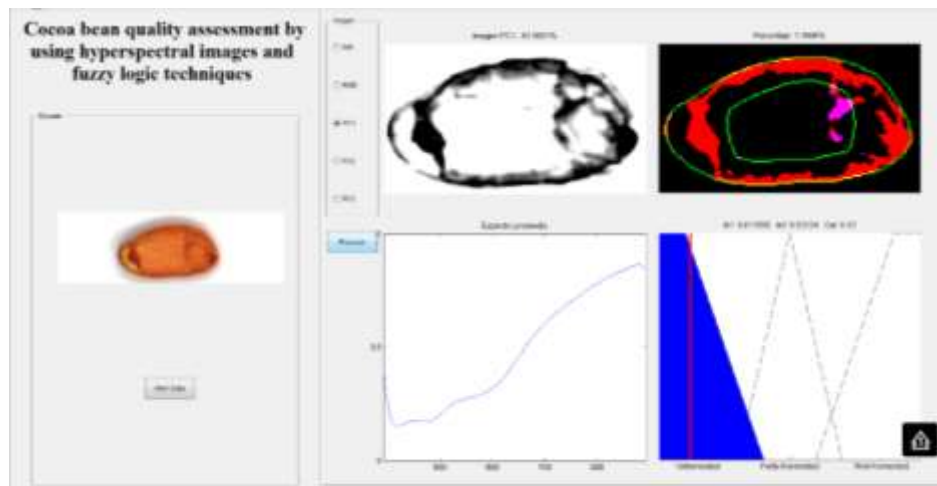


Figure 3. Quality assessed by fuzzy blocks - unfermented beans.

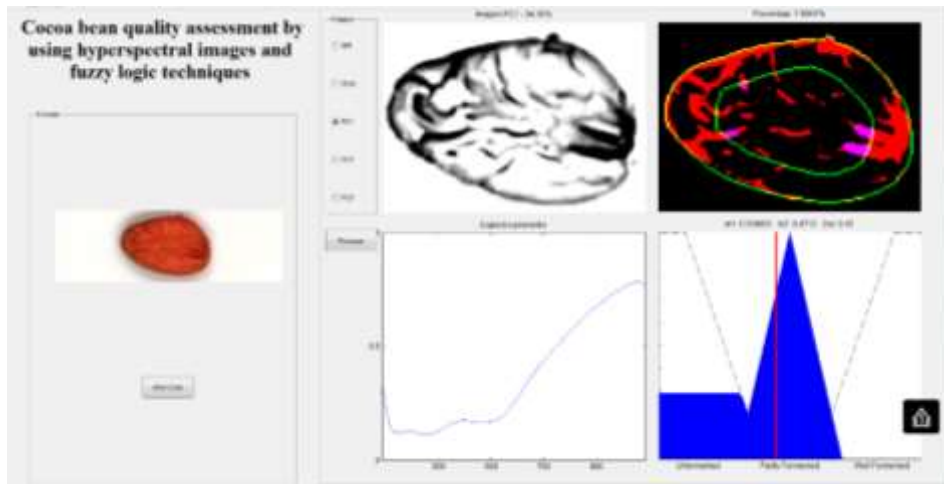


Figure 4. Quality assessed by fuzzy blocks – partly fermented beans.

