

# ANÁLISIS DE TEXTURA EN PANES USANDO LA MATRIZ DE COOCURRENCIA

Juan Sebastián Botero Valencia<sup>1</sup>, Alejandro Restrepo Martínez<sup>2</sup>

<sup>1</sup> Juan Sebastián Botero Valencia: Ingeniero Electrónico. Estudiante de Maestría en Automatización y Control Industrial. Investigador INSTITUTO TECNOLÓGICO METROPOLITANO.

e-mail: [juanbotero@itm.edu.co](mailto:juanbotero@itm.edu.co)

<sup>2</sup> Alejandro Restrepo Martínez PhD. Docente Investigador INSTITUTO TECNOLÓGICO METROPOLITANO.

e-mail [alejandrorestrepo@itm.edu.co](mailto:alejandrorestrepo@itm.edu.co)

## RESUMEN

El análisis de textura en imágenes digitales busca la descripción de patrones en superficies analizando en nuestro caso la ocurrencia. En el control de calidad de la industria panificadora es necesario para determinar la porosidad en los productos. En este trabajo se propone una estrategia de clasificación y análisis de panes usando medidas de textura con la matriz de coocurrencia. El objetivo del análisis es determinar las diferencias en las texturas internas de los panes, la correlación de los patrones con las diferencias en el proceso de producción y además de encontrar datos atípicos que evidencien fallos en el proceso de producción.

**Palabras clave:** Coocurrencia, homogeneidad, pan, reconocimiento de patrones, textura.

Recibido 10 de Mayo de 2010. Aceptado 16 de Junio de 2010

Received: May 10, 2010 Accepted: June 16, 2010

## TEXTURE ANALYSIS IN THE BREAD USING COOCURRENCE MATRIX

### ABSTRACT

*Texture analysis in digital images look the description of surface patterns in our case by analyzing the occurrence. In the quality control of the baking industry is needed to determine the porosity in the products. This paper proposes a strategy for classification and analysis of bread using texture measures with the co-occurrence matrix. The analysis aims to identify differences in the internal texture of the bread, the correlation of patterns with differences in the production process and also to find outliers that demonstrate flaws in the production process.*

**Keywords:** Co-occurrence, homogeneity, bread, pattern recognition, texture.

## 1. INTRODUCCIÓN

Aunque el concepto de textura es algo confuso, la textura se puede tomar como un descriptor usado para medir la estructura de una superficie y está directamente ligado a la percepción humana. Debido a la importancia de la porosidad en las propiedades del pan, es pertinente explorar relaciones que con el análisis digital de imágenes se puedan encontrar entre porosidad de la imagen y propiedades de los panes. En la literatura se encuentran varios trabajos orientados con esta tendencia, así: en [1] se hacen mediciones de porosidad sobre imágenes de microscopía electrónica de barrido y se relacionan los resultados a propiedades de pan. En [2] se relacionan, las medidas de los espesores de las paredes de las celdas que conforman la textura de la superficie del pan, con diversas propiedades mecánicas de los panes. En [3] se hace un análisis granulométrico 2D and 3D, usando la morfología matemática con los que se describen los granos del pan, además encontraron que las medidas en 2D, que son más económicas, fáciles de usar, y presentan un adecuado desempeño en describir las características de los panes. En [4] [5] se enfatiza en como una objetiva segmentación es requerida además presenta como algunas propiedades mecánicas del pan se relacionan a la imagen.

Para nuestro trabajo nos enmarcamos en la comparación de la textura de dos tipos de pan tajado (blanco e integral). Las texturas generadas por la porosidad de los panes, son analizadas con matrices de coocurrencia, método ampliamente utilizado en la descripción de texturas [6] [7] [8]. Para las muestras analizadas se establecen las características más relevantes y se relacionan los resultados, con los comportamientos de las imágenes de cada grupo de pan analizado. Vale aclarar que el enfoque del trabajo es lograr la diferenciación de dos tipos de pan analizando la porosidad, sin embargo, para trabajos posteriores sería pertinente relacionar las características de las imágenes con propiedades mecánicas.

