

PROCESAMIENTO DE IMÁGENES INDUSTRIALES: UNA APLICACIÓN AL CONTROL DEL TOSTADO DEL MANÍ

INDUSTRIAL IMAGE PROCESSING: AN APPLICATION TO CONTROL THE ROAST OF THE PEANUT

Juan José Palma

Universidad Tecnológica Nacional, Regional Córdoba
Marcelo López esquina Cruz Roja Argentina Ciudad Universitaria. Córdoba.
juanjosepalma2004@yahoo.com.ar

Silvia María Ojeda

Facultad de Matemática, Astronomía y Física. Universidad Nacional de Córdoba
Haya de la Torre y M Allende s/n. Ciudad Universitaria. Córdoba. Argentina.
ojeda@mate.uncor.edu

Mario Roberto Modesti

Universidad Tecnológica Nacional, Regional Córdoba
Marcelo López esquina Cruz Roja Argentina Ciudad Universitaria. Córdoba.
mmodesti@scdt.frc.utn.edu.ar

RESUMEN: Este trabajo se enmarca en el ámbito del control automático aplicado al sector agroindustrial. Responde a la creciente incorporación de tecnología en el sector primario por parte de algunas empresas dedicadas a la elaboración de derivados de la pasta de maní, en su búsqueda de un aumento en la productividad y calidad de sus productos. Concretamente, se presenta una metodología automática y en línea, para distinguir distintos niveles de tostado de maní pelado sin piel a granel e introducir, en caso de fallas, correcciones en el proceso de tostado. El método propuesto utiliza la información proporcionada por sensores ópticos instalados a la salida del horno de tostado y está basado en el modelo de regresión logística.

Palabras clave: Maní a Granel. Regresión Logística. Procesamiento de Imágenes. Redes de Computadoras. Horno de Tostado de Maní.

ABSTRACT: This work deals with automatic control applied to an agro-industrial sector. It responds to the increasing interest of the peanut industry to improve productivity and product quality. Precisely, an online automatic methodology to distinguish the different roast levels in the peanut roasting process is introducing. This helps to detect failures in the peanut roasting process so that corrections can be applied if needed. The proposed method is based on logistic regression with information provided by optical sensors installed near to the oven where the roasting process is carried out.

Keyword: Bulk Peanut. Logistic Regression. Image Processing. Computer Networks. Toasting Oven of Peanut.

1 INTRODUCCIÓN

La globalización de la economía hace imprescindible un esfuerzo por parte de las empresas de ofrecer productos de alta calidad, plenamente garantizados y con menor costo de producción. Por otro lado, automatizar las empresas implica reducir sus costos de producción, eliminando el trabajo monótono de un gran número de personas, ya que una cadena de producción automática requiere para las diversas operaciones de control, mantenimiento, etc., la participación de un número de personas bastante menor del que sería necesario en el caso de una producción análoga organizada con métodos tradicionales. En paralelo con el desarrollo de cadenas de producción, la aceptación mundial de redes de computación en diferentes entornos ha generado el deseo de expandir su aplicación al piso de planta. Las empresas de hoy en día están cada vez más interesadas en conectar sus sistemas de control a redes de computación industriales. Además desean más y mejores diagnósticos, menor tiempo de parada y costo reducido de instalación y mantenimiento, impulsando a numerosos usuarios a invertir en soluciones con conexiones en red (NAGHSHVAR; ZHUANG; JAVIDI, 2009).

Los recientes avances en el área de procesamiento de imágenes y sistemas de computación, permiten implementar aplicaciones críticas de inspección en conexión con las redes de automatización industrial existentes. Con el aumento de la funcionalidad, aumenta el número de dispositivos instalados y software de control, llevando esto a un mayor tráfico de datos en la red. Las redes proveen un medio adecuado para recuperar diagnósticos de estos dispositivos y mantenimiento de programas. Por ejemplo, la información específica a un instrumento de un desperfecto, como ser la detección de un nivel bajo de temperatura de un horno de tostado de granos, puede ser comunicada sobre la red al sistema de control durante su operación. La red entrega el diagnóstico a la interfase del operador del sistema, alertando al personal de planta del problema. La temperatura puede ser regulada en el momento adecuado, antes de que surja un inconveniente en el proceso. Así, localizar fallas en un dispositivo, leer sus códigos de falla, actualizar registro de datos, todo sin afectar el intercambio de datos de control de entrada/salida remotas entre sí misma y otros nodos, es mucho más que una necesidad.

Se sabe que la vista proporciona el 75 % de la información que recibe el ser humano (BEGBIE, 1977). Consecuentemente, dotar a las máquinas de esta fuente de información (visión artificial), proporciona una mejora significativa en el proceso de control automático de medición del color en la industria alimenticia (NAGHSHVAR; ZHUANG; JAVIDI, 2009). La importancia del control colorimétrico en alimentos sienta sus bases esencialmente en tres razones. La primera es la caracterización de la calidad de un producto a partir del color. Un consumidor espera que todas las unidades de una cierta marca de alimento tengan el mismo color. Si una es diferente,

inmediatamente es sospechosa y probablemente será retirada de las estanterías. La segunda es el uso del color como índice de valor económico. Por ejemplo, la madurez óptima de los tomates está asociada con el desarrollo óptimo del color y del sabor. Puesto que es más fácil medir el color que el sabor, se utiliza el color como medida del valor del lote. La tercera razón es mejorar un producto dado. Esto se aplica principalmente a alimentos preparados o a ingredientes de alimentos en los que el tecnólogo tiene libertad para manipular el color, como en el caso del proceso de tostado de cereales.