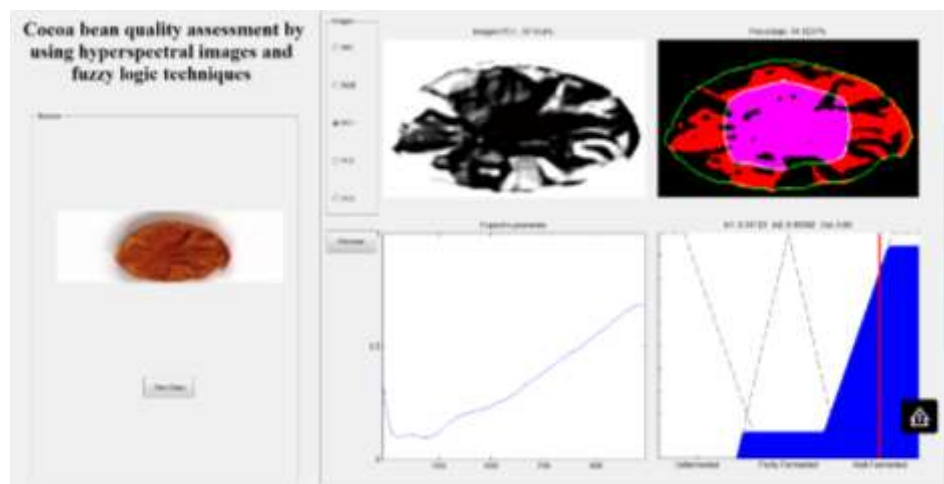


Figure 5. Quality assessed by fuzzy blocks – well fermented beans.

The Fuzzy block's output quantified the cocoa bean quality in the form of an index from 0 to 1 (values near to 0 stand for unfermented beans). The Fig. 6 shows these indexes for each grain analyzed.

There is a large gap between the quality classification thresholds. The unfermented cocoa beans presented a value of 0.12. As for partly fermented grains, the values ranged from 0.4 to 0.57. The values related to well fermented cocoa beans, fluctuate between 0.73 to 0.9. Therefore, the fuzzy block's output can be considered as an index which is able to classify the cocoa beans quality in three levels: unfermented, partly fermented and well fermented.

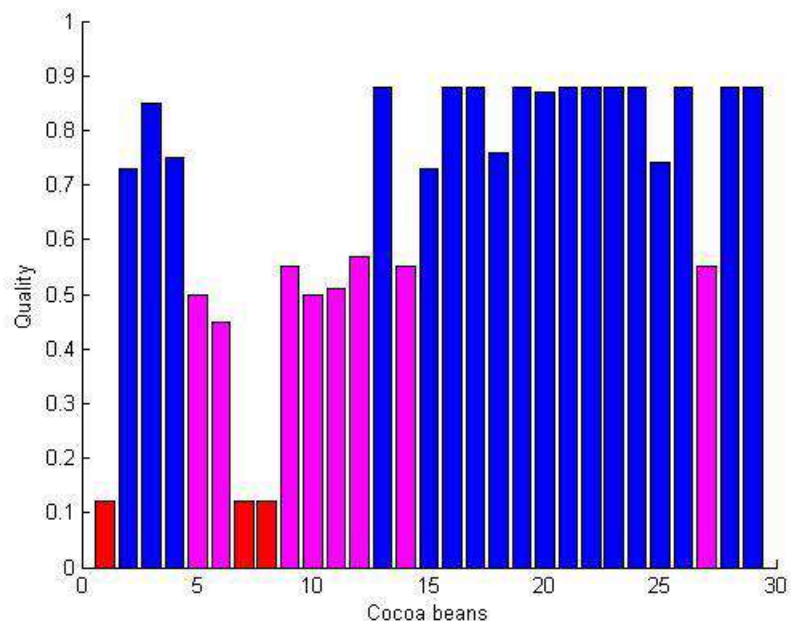


Figure 6. Quality assessed by fuzzy block's output. More gray bars: unfermented beans. Less gray bars: partly fermented beans. Black bars: well fermented beans.

CONCLUSIONS

The principal component PC1 obtained has a negative or close to zero correlation with the internal cracked zones. Additionally, this correlation is directly proportional with how deep the cracks are. The Fig.2B y Fig.2F, represent these zones with a black color.

The wavelengths range, which has the best way to distinguish the three degree of quality, is delimited from 640 to 785nm (these are the channels from 123 to 189 of the hyperspectral cube). This wavelengths range includes the red and near infrared spectrum.

Finally, the fuzzy block is able to classify the cocoa beans quality in three degrees.