Si se realiza una revisión del estado del arte, se encuentra que para el estudio de la medición de la porosidad de panes, se utilizan diversos métodos de análisis de texturas [9] [10] entre los que se encuentran: las matrices de coocurrencia, las matrices de diferencia de los niveles de gris del vecindario, el análisis estadístico de las características geométricas y el estudio de

características texturales a través de las frecuencias espaciales en el dominio de Fourier. Sin embargo, la utilización de las matrices de coocurrencia en el presente trabajo, no riñe con el aporte de este artículo, debido a las siguientes razones: El presente trabajo, contiene una diferencia metodológica con respecto a las referencias en la literatura, en procedimientos tales como: el mejoramiento, la segmentación y la extracción de las regiones de interés en la imagen. Además, que para la sección de las áreas de interés, no nos limitados a analizar, sólo la zona central del pan, situación recurrente en los trabajos revisados en la literatura. Esto es importante, ya que las zonas cercanas a los bordes del pan, en general no son consideradas, debido a que los poros en estas regiones presentan elongaciones, que generan texturas que no se encuentran en la zona central del pan. Adicionalmente, el presente trabajo, se aprovecha de la diferencias de las texturas dentro del pan, para generar de manera automática un crecimiento de regiones, basado en los parámetros texturales de la matrices de coocurrencia, que experimentalmente, nos presentaron una mayor relevancia para describir las texturas analizadas.

A causa de las condiciones de escala utilizadas en el trabajo, se encontró que los parámetros más relevantes de las matrices de coocurrencia, son diferentes a los reportados en la literatura. Esto, es un resultado importante, ya que no se puede considerar, que el desempeño de las matrices de coocurrencia, va a ser similar sin importar la escala en la que se analicen los poros. Se debe, decir que la escala utilizada en este trabajo es consecuencia de utilizar una óptica no muy especializada en el montaje.

Finalmente, se considera importante, divulgar en la comunidad científica y en la industria nacional, las posibilidades tecnológicas que tienen las matrices de coocurrencia para estudiar la porosidad de panes. Adicionalmente consideramos, que es necesario sensibilizar en el tema del análisis de la textura, ya que como se plantea en [11] [12] el estudio de la micro y macro estructura de los panes sirve para entender procesos térmicos y la influencia de la composición. También, es necesario destacar, que en general, para la industria de alimentos, el análisis de textura es una herramienta fundamental, para analizar propiedades y cualidades en los alimentos como plantea [13]. Y

debido, a que las matrices de coocurrencia, son una tan herramienta tan popular y de fácil acceso, es relevante, recalcar las posibilidades de su uso en la inspección visual automática aplicada a alimentos.

Para terminar la discusión de la revisión de la literatura, se debe destacar, que en el análisis de imagen aplicado a panes también se utilizan las mediciones de las distribuciones de tamaño de poro y las fracciones de área de los poros, como herramientas para describir y caracterizar los fenómenos que ocurren en el pan, pero el tipo de segmentación que se utiliza afecta, los anteriores descriptores como se plantea en [14]. Por esta razón, el uso de uso de estrategias texturales, es una opción muy importante en la descripción de las características micro y macro estructurales de los panes.

## 2. MATERIALES Y MÉTODOS

### 2.1 Base de datos

Para construir el conjunto de entrenamiento se capturaron 40 imágenes digitales (20 por clase) con 2848 x 4272 px (12 Mpx) de resolución en formato JPG a 8 bits. Se consideraron algunas reglas mínimas para el control del ambiente de captura, con el objetivo de no introducir problemas de iluminación y ruido que no se presentan en un ambiente industrial de trabajo. La única fuente de luz presente en el experimento fue una lámpara fluorescente de 32 W, ubicada perpendicularmente a 1.60 metros de altura. En la Figura 1 se muestra un ejemplo de la imagen obtenida de un pan blanco, y en la Figura 2 de un pan integral. Se uso un patrón de calibración de un centímetro y se obtuvo su equivalencia en pixeles