Los hornos de tostado de cereales generalmente han sido diseñados para el tueste de frutos secos o similares, tales como maní, semillas de girasol, semillas de calabaza, café, almendras, avellanas, pistachos, etc. En particular, el proceso del tostado del maní debe controlarse cuidadosamente (BUCKHOLZ; DAUN; STIER, 1980). Por un lado, un tueste pasado o muy oscuro puede producir una disminución de la calidad de las proteínas que contiene el fruto; mientras que por otro lado, un tostado adecuado intensifica el aroma y el sabor del maní y sus derivados, mejorando su digestibilidad.

En este trabajo se presenta una propuesta para el desarrollo de un sistema automático de control del proceso de tostado de maní. Con esto se pretende modernizar y automatizar procesos ya existentes logrando de esta forma mejorar la calidad del producto. Concretamente, se presenta una metodología automática y en línea, para distinguir distintos niveles de tostado de maní pelado, sin piel, a granel e introducir correcciones en el proceso de tostado, según características de calidad previamente definidas. El método propuesto utiliza la información proporcionada por una cámara IP instalada a la salida del horno de tostado y analiza en red los datos obtenidos.

La información lograda a partir de estas imágenes se aplica a un modelo de regresión logística para ajustar la temperatura a la cual debe trabajar el horno. De esta forma se procura aplicar tecnologías informáticas y de procesamiento estadístico de imágenes a las actividades de control de producción, para minimizar la intervención humana, logrando que la temperatura del horno permanezca fija o cercana a un valor deseado a lo largo del tiempo, por medio de una acción correctiva constante. Al independizar el funcionamiento del proceso de las habilidades del operador se puede asegurar homogeneidad del tono de tostado del maní en el tiempo, y repetitividad del color obtenido en el tostado de otros lotes, garantizando características de calidad de exportación exigidas para diferentes productos finales de pasta de maní obtenidos en un proceso posterior a la molienda.

El trabajo se organiza de la siguiente manera: en la sección 2 se comenta el mercado global y argentino del maní, además de algunos productos comerciales elaborados con pasta de

maní. En la sección 3 se expone brevemente la propuesta de control de tostado del horno, la cual es complementada en las conclusiones con un diagrama en bloque de la arquitectura final de la aplicación. En el apartado 4 explica el perfil de la cabina de iluminación que se construyó en el laboratorio para este estudio. En la parte 6 se efectúa un análisis de los estadísticos que tienen mayor influencia en la discriminación de las imágenes de maní a granel con diferentes niveles de tostado y se formula la ecuación que discrimina el nivel de tostado deseado del no deseado.

2 EL MERCADO DEL MANÍ

El maní es una planta anual de la familia de las leguminosas, que produce frutos secos comestibles. Es originario de América del Sur y la planta presenta aspecto de matorral, hojas alternas, compuestas por dos pares de folíolos ovalados, con flores axilares, papilionáceas, de color amarillo estriado de rojo. El fruto de esta leguminosa, que madura bajo tierra, a una profundidad de 5 a 10 centímetros, es indehiscente, alargado, con tendencia a estrangularse, redondeado en ambos extremos, rugoso en el exterior y de color pajizo o terroso; contiene de una a cuatro semillas ovoideas, formadas por dos cotiledones hemisféricos y revestidos por una película rojiza. Estas semillas contienen en promedio, entre 40 y 50% de aceite comestible, de 25 a 27% de proteínas, de 4 a 12% de agua, alrededor de 2 a 3,1% de celulosa y de 1 a 3% de cenizas (BATAL; DALE; CAFÉ, 2005; SANDERS, 2007).

El mercado mundial de maní es reducido, concentrado en cinco países, a los que se atribuye el 79% de la producción mundial. Estos países son China, India, Nigeria, EEUU e Indonesia. Los precios internacionales dependen de la producción de los grandes países productores y exportadores, la variedad y el tamaño del grano. El 80% del comercio del maní es descascarado, ya sea como grano o bien transformado (pasta de maní y aceite de maní) (MAGFOR, 2010). Argentina ocupa el decimocuarto lugar en la producción mundial, siendo la provincia de Córdoba la principal zona productora (FERNÁNDEZ; GIAYETTO, 2006).

En lo que a exportación se refiere, Argentina es el segundo país exportador, detrás de China cuyas exportaciones se concentran principalmente en los países asiáticos. La calidad del fruto se logra durante todas las etapas y operaciones del cultivo, cosecha y post-cosecha, con un solo propósito: preservar la integridad física y química, en primera instancia de la vaina y finalmente del grano. El objetivo final del acondicionamiento de post-cosecha es obtener granos de maní sanos, secos, limpios, libres de contaminantes (químicos o biológicos) y de excelente sabor.

La mantequilla o manteca de maní es una pasta elaborada con maníes molidos sin piel generalmente tostados, ya sea salada, o bien endulzada. Es consumida principalmente en Estados

Unidos, Australia, Reino Unido y Canadá y se la utiliza en la elaboración de distintas comidas o, según los gustos, se la consume sola sustituyendo a la manteca de leche. También se emplea para elaborar golosinas, alfajores, turrone, obleas, etc. (SAGPYA, 2010).