The final goal of these researches is to establish a relation between the internal features and the external physical changes experienced by the grains over the fermentation and drying process. This relation would allow to assess the cocoa bean quality in a non-invasive and accurate way.

REFERENCES

- [1] Aculey, P., Snitkjaer, P., Owusu, M., Bassompierre, M., Takrama, J., & Norgaard, L. "Ghanaian cocoa bean fermentation characterized by spectroscopic and chromatographic methods and chemometrics," *Journal of Food Science*, 75(6), 300-307 (2010).
- [2] Romero Cortes, T., Salgado Cervantes, M., García Alamilla, P., García Alvarado, M., & Hidalgo Morales, M. "Relationship between fermentation index and other biochemical changes evaluated during the fermentation of Mexican cocoa (*Theobroma cacao*) beans," *Journal of the Science of Food and Agriculture*, 93(10), 2596-2604 (2013).
- [3] Saltini, R., Akkerman, R., & Frosh, S. "Optimizing chocolate production through traceability: A review of the influence of farming practices on cocoa bean quality," *Food Control*, 29(1), 167-187 (2013).

[4] Teye, R., Huang, X., Lei, W., & Dai, H. (2014). "Feasibility study on the use of Fourier transform near-infrared spectroscopy together with chemometrics to discriminate and quantify adulteration in cocoa beans," *Food Research International* 55, 288-293 (2014).

Conclusiones

La componente principal PC1 extraída tiene una correlación con las zonas agrietadas del grano. Esta correlación tiene una proporción directa con la profundidad de las grietas internas.

Se presenta el uso de un clasificador basado en lógica difusa cuyas dos entradas y salida son: el porcentaje de área de grietas, el color y la calidad respectivamente.

El índice relacionado con las grietas del grano fue descrito en el rango de [0,1] en 4 funciones de membresía: bajo, medio, alto y elevado; cada una del tipo trapezoidal.

El índice relacionado con el color interno fue descrito en el mismo rango y tipo. Se consideraron 3 funciones de membresía: bajo, medio y alto.

La salida relacionada con la calidad del grano, fue descrita en el rango de [0,1]. Se estimaron 3 funciones de membresía: malo, parcial y bueno. La función “parcial” fue descrita del tipo triangular. Las otras dos funciones del tipo trapezoidal. Un total de 12 reglas fueron planteadas, el método de la bisectriz fue considerado para la defusificación.

El análisis de imágenes RGB externas no es indicador para caracterizar el estado de la calidad del cacao.

El rango de longitudes de onda que mejor describe la diferencia entre las 3 calidades de granos se encuentra entre los 648 a 785 nm, que corresponden a las dimensiones 123 a 189 del cubo hiperespectral. Este rango de longitudes de onda abarca las longitudes del rojo visible e infrarrojo cercano.

La salida del bloque difuso, muestra una marcada diferencia entre los granos mal fermentados respecto a los granos parcialmente fermentados y bien fermentados, asignándoles un valor de 0.12. En cuanto a los granos parcialmente fermentados, oscilan entre los valores 0.4 a 0.57. Por otro lado, el mínimo valor que presentan los granos bien fermentados es de 0.73; asimismo, el máximo valor de éstos bordea el valor de 0.9.

De esta manera se puede considerar a la salida del bloque difuso como un índice capaz de describir la calidad del cacao en tres niveles de fermentación.