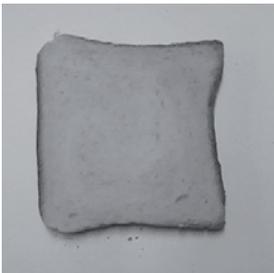


Figura 1. Pan Blanco



Figura 2. Pan Integral

### 2.2 Preprocesamiento

Luego de la adquisición de las imágenes fue necesario aplicar un proceso de adecuación con el

objetivo de hacerlas aptas para obtener las características. En primer lugar una umbralización de la imagen como se muestra en Figura 3

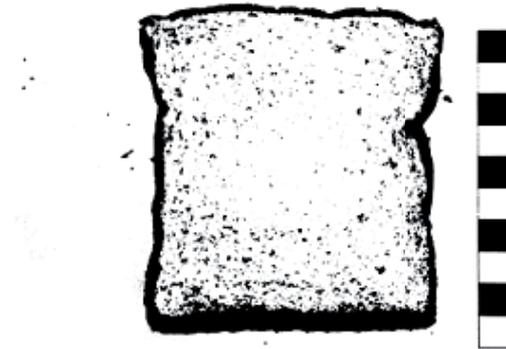


Figura 3. Umbralización

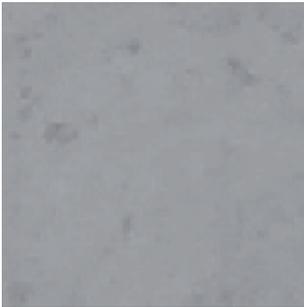
Después de obtener la imagen umbralizada se aplica un operador gradiente para encontrar una aproximación al borde. Con la estimación se selecciona un área de como se muestra en la Figura 4, imagen esta se divide en 5 subregiones. Con los índices de las posiciones extraídas se almacenó en un archivo la imagen resultante con espacio de color RGB, estas imágenes se constituyeron como el nuevo conjunto de prueba. Las subregiones se usan para calcular la matriz de coocurrencia por separado a 5 subregiones (de las 25) seleccionadas aleatoriamente sin repetición y con la idea de promediar el valor de la característica estimada.



Figura 4. Región de interés seleccionada

Adicionalmente fue necesario aumentar el contraste, con el objetivo de mejorar el rendimiento del análisis de texturas y determinar las características con mejor claridad, en la Figura 5 se observa un ejemplo de una imagen correspondiente al pan blanco antes de aumentar el contraste y en la Figura 6 después de aumentarlo.

Los valores para mapear la imagen se obtuvieron adaptativamente normalizando el histograma de cada imagen y considerando las distribuciones normales se mapeo de  $[\mu - \sigma \ \mu + \sigma]$  a  $[0 \ 1]$ , donde  $\mu$  es la media y  $\sigma$  la desviación estándar.



**Figura 5.** Pan blanco sin procesamiento

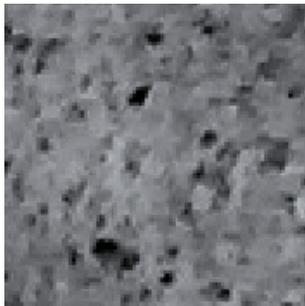


**Figura 6.** Pan blanco con procesamiento

En la Figura 7 y la Figura 8 se observa el aumento del contraste en una muestra de pan integral.



**Figura 7.** Pan integral sin procesamiento



**Figura 8.** Pan integral con procesamiento

### 2.3 Matriz de Coocurrencia

La textura en imágenes digitales describe la estructura de una superficie, esta descripción se puede obtener con propiedades como la granulosidad, el contraste y la dirección de las fibras. El análisis de texturas se fundamenta en el análisis de información redundante, y es por eso que puede presentar errores considerables cuando se aplica a zonas fronterizas donde existen cambios de textura. El análisis de textura es dependiente de que la distribución del vecindario sea uniforme con respecto a una medida y un periodo determinado.