El maní tostado se ofrece en distintos calibres, cumpliendo las especificaciones del cliente. Los hornos comerciales actuales monitorean la temperatura en su interior a través de sensores de temperatura, obteniendo de este modo distintos niveles de tostado, regulando la temperatura del horno. El proceso del tostado del maní debe controlarse cuidadosamente (BUCKHOLZ; DAUN; STIER, 1980). Por un lado, un tueste pasado o muy oscuro hace que el grano suelte más aceites y es en esta etapa donde puede ocurrir una disminución de la calidad de las proteínas; por otro lado, el tostado intensifica el aroma y el sabor del maní, mediante la dextrinización del almidón, mejora la digestibilidad y contribuye a reducir la voluminosidad de los preparados alimenticios. También destruye los microorganismos e insectos y reduce la actividad enzimática, con lo cual favorece a su conservación.

3 PROPUESTA DE CONTROL DEL NIVEL DE TOSTADO DEL MANÍ

Se propone implementar un colorímetro basado en una cámara IP conectada en una red de computadoras. La cámara está enfocada hacia la cinta transportadora de Maní que se encuentra a la salida del horno y permite adquirir imágenes que luego de ser procesadas proporcionan información útil para ajustar la temperatura del horno alrededor de un valor deseado, cerrando de este modo el lazo de realimentación.

En la Figura 1, se muestran los componentes que forman parte del colorímetro: cámara, sistema de iluminación, la cinta transportadora, el PLC y la PC donde se instalará el programa que se debe diseñar para distinguir los distintos niveles de tostado de maní.

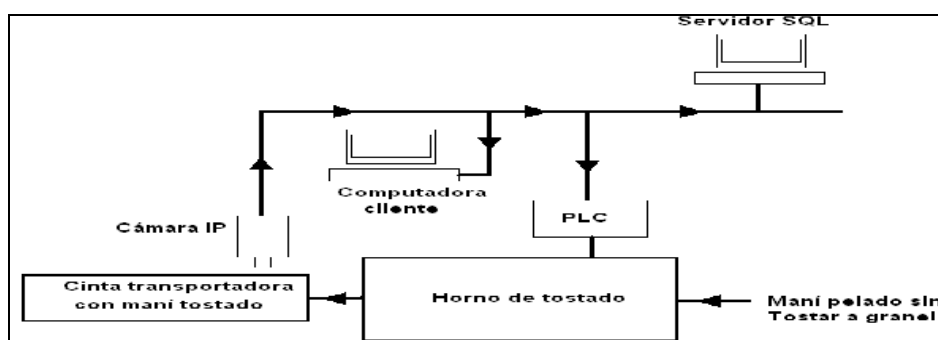


Figura 1 – Arquitectura de la aplicación

La arquitectura del sistema se plantea del tipo cliente-servidor y se espera implementarla a través de una red basada en protocolos TCP/IP (DECOTIGNIE, 2005). El cliente está constituido por un computador y una cámara IP. En la computadora cliente, se instalará el programa que es

capaz de diferenciar los distintos niveles de tostado, junto con los *Socket* (los que ya fueron implementados) para el establecimiento de una comunicación segura entre las computadoras cliente y servidor a través de Internet. Los datos procesados con la información del nivel de tostado son enviados al PLC que es el responsable de regular la temperatura del horno. El PLC ejecuta correlativamente las instrucciones indicadas en el programa de usuario almacenado en su memoria, generando las ordenes o señales de mando a partir de las señales de entrada que fueron enviadas por la computadora cliente; al detectarse cambios en éstas señales, el autómatas reacciona según el programa y ajusta su salida para obtener el nivel de tostado buscado. Esta secuencia se ejecuta continuamente para conseguir el control actualizado del proceso.

Para validar la propuesta fue necesario simular el proceso de tostado de maní. Ello se logró considerando bandejas de maní ya tostado, sin piel, a granel y en tres niveles de tostado diferente: poco tostado, tostado (nivel de tostado óptimo) y muy tostado. Las bandejas conteniendo el maní simulaban en este caso la cinta transportadora conteniendo el maní a la salida del horno de tostado. Las bandejas con maní fueron fotografiadas en el interior de una cabina de iluminación especialmente construida a tal efecto. De esta manera se lograron muestras de los tres niveles de tostado señalados, que fueron procesadas permitiendo definir una regla basada en el modelo de regresión logística para controlar en línea la temperatura del horno de tostado.

4 PLATAFORMA DE DESARROLLO

La cabina de iluminación para la obtención de las muestras, se construyó en cartón, con una altura de 745 mm, 490 mm de ancho, y 450 mm de profundidad, con paredes internas en color beige opaco. La fuente de iluminación se instaló en la parte superior, por encima de la bandeja conteniendo el maní. La distancia entre la bandeja y la fuente de luz se estableció en 600 mm. Sobre fotografías preliminares de bandejas con maní tostado, se verificó que en general, el promedio de la matriz componente de color rojo es mayor que el promedio sobre la matriz de color verde y ésta a su vez mayor que el promedio sobre la matriz azul; por otra parte, del estudio de histogramas de fotografías de maní a granel, se deduce mayor saturación de los componentes espectrales rojos y verdes que azules. De aquí que para iluminar la cabina de obtención de fotografías de maní a la salida del horno de tostado, se eligió una lámpara de tipo incandescente capaz de generar un espectro continuo (todo el espectro visible) de luz blanca, con una distribución espectral en que predomina el rojo y verde sobre el azul, a una temperatura de 2700° K con una emisión de luz de 800 lúmenes, eficiencia luminosa de 57 lúmenes/watt y potencia de 14 watt.