Referencias

- [1] F. Ayuso y V. Ochoa, «Aplicación de la Tecnología Grid al Tratamiento de Imágenes Hiperespectrales,» Universidad de Extremadura, España, 2008.
- [2] . V. Maider y J. M. Amigo, «Pre-processing of hyperspectral images. Essential steps before image analysis,» *Chemometrics and Intelligent Laboratory Systems*, vol. 117, pp. 138-148, 2012.
- [3] Aculey PC, Snitkjaer P, Owusu M, Bassompierre M, Takrama J, Nørgaard L, Petersen MA y Nielsen DS, «Ghanaian cocoa bean fermentation characterized by spectroscopic and chromatographic methods and chemometrics,» *Journal of Food of Science*, vol. 75, pp. 300-307, 2010.
- [4] T. Romero-Cortes, M. A. Salgado-Cervantes, P. García-Alamilla, M. A. García-Alvarado, G. Rodriguez-Jimenes y M. Hidalgo-Morales, «Relationship between fermentation index and other biochemical changes evaluated during the fermentation of Mexican cocoa (Theobroma cacao) beans,» *Journal of the Science of Food and Agriculture*, vol. 93, pp. 2593-2604, 2013.
- [5] A. Diwan P. , Renfu Lu y . D. E. Guyer, «Near-infrared hyperspectral reflectance imaging for detection of bruises on pickling cucumbers,» *Computers and Electronics in Agriculture*, vol. 53, p. 60–70, 2006.
- [6] A. M. Lefcourt, M. S. Wiederoder, N. (. Liu, M. S. Kim y Y. M. Lo, «Development of a portable hyperspectral imaging system for monitoring the efficacy of sanitation procedures in food processing facilities,» *Journal of Food Engineering*, vol. 117, pp. 59-66, 117.
- [7] E. Teye, X. Huang, H. Dai y Q. Chen, «Rapid differentiation of Ghana cocoa beans by FT-NIR spectroscopy coupled with multivariate classification,» *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 114, p. 183–189, 2013.
- [8] J. Gomez, Detection automática de podredumbres en cítricos mediante procesado avanzado de imágenes hiperespectrales., España: Universdad de Valencia, 2010.
- [9] Di Wu y D. Sun, «Novel non-invasive distribution measurement of texture profile analysis (TPA) in salmon fillet by using visible and near infrared hyperspectral Imaging,» *Food Chemistry*, vol. 145, p. 417–426, 2014.
- [10] Wayan Astika, Mohamad Solahudin, Andri Kurniawan y Yunindri Wulandari, «Determination of Cocoa Bean Quality with Image Processing,» de *AFITA 2010 International Conference, The Quality Information for Competitive Agricultural Based Production System and Commerce*, Indonesia, 2010.
- [11] A. Girolami, F. Napolitano, D. Faraone y A. Braghieri, «Measurement of meat