Los elementos de la matriz de coocurrencia,  $P(i,j)$ , representan las frecuencias relativas de los niveles de gris  $i$  y  $j$ , tomando los píxeles dos a dos y separados una distancia  $d$  según una dirección dada. Así, cuanto mayores sean los valores de su

diagonal principal más homogénea será la textura que representa, mientras que cuanto más repartidos estén los valores fuera de la diagonal más heterogénea será. A partir de esta matriz se calcularon 8 variables estadísticas de segundo orden, propuestas por (Haralik, Shanmugan, & Dinstein, 1973), las cuales describen propiedades como uniformidad, contraste, media, entropía, varianza, momento producto, correlación y momento diferencia inverso.

Por la variedad de descriptores que es posible obtener de esta matriz, se puede caracterizar un conjunto de valores cuantificables para cada imagen analizada. Lo cual es importante para posteriormente alimentar un clasificador automático, el cual podría separar un conjunto de imágenes en diferentes clases.

#### 2.3.1 Contraste

El contraste de una textura proporciona información acerca de las variaciones bruscas de color en la imagen. La extracción de este descriptor está dada por la expresión mostrada en (1).

$$Co = \sum_{i=1}^n \sum_{j=1}^n |i - j|^2 c_{i,j} \quad (1)$$

$n = 8$

#### 2.3.2 Correlación

La correlación de la textura es una medida de la probabilidad que mide la relación entre las diferentes intensidades presentes en el entorno definido en la matriz de coocurrencia. Matemáticamente la correlación se describe en (2).

$$Cr = \frac{1}{\sigma_i \sigma_j} \sum_{i=1}^n \sum_{j=1}^n (i - \mu_i)(j - \mu_j) c_{i,j} \quad (2)$$

$n = 8$

#### 2.3.3 Energía

La propiedad de energía da una idea de la suavidad de la textura, y esto se refleja en la ubicación de sus probabilidades en la matriz de coocurrencia. De modo que si una mayor cantidad de píxeles de colores iguales cumplen con la condición de ocurrencia, será posible observar un pico de intensidad en la diagonal principal de la matriz. La forma de calcularla se muestra en (3).

$$En = \sum_{i=1}^n \sum_{j=1}^n c_{i,j}^2 \quad (3)$$

$n = 8$

### 2.3.4 Homogeneidad

El descriptor de homogeneidad local proporciona información sobre la regularidad local de la textura. Indica que mientras los elementos de la matriz de coocurrencia estén más próximos a la diagonal principal mayor será el valor de la homogeneidad local. Se muestra en (4).

$$H_o = \sum_{i=1}^n \sum_{j=1}^n \frac{c_{i,j}}{1 - |i - j|} \quad (4)$$

$n = 8$

## 3. RESULTADOS

Usando la matriz de coocurrencia se obtuvieron métricas con el objetivo de encontrar las características de mayor relevancia. Se tuvieron en cuenta, el contraste, la homogeneidad, la correlación y la energía por presentar mayor independencia estadística en la clasificación. Obtenidas las características descritas anteriormente, se obtuvo el espacio de características 2D. En la Figura 9 se muestra el espacio obtenido, usando el contraste y la correlación como medidas sobre la matriz de coocurrencia. Se observa que existe un alto nivel de separación entre las clases, y que el contraste en el pan blanco es uniforme.

La Figura 10 muestra el espacio de características usando la correlación y la energía. La alta dispersión de la clase Blanco indica que no se puede agrupar con facilidad en un clúster para usar la característica de forma discriminante. En el caso de la energía indica que la suavidad intra clase es variable para la clase Blanco, y estable para la clase Integral.