Para captar las imágenes del maní en las bandejas, simulando cintas se utilizó una cámara IP, con o transductor de tipo CMOS (IP INTELLINET, 2010). La cámara se ubicó al costado de la bandeja con maní, captando imágenes desde arriba. La distancia entre el centro de la lente y la muestra es de 170 mm. Se utilizó una lente de 1/3" con montura CS estándar, distancia focal de 6.0 mm., vista angular de 54°. La Figura 2, muestra en a) las bandejas de maní en los tres niveles de tostado y en b) la cabina de iluminación.

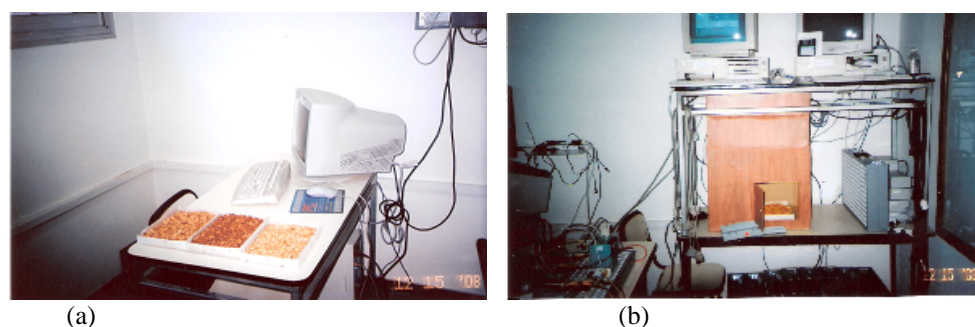


Figura 2 – a) Bandejas conteniendo maní en tres niveles de tostado. b) Cabina de iluminación construida en el laboratorio

Se obtuvieron en el laboratorio 3900 fotos (1300 para cada nivel de tostado), en formato *jpg*, modelo RGB, de dimensión 640 x 480. Se realizó un análisis descriptivo por grupo, tanto sobre las imágenes compuestas (composición de bandas de color) como considerando cada banda (verde, azul y roja) por separado. Para cada nivel de tostado, la muestra se dividió en dos grupos; una muestra de trabajo (muestra de estimación), compuesta por mil fotos a través de la cual se obtuvo información de diversos parámetros que caracterizan a la población de la cuál fue extraída, y otra de validación constituida por trescientas fotos, que fue utilizada para probar la capacidad predictiva del modelo de regresión logística finalmente propuesto.

5 ANÁLISIS DE DATOS

5.1 Análisis Preliminar de Imágenes de maní

Al observar las imágenes de maní pelado, sin piel, a granel, obtenidas en el laboratorio, se destaca de manera notable la presencia tanto de píxeles oscuros, como excesivamente claros; los primeros se deben a sombras en los espacios entre granos; en tanto que la presencia de los píxeles claros obedece a la existencia de pequeños fragmentos de maní que adquieren un tostado menor (Figura 3).

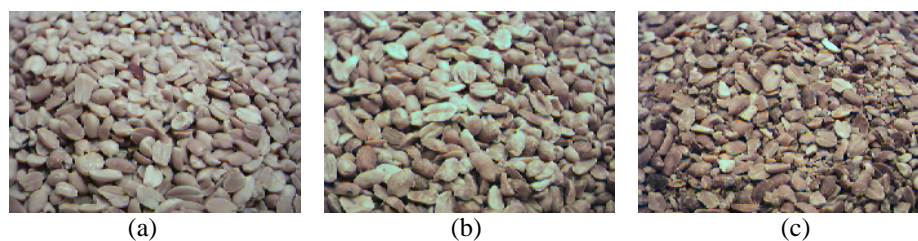


Figura 3 – Imágenes de bandejas de maní tostado en formato jpg, modelo RGB de 640x480 píxeles. a) Maní sin tostar (tostado deficiente). b) Maní tostado (nivel de tostado óptimo). c) Maní muy tostado

La intensidad de la imagen es una variable que puede asumir cualquier número entero entre 0 y 255. En el caso de la imagen compuesta, esta variable da cuenta del color en la imagen; en tanto que para el caso de una banda, indica el nivel de gris de la imagen en dicha banda. Tanto en imágenes de una banda como en imágenes compuestas, la intensidad promedio presenta valores altos para el maní que no fue tostado y más bajos a medida que el nivel de tostado aumenta. De ahí que los píxeles que en la imagen se presentan excesivamente oscuros se corresponden con valores inusualmente bajos de intensidad; en tanto que los píxeles marcadamente claros lo hacen con valores notablemente altos. Los valores muy bajos como muy altos de intensidad se denominan valores atípicos y no se relacionan con la intensidad característica del grano de maní tostado destinado al consumo. La Figura 4 muestra que la variable intensidad presenta valores de intensidad atípicos en los tres niveles de tostado estudiados (BARON, 2007; FIEGUTH, 2011).