- color using a computer vision system,» *Meat Science*, vol. 93, p. 111–118, 2013.
- [12] Z. L. Sandoval Niño¹ y F. A. Prieto Ortiz, «Caracterización de café cereza empleando técnicas de visión artificial,» *Revista Facultad Nacional de Agronomía*, vol. 60, pp. 4105-4127, 2007.
- [13] J. Gómez Sanchis, G. Camps Valls y J. Blasco Ivars, Detección automática de podredumbres en cítricos mediante procesamiento avanzado de imágenes hiperespectrales, España: Universitat de València. Departament d'Enginyeria Electrònica, 2010.
- [14] Yao-Ze Feng, Gamal ElMasry, D. Sun, A. G. Scannell, Des Walsh y Noha Morcy, «Near-infrared hyperspectral imaging and partial least squares regression for rapid and reagentless determination of Enterobacteriaceae on chicken fillets,» *Food Chemistry*, vol. 138, p. 1829–1836, 2013.
- [15] Dan Liu, D. Sun, Jiahuan Qu, Xin-An Zeng, Hongbin Pu y Ji Ma, «Feasibility of using hyperspectral imaging to predict moisture content of porcine meat during salting process,» *Food Chemistry*, vol. 152, pp. 197-204, 2014.
- [16] C. C. Meulemans, Untung Surapati T y Ardin Tjatjo, «Colorimetric Measurements of Cocoa Beans,» *Indonesian Journal of Agricultural Science*, vol. 3, pp. 52-57, 2002.
- [17] L. R. Lleo, A. Herrero-Langreo, J. Roger, B. Diezma-Iglesia y P. Barreiro, «Comparison of multispectral indexes extracted from hyperspectral images for the assessment of fruit ripening,» *Journal of Food Engineering*, vol. 104, pp. 612-620, 2011.
- [18] Kıvanç Kılıç, İsmail Hakki Boyacı, Hamit Köksel y İsmail Küsmenoğlu, «A classification system for beans using computer vision system and artificial neural networks,» *Journal of Food Engineering*, vol. 78, p. 897–904, 2007.
- [19] C. Garrido-Novell, D. Pérez-Marin, J. M. Amigo, J. Fernández-Novales, J. E. Guerrero y A. Garrido-Varo, «Grading and color evolution of apples using RGB and hyperspectral imaging vision cameras,» *Journal of Food Engineering*, vol. 113, pp. 281 -288, 2012.
- [20] W. Jonfia-Essien, G. West, P. Alderson y G. Tucker, «Phenolic content and antioxidant capacity of hybrid variety cocoa beans,» *Food Chemistry*, vol. 106, pp. 967-975, 2008.
- [21] R. Saltini, R. Akkerman y S. Frosch, «Optimizing chocolate production through traceability: A review of the influence of farming practices on cocoa bean quality,» *Food Control*, vol. 29, p. 167–187, 2013.
- [22] Ernest Teye, Xing-yi Huang, Wu Lei y Huang Dai, «Feasibility study on the use of Fourier transform near-infrared spectroscopy together with chemometrics to discriminate and quantify adulteration in cocoa beans,» *Food Research International*, vol. 55, p. 288–293, 2014.
- [23] J. M. Rubio Amigo, I. Marti and A. Gowen, "Hyperspectral imaging and chemometrics : a perfect combination for the analysis of food structure, composition and quality," *Chemometrics in Food Chemistry*, vol. 28, pp. 343-370, 2013.
- [24] A. Holten Sivertsen, K. Heia, K. Hindberg y F. Godtliebse, «[6] Sivertsen, A.H.,Heia, Automatic nematode detection in cod fillets (*Gadus morhua* L.) by hyperspectral imaging,» *Journal of Food Engineering*, vol. 111, p. 675–681., 2012.
- [25] P. A. Coelho, M. E. Soto, S. N. Torres, D. G. Sbarbaro y J. E. Pezoa, «[10]

- CoelhoHyperspectral transmittance imaging of the shell-free cooked clam *Mulinia edulis* for parasite detection,» *Journal of Food Engineering*, vol. 117, p. 408–416, 2013.
- [26] E. Ivorra, J. Girón, A. J. Sánchez, S. Verdú, J. M. Barat y R. Grau, «Detection of expired vacuum-packed smoked salmon based on PLS-DA method using hyperspectral images,» *Journal of Food Engineering*, vol. 117, pp. 342 -349., 2013.
- [27] M. Arlorio, M. Locatelli, F. Travaglia, J.-D. Coisson, E. Del Grosso, A. Minassi, G. Appendino y A. Martelli, «Roasting impact on the contents of clovamide (N-caffeoyl-L-DOPA) and the antioxidant activity of cocoa beans (*Theobroma cacao* L.),» *Food Chemistry*, vol. 106, p. 967 – 975, 2008.
- [28] L. Kothe, B. F. Zimmermann y R. Galensa, «Temperature influences epimerization and composition of flavanol monomers, dimers and trimers during cocoa bean roasting,» *Food Chemistry*, vol. 141, p. 3656 – 3663., 2013.
- [29] S. Serranti, D. Cesare, F. Marini and G. Bonifazi, "Classification of oat and groat kernels using NIR hyperspectral imaging," *ElsevierB.V. Talanta*, vol. 103, pp. 276-284, 2013.
- [30] Gamal ElMasry, S. Cubero, E. Moltó y J. Blasco, «In-line sorting of irregular potatoes by using automated computer-based machine vision system,» *Journal of Food Engineering*, vol. 112, p. 60–68, 2012.