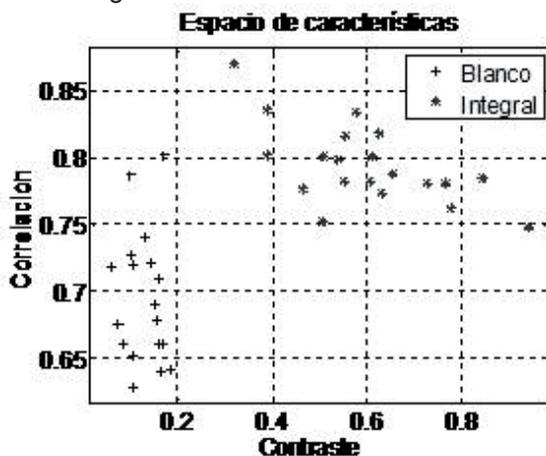


Figura 9. Espacio de características<sub>1</sub>

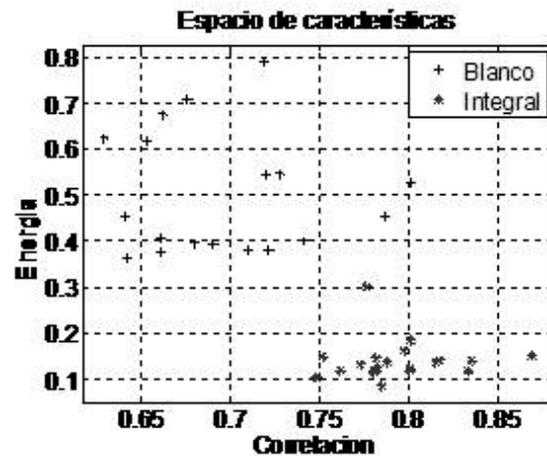


Figura 10. Espacio de características<sub>2</sub>

La Figura 11 muestra el espacio de características generado por la energía y la homogeneidad. Se observa que existe una relación casi lineal entre las características para ambas clases, lo que indica dependencia. Debido a la alta concentración de estas características en clúster mejor definidos, pueden ser usadas para identificar datos atípicos en la adquisición o en el control de calidad. Se observa también la alta homogeneidad en la clase Blanco lo que corresponde a una superficie más suave debida a su proceso de producción.

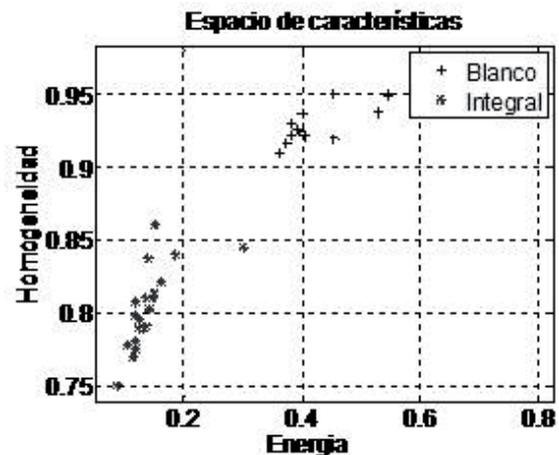


Figura 11. Espacio de características<sub>3</sub>

En la Figura 12 se observa el rendimiento de implementar un clasificado lineal y un clasificador bayesiano con el espacio de características<sub>2</sub>. El clasificador lineal presenta mejor rendimiento, aparte de tener una implementación más simple. El clasificador bayesiano esta clasificando mal un

objeto de la clase Integral, y se está sobre ajustando sobre los datos de esta clase.

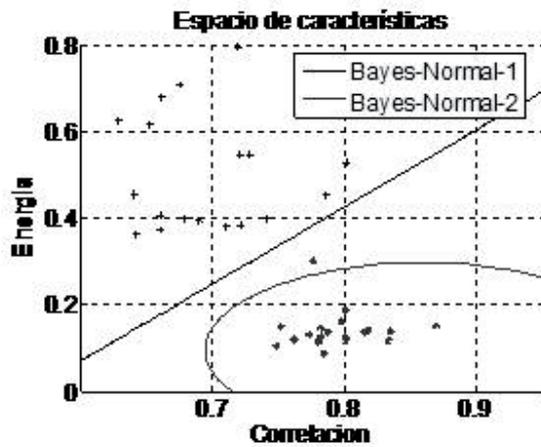


Figura 12. Espacio de características<sub>2</sub>

La Figura 13 muestra la Función de Distribución (FD) de las clases para la característica contraste. Es claro que el solapamiento es pequeño, y que esta característica sería suficiente para realizar la clasificación. Además se puede observar la aparición de una subclase en la clase blanco.