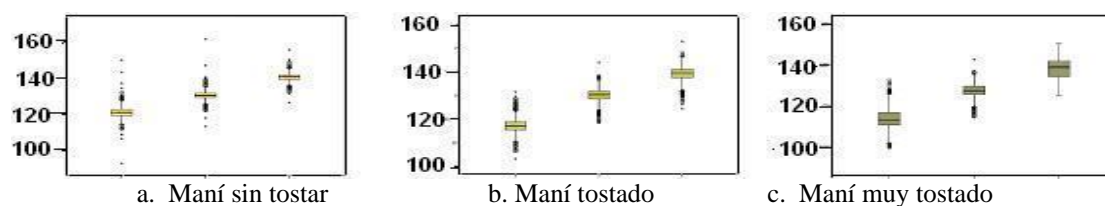


Figura 4 – *Box-plot* de la variable intensidad por banda de color, en los tres niveles de tostado. a) Maní sin tostar. b) Maní tostado. c) Maní muy tostado

En cuanto a las distintas bandas de color (roja, verde azul), quedó evidenciado que, en términos de nivel de gris de la imagen, éstas aportan, en general, un conocimiento diferente (observe las escalas en los gráficos de la Figura 5). Esta hipótesis se confirmó al aplicar la prueba no paramétrica de *Kruskal-Wallis* ($p < 0.000$).

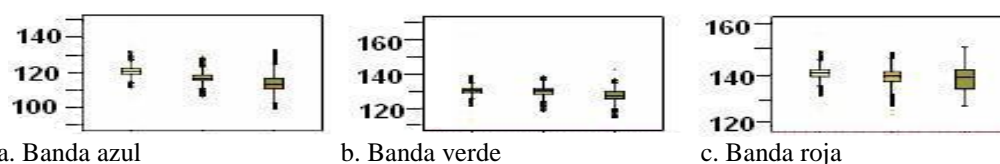


Figura 5 – *Box-plot* del nivel de gris promedio para los distintos niveles de tostado de maní (sin tostar, tostado, muy tostado respectivamente)

Para cada nivel de tostado se construyó un intervalo de confianza (95%) para la media de la variable intensidad, en base a 1000 imágenes de bandejas de maní pelado, sin piel, obtenidas

en el laboratorio. Intervalos correspondientes a dos niveles diferentes de tostado resultaron disjuntos; sin embargo no fue posible definir un criterio de clasificación de imágenes a partir de estos intervalos, debido a su exagerada precisión (longitud menor a uno); tal precisión para estimar la media de intensidad de cada grupo determina que en muchos casos el valor de intensidad promedio de una dada imagen no pertenezca a ninguno de los tres intervalos obtenidos. Estos resultados son válidos tanto en el análisis de la imagen compuesta como por bandas. La Tabla 1 muestra un análisis descriptivo de la intensidad para el caso de imágenes compuestas. Observe la longitud de los intervalos hallados.

Tabla 1 – Análisis descriptivo de la variable intensidad (imagen compuesta) por nivel de tostado

Intensidad	Nivel	Estadístico	
	Intensidad	Sin Tostar n=1000	Media Muestral
Intervalo de confianza (95%) para la intensidad promedio.			130,39 - 130,67
Mediana			130,57
Varianza			4,91
Desvío Estándar			2,21
Mínimo			116,72
Máximo			155,75
Tostado n=1000		Media Muestral	128,62
		Intervalo de confianza (95%) para la intensidad promedio	128,45 - 128,77
		Mediana	128,82
		Varianza	6,70
		Desvío Estándar	2,59
		Mínimo	117,65
		Máximo	137,22
Muy Tostado n=1000	Media Muestral	126,40	
	Intervalo de confianza (95%) para la intensidad promedio	126,22 - 126,57	
	Mediana	126,48	
	Varianza	7,99	
	Desvío Estándar	2,82	
	Mínimo	114,26	
		Máximo	138,32

5.2 Medidas Estadísticas para Imágenes de Maní

El análisis preliminar de los datos, evidenció limitaciones de la muestra para caracterizar el nivel de tostado óptimo. Por otra parte clasificar imágenes de maní según su nivel de tostado, considerando intervalos de confianza para el promedio de la intensidad, no fue viable. Consecuentemente se plantó la necesidad de estudiar otras medidas estadísticas que pudieran caracterizar en nivel de tostado óptimo de forma más adecuada que la media de intensidad. Se analizaron diferentes medidas:

5.3 La media Recortada

Como se vio en la sección anterior la intensidad promedio de una imagen de maní se ve afectada por la presencia de valores atípicos de intensidad debidos tanto a espacios entre granos de maní y como a brillos de fragmentos pequeños. Una manera de amortiguar tal distorsión, consiste reemplazar la intensidad media por la *intensidad media recortada*. Esta medida se obtiene ordenando de menor a mayor el nivel de intensidad de los píxeles de cada imagen, y calculando la media de los valores que quedan después de eliminar un cierto porcentaje de los valores del extremo inferior y/o superior. Por ejemplo, la media recortada al 15% consiste en 1°) eliminar de la imagen el 15% de las observaciones extremas en la intensidad de la imagen con la posibilidad de eliminar en forma equilibrada o no, un $\alpha\%$ de los valores más bajos y un $\beta\%$ de los valores más altos ($\alpha+\beta=15$) y 2°) promediar los valores que no fueron eliminados. Se estudió la media recortada al 2%, al 5%, al 10% y al 15%, calculando estas medidas a cada imagen de la muestra.

La distribución de los valores obtenidos muestra una notable disminución de la cantidad de datos atípicos, con respecto a la cantidad de datos atípicos observada en la distribución de la media de intensidad a secas (sin recortar), tanto para imágenes compuestas como para imágenes de una sola banda de color. La Figura 6 exhibe el perfil de la media recortada al 15%, con $\alpha=10$ y $\beta=5$, sobre la muestra de estimación (imágenes compuestas) para cada nivel de tostado. Al comparar esta figura con los gráficos de la Figura 5, se observa una notable disminución en la cantidad de datos atípicos (BARON, 2007; FIEGUTH, 2011).

