



FACULTAD DE INGENIERÍA

CARRERA DE INGENIERIA INDUSTRIAL

“SELECCIÓN DE CARACTERISTICAS DE INTERÉS EN LA
CLASIFICACIÓN DE GRANOS DE CAFÉ MEDIANTE UN
SISTEMA DE VISIÓN POR COMPUTADORA.”

Tesis para optar el título profesional de:

Ingeniero Industrial

Autor:

Perez Miranda, Stefany Carolyn

Rosell Llanos, Luis Adrián

Asesor:

Dr. Wilson Manuel Castro Silupu

Cajamarca – Perú

2016

ÍNDICE DE CONTENIDOS

Contenido

CAPÍTULO 1.INTRODUCCIÓN.....	11
1.1. Realidad problemática.....	11
1.2. Formulación del problema.....	12
1.3. Justificación.....	12
1.4. Limitaciones.....	13
1.5. Objetivos.....	13
CAPÍTULO 2.MARCO TEÓRICO.....	14
2.1. Antecedentes.....	14
2.2. Bases teóricas.....	18
2.2.1. <i>Café</i>	18
2.2.2. <i>Clasificación Visual</i>	22
2.2.3. <i>Espacios de color</i>	22
2.2.4. <i>Parámetros morfo geométricos</i>	25
2.2.5. <i>Determinación de características de interés</i>	25
2.2.6. <i>Clasificadores</i>	28
2.3. Hipótesis.....	36
CAPÍTULO 3.METODOLOGÍA.....	37
3.1. Diseño de investigación.....	37
3.2. Unidad de estudio.....	37
3.3. Población.....	37
3.4. Muestra (muestreo o selección).....	37
3.5. Materiales.....	37
3.5.1. <i>Muestras de granos de café</i>	37
3.5.2. <i>Sistema de adquisición de imágenes</i>	37
3.6. Métodos.....	40
3.6.1. <i>Clasificación visual</i>	41
3.6.2. <i>Obtención de Imágenes en RGB</i>	42
3.6.3. <i>Establecimiento de características de interés</i>	43
3.6.4. <i>Selección de parámetros de mayor importancia</i>	46
3.6.5. <i>Entrenador del clasificador</i>	46
CAPÍTULO 4.RESULTADOS.....	48
4.1. Diseño y construcción del sistema de visión por computadora.....	48
4.2. Adquisición y procesamiento de Imágenes RGB.....	49

4.3.	Obtención de los parámetros de patrones de color y características morfo geométricos	50
4.4.	Determinación de características de interés	54
4.5.	Generación y evaluación de clasificadores	54
CAPÍTULO 5. DISCUSIÓN		58
CAPÍTULO 6. CONCLUSIONES		59
CAPÍTULO 7. RECOMENDACIONES		59
REFERENCIAS BIBLIOGRÁFICAS		60

ÍNDICE DE FIGURAS

Figura 1: Producción total en los últimos cuatro años	20
Figura 2: Imagen RGB en formato matricial.....	22
Figura 3: Representación cubo de color RGB	23
Figura 4: Representación cúbica HSV	24
Figura 5: Representación gráfica de máquina de soporte vectorial (SVM)	30
Figura 6: Calculo