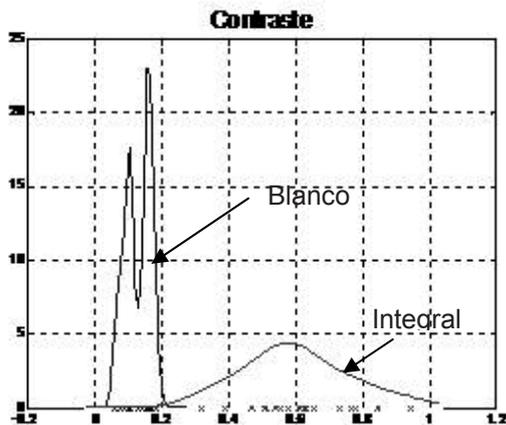


Figura 13. FD Contraste

La Figura 14 se observa la FD para la característica correlación. Se observa un pequeño solapamiento, y la presencia de varios máximos locales sobre la clase Blanco (varios clúster), las clases no pueden ser separadas únicamente con esta característica.

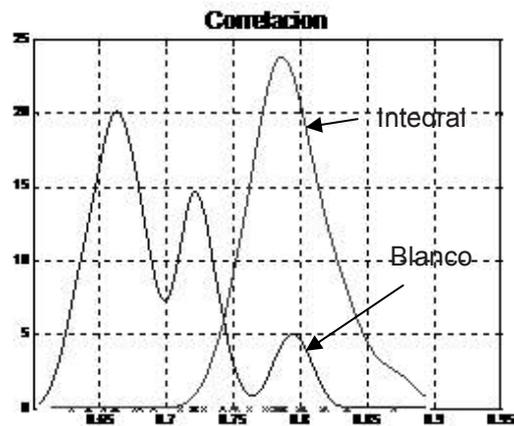


Figura 14. FD Correlación

En la Figura 15 se observa la FDP de la característica homogeneidad. De todas las características esta es la que posee la mayor capacidad discriminante. Es clara la alta homogeneidad en la clase blanco.

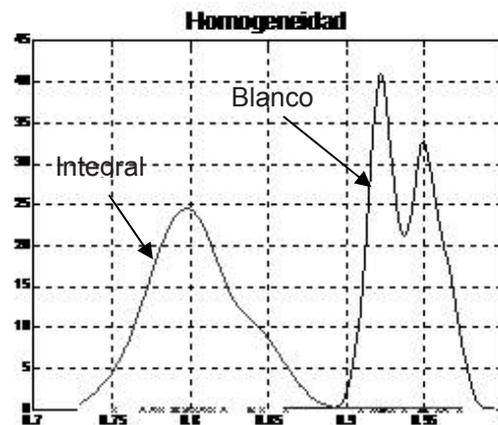
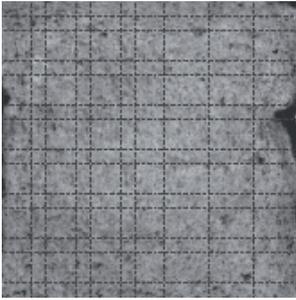
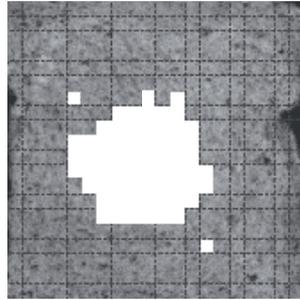


Figura 15. FD Homogeneidad

En la Figura 16 se observa el resultado de dividir la imagen segmentada en 20 subregiones. En la Figura 17 se muestra el resultado de buscar las subregiones con un umbral de [0.8 0.85] estimado a partir del valor medio de la característica homogeneidad en la clase pan blanco y tomando una desviación estándar a cada lado ( $[\mu - \sigma \mu + \sigma]$ ). Con la búsqueda de las subregión y usando la característica de homogeneidad se halla una región con textura homogénea como se muestra en la Figura 17.