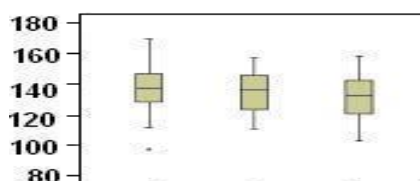


Figura 6 – Box-plot de la media recortada al 15% (10-5), para imágenes compuestas, según nivel de tostado

5.4 El Área

Esta medida surgió a partir del estudio del histograma de las imágenes de maní, el cual presenta en todas las bandas un perfil característico. Se trata de histogramas asimétricos con sesgo a izquierda, y tales que si se lo aproxima mediante una curva suave, ésta resulta cóncava hacia abajo en el primer tramo, (valores bajos de la variable intensidad) y luego cóncava hacia arriba, con crecimiento relativamente suave, hasta el valor de la moda. Para valores mayores a la moda, esta curva decrece abruptamente con concavidad hacia arriba (Figura 7). Se entiende que el comportamiento cóncavo de la curva del histograma en la región de valores bajos de intensidad se debe a los espacios de sombras entre los granos de maní, que no es relevante para informar sobre el nivel de gris de la imagen. También puede observarse, que para diferentes

niveles de tostado, en el tramo donde la curva crece, la concavidad difiere, según de qué nivel de tostado que se trate; así, la relación entre el área acumulada bajo la curva hasta el valor de la moda, A_1 , y el área total A_2 , depende en general del nivel de tostado (Figura 8). Esta observación motivó la definición de la siguiente medida que llamamos Área (Ecuación 1):

$$A = \frac{A_1}{A_2} \quad (1)$$

Operativamente A , se obtuvo como el cociente entre la suma de todas las intensidades menores o iguales a la intensidad modal y la cantidad total de píxeles en la imagen.

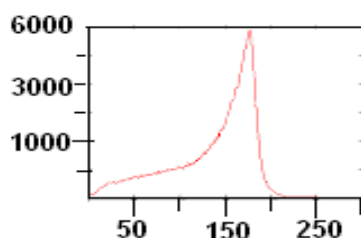


Figura 7 – Histograma de una imagen de maní tostado (imagen compuesta)

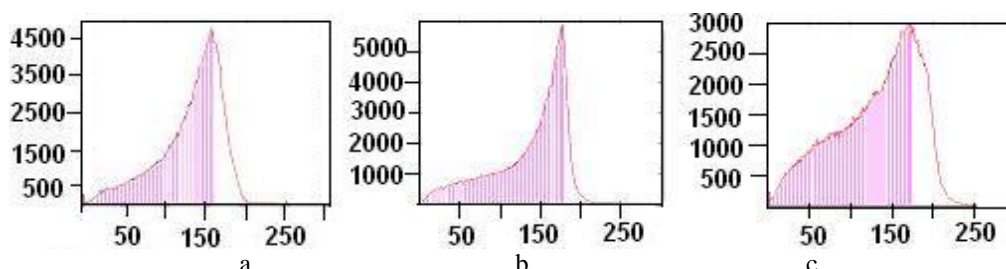


Figura 8 – Histogramas de tres imágenes en la banda roja. En todos los casos el sector sombreado, corresponde al área acumulada bajo la curva hasta el valor modal de la intensidad. a) Maní sin tostar. b) Maní tostado y c) Maní muy tostado

Otras medidas estadísticas examinadas, fueron la mediana de la intensidad en la imagen y la mediana de la intensidad de los datos previamente recortados (al 2%, 5% y 10%). Como medidas de dispersión, se consideraron el desvío estándar en la imagen de intensidad, y la desviación absoluta media (MAD). Esta medida, basada en la mediana está dada por (Ecuación 2):

$$MAD = \text{mediana} [\text{abs} (x_i - \text{mediana } x)] \quad (2)$$

donde $\text{abs} (y)$ calcula el valor absoluto de y .

Para todas las medidas consideradas, se realizó un estudio de normalidad. En todos los casos se rechazó la hipótesis de normalidad (test de *Kolmogorov-Smirnov* $p < 0.000$), lo que impidió la aplicación de metodologías clásicas tales como ANOVA y regresión lineal múltiple.

5.5 Análisis de Regresión Logística

La regresión logística es un método de análisis estadístico adecuado para modelizar una variable de respuesta binaria (HOSMER; LEMESHOW, 2000). Se trata de un modelo flexible en cuanto los supuestos que deben cumplir las variables del modelo, que permite expresar la probabilidad de que ocurra un evento, como función de ciertas variables independientes, llamadas regresoras, que se consideran potencialmente influyentes. De acuerdo con este enfoque, en este trabajo se propone incorporar esta metodología con el propósito de construir un modelo que permita establecer un criterio para clasificar imágenes de maní a la salida del horno de tostado.

Sea la variable dicotómica Y definida como (Ecuación 3):

$$Y(i) = \begin{cases} 1 & \text{si el nivel de tostado de la } i\text{-ésima imagen es } \textit{óptimo} \\ 0 & \text{en caso contrario} \end{cases} \quad (3)$$

Según el modelo de regresión logística, es posible expresar el cociente entre la probabilidad de que el nivel de tostado del grano sea el deseado ($P(Y=1)$) y la probabilidad del complemento de este evento ($P(Y=0)$), como una combinación lineal de n variables independientes, propuestas a priori. En la Tabla 2, se presentan las variables consideradas como posibles variables regresoras. La construcción del modelo se basó en 3000 fotografías de bandejas de maní (1000 para cada nivel de tostado). La validación y evaluación de la capacidad predictiva del modelo finalmente construido se realizó en base a una muestra de validación de 900 imágenes de bandejas de maní tostado (300 de cada nivel de tostado).