**Figura 16.** Región de interés



**Figura 17.** Región central con valores homogéneos de textura.

#### 4. CONCLUSIONES

El uso de las métricas sobre la matriz de coocurrencia tiene un excelente resultado para este problema. La homogeneidad y la energía son las características más discriminantes.

En la representación inicial en escala de grises algunos problemas de adquisición hacían cercanas las características, pero el aumento del contraste usando operaciones de mapeo sobre el histograma, mejorar el rendimiento del algoritmo de clasificación.

Las propiedades de la matriz de coocurrencia están relacionadas con el proceso de fabricación de los panes, la textura en el pan blanco es más homogénea y posee un menor contraste que la del pan integral.

El sistema puede ser mejorado considerando un mayor número de objetos, y mejorando la resolución de las imágenes capturadas. Esto con el objetivo de encontrar características que se puedan relacionar con la composición de los panes.

#### 5. AGRADECIMIENTOS

Este trabajo está enmarcado en labores de investigación del Grupo MIRP, del Centro de Investigación del Instituto Tecnológico Metropolitano – Medellín Colombia.

#### 6. REFERENCIAS BIBLIOGRÁFICAS

[1] S. Ozge Ozkoc, G. Sumnu, and S. Sahin, "The effects of gums on macro and micro-structure of breads baked in different ovens," *Food Hydrocolloids*, p. 2182–2189, 2009.

[2] M. C. Zghal, M. G. Scanlon, and H. D. Sapirstein, "Cellular Structure of Bread Crumb and its Influence on Mechanical Properties," *Journal of Cereal Science*, pp. 167-176, 2002.

[3] N. Lassoued, P. Babin, G. Della Valle, M.-F. Devaux, and A.-L. Reguerre, "Granulometry of bread crumb grain: Contributions of 2D and 3D image analysis at different scale," *Food Research International*, pp. 1087-1097, 2007.

[4] M. G. ., Z. M. C. Scanlon, "Bread properties and crumb structure," *Food Research International*, pp. 841-864, 2001.

[5] Z. Liu and M. G. Scanlon, "Predicting mechanical properties of bread crumb," *Trans IChemE*, pp. 224-238, 2003.

[6] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, óth ed. New Jersey, United States: Prentice-Hall, 2008.

[7] R. M. Haralik, K. Shanmugan, and I. Dinstein, "Texture features for image classification," *IEEE Transactions on Systems, man and cybernetics*, pp. 610-621, 1973.

[8] M. Nixon and A. Aguado, *Feature extraction & image processing*. Bodmin: Newnes, 2002.

[9] U.A. Gonzales-Barron and F. Butler, "Statistical and spectral texture analysis methods for discrimination of bread crumb images," *IUFoST*, pp.749-759, 2006.

[10] U. Gonzales-Barron and F. Butler, "Discrimination of crumb grain visual appearance of organic and non-organic bread loaves by image texture analysis," *Journal of Food Engineering*, pp.480-488, 2008

[11] M.A Brescia, D. Sacco, A. Sgaramella, A.Pasqualone R. Simeone and G.P. A. Sacco, "Characterization of different typical Italian breads by means of traditional, spectroscopic and image analyses," *Food Chemistry*, Vol. 104, pp. 429–438, 2007

[12] S. O. Ozkoc, G.Sumnu and S. Sahin, "The effects of gums on macro and micro-structure of breads baked in different ovens", *Food Hydrocolloids*, Vol 23, pp. 2182–2189, 2009

[13] C. Zheng, D.W. Sun and L. Zheng, "Recent developments and applications of image features for food quality evaluation and inspection review", *Trends in Food Science & Technology*, Vol.17 pp.642-655, 2006

[14] U. Gonzales-Barron and F. Butler, "A comparison of seven thresholding techniques with the k-means clustering algorithm for measurement of bread-crumbs features by digital image analysis," *Journal of Food Engineering*, Vol. 74, pp. 268–278, 2006.