Tabla 2 – Variables potencialmente regresoras en el modelo

Notación	Variable
X_1	Media de Intensidad
X_2	Mediana de Intensidad
X_3	Desvío Estándar de Intensidad
X_4	Área
X_5	Media de Intensidad Recortada al 15% (10-5)
X_6	MAD

Un estudio de colinealidad entre las variables de la Tabla 2 determinó el conjunto de variables que finalmente se incorporaron como variables regresoras en el modelo: X_2 (Mediana), X_4 (Área) y X_6 (MAD). Los gráficos de la Figura 9, muestran para la banda roja, la correlación entre las variables de la Tabla 2 (a) y entre las tres variables finalmente seleccionadas (b). Para las bandas azul y verde este estudio arrojó resultados similares.

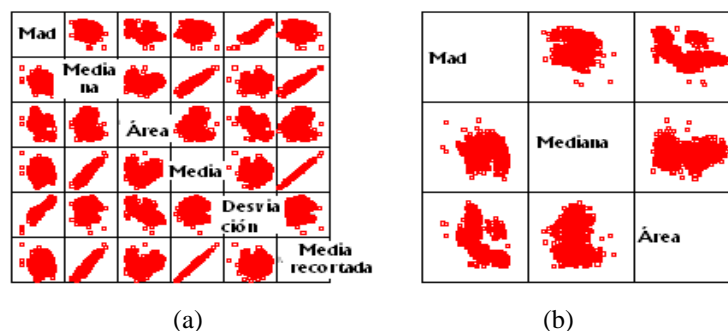


Figura 9 – Correlación entre las variables del modelo a) Variables propuestas inicialmente. b) Variables seleccionadas

El modelo logístico fue rechazado para las bandas azul y verde; no así para la banda roja. Los resultados en la banda roja se presentan en la Tabla 3. Se observa que en el modelo finalmente seleccionado resultó eliminada la Mediana. También se observa la estimación de máxima verosimilitud de los parámetros del modelo para las variables MAD y Área. Con un nivel de significación del 5%, los coeficientes de dichas variables son estadísticamente diferentes de cero (test de *Wald*). Ello indica que las mismas explican el cociente de riesgo del modelo y son capaces de predecir la probabilidad de que una imagen tomada a la salida del horno de tostado se corresponda con el nivel de tostado deseado del grano (HOSMER; LEMESHOW, 2000; MENARD, 1995; PREGIBON, 1981).

Tabla 3 – Variables en regresoras en el modelo de regresión logística seleccionado

Variabes	B	E.T	Wald	GI	Sig
Área	-101.064	4.389	530.369	1	0.000
Mad	-.754	0.034	504.197	1	0.000
Constante	39.208	1.702	530.614	1	0.000

La estimación de máxima verosimilitud de los parámetros del modelo permite expresar a la $P(Y=1/ X_4, X_6)$ como (Ecuación 4):

$$\hat{P}(Y=1/ X_4, X_6) = \frac{e^{39,208101,064* \text{Área} - 0,754* \text{MAD}}}{1 + e^{39,208101,064* \text{Área} - 0,754* \text{MAD}}} \quad (4)$$

La Ecuación 3, permite clasificar cada imagen captada a la salida del horno de tostado como aceptable o no aceptable, conociendo los valores de las variables X_4 (Área) y X_6 (MAD) para dicha imagen. Si el valor obtenido a partir de 3 es mayor que 0.5, la imagen será clasificada dentro del grupo de tostado deseado; por el contrario, si este valor es inferior o igual a 0.5, la imagen será clasificada en el grupo de imágenes con nivel de tostado no deseable. Los valores de las pruebas asociadas con el modelo hallado se presentan en las Tablas 4 y 5.

Tabla 4 – Prueba de Hosmer Lemeshow

Chi cuadrado	GI	Sig.
--------------	----	------

Chi cuadrado	Gl	Sig.
9,49	8	,303

Tabla 5 – Resumen del modelo

-2 log de la Verosimilitud	R cuadrado de Cox Snell	R cuadrado de Nagelkerke
1562,727	,523	,727

Finalmente resulta de interés analizar la capacidad predictiva del modelo, a partir de la clasificación que produjo el modelo sobre la muestra de validación. Estos resultados pueden apreciarse en la Tabla 6. El 88,6 % de bandejas de la muestra de validación fueron clasificadas correctamente. Más precisamente de las bandejas con tostado adecuado, el modelo clasificó bien al 82,6%; en tanto que de las bandejas con tostado inadecuado (por exceso o por defecto), clasificó exitosamente al 91,7%. Consecuentemente el modelo clasifica mejor las bandejas cuyo nivel de tostado es incorrecto.

Tabla 6 – Capacidad predictiva del modelo

Observados	Pronosticado					
	Casos seleccionado			Casos no seleccionado		
	TCORRECT 0	1	Porcentaje correcto	TCORRECT 0	1	Porcentaje correcto
T CORRECT	1832	157	91.5	549	50	91.7
1	158	839	84.2	52	247	82.5
Porcentaje Global			89.2			88.6

6 CONCLUSIONES

En este trabajo se propone una metodología automática y en línea, para distinguir distintos niveles de tostado de maní pelado sin piel a granel e introducir, en caso de fallas, correcciones en el proceso de tostado. El método propuesto utiliza la información brindada por sensores ópticos instalados a la salida del horno de tostado y está basado en el modelo de regresión logística. La construcción del modelo determinó que las variables Área y MAD son capaces de explicar la probabilidad de que el nivel de tostado de maní a la salida del horno sea el adecuado.

La arquitectura final propuesta consiste de una cámara IP la cual capta las imágenes de maní tostado a granel que salen del horno mientras son trasladadas por una cinta transportadora; estas imágenes son enviadas a la computadora cliente, la cuál recibe y procesa las imágenes enviadas por la cámara. Sobre cada imagen se calculan las medidas MAD y Área. Estos valores son introducidos en la ecuación del modelo de regresión logística hallado. Un resultado mayor que 0.5 indica que la temperatura del horno es adecuada para mantener el nivel de tostado

deseado; por el contrario, un resultado menor o igual a 0.5, indica que la temperatura del horno no es la correcta; esta señal se envía al PLC, el cual produce la corrección en tiempo real.

Una cuestión importante a tener en cuenta en la implementación del sistema, se refiere a los retardos. Los mismos se deben fundamentalmente a colisiones que se producen al intentar emitir y recibir datos hacia y desde el PLC. En una red Ethernet industrial deberían estar comprendidos entre 5 ms y 20 ms como máximo. Para reducir los retardos el sistema debe programarse de tal forma que el PLC intervenga sólo en situaciones excepcionales, en lugar de hacerlo continuamente; por ejemplo, el PLC debería actuar cuando se detecten variaciones de temperatura en el horno de tostado que no sean acordes a los niveles de tostado requeridos; en tanto que cuando las variaciones en la temperatura estén dentro de límites normales, el PLC no debería intervenir.

Al ser un control en tiempo real, los tiempos de cálculo entre la recepción de la medición de las variables del proceso y la señal de control, deben ser mínimos. Una ventaja del modelo logístico obtenido, es que al trabajar con una sola banda de color RGB, el algoritmo es mucho más rápido, ya que consume menos memoria que si se contemplan las tres bandas. Asimismo al trabajar con el modelo de color RGB, el cuál emplea el programa que envía las imágenes desde la cámara IP a la computadora cliente en la cuál se efectúa el procesamiento de la imagen, no es necesario transformar a otro modelo de color, abreviando el tiempo de procesamiento de la imagen. La ecuación obtenida es muy sencilla, no tiene operaciones ni funciones matemática complicadas, pudiéndose implementar en un programa a tiempo real, con un costo computacional muy bajo.

REFERENCIAS

BARON, M. **Probability and statistics for computers scientists**. Champan & Halls/CRC. Taylor & Francis Group. Boca Raton, London, New York, 432p, 2007.

BATAL, A.; DALE, N.; CAFÉ, M. Nutrient composition of peanut meal. **Journal of Applied Poultry Research**, v. 14, p. 254-257, 2005.

BEGBIE, G. H. **La visión y el ojo**. Editorial Universitaria de Buenos Aires (EUDEBA). 288p, 1977.

BUCKHOLZ, L. Jr.; DAUN, H.; STIER, E.; TROUT, R. Influence of roasting time on sensory attributes of fresh roasted peanuts. **Journal of Food Science**, v. 45, n. 3, p. 547-554, May, 1980.

DECOTIGNIE, J. Ethernet-based real-time and industrial communications. **Proceedings of the IEEE**. v. 93, n. 6, p.1102-1117, 2005.

FERNÁNDEZ, E.; GIAYETTO, O. **El cultivo del Maní en Córdoba**. Editorial de la Universidad Nacional de Río Cuarto, 280p, 2006.

FIEGUTH, P. **Statistical image processing and multidimensional modeling**. Springer Science+Business Media, LLC 2011. New York Dordrecht Heidelberg London Springer, 454p, 2011.

HOSMER D.; LEMESHOW S. **Applied logistic regression**. New York: Wiley 2nd Edition, 373p, 2000.

INTELLINET-NETWORK. **IP Intellinet 550710**. Disponible en: <<http://www.intellinet-network.com>> 2010.

MAGFOR. **Ministerio Agropecuario y Forestal. República de Nicaragua**. La Cadena Agroindustrial del Maní. Disponible en: <<http://cedoc.magfor.gob.ni/documentos/cedoc/E21-0162.pdf>> 2010.

MENARD S. **Applied logistic regression analysis**. 2nd edition. Thousand Oaks, CA: Sage Publications. 111p, 1995.

NAGHSHVAR, M.; ZHUANG, H.; JAVIDI, T. A General class of throughput optimal routing policies in multi-hop wireless networks". NJ.USA. ISBN 978-1-4442-5870-7. 47th. Allerton 09 **Proceedings of the 47th annual Allerton conference on Communication, control, and computing**. p. 1395-1402, 2009.

PREGIBON D. Logistic Regression Diagnostics. **Annals of Statistics**, v. 9, p. 705-724, 1981.

SAGPYA. **Secretaría de Agricultura, Ganadería Pesca y Alimentos**. República Argentina. Indicadores del Sector Manicero 2003/2004. Disponible en: <<http://www.sagpya.mecon.gov.ar/new/0-0/agricultura/otros/granos/mani.php>> 2010.

SANDERS, T. H. Effects of variety and maturity on lipid class composition of peanut oil. **Journal of the American Oil Chemists' Society**. v. 57, n. 1, p. 8-11, 2007.

Originais recebidos em: 27/09/11

Aceito para publicação em: 10/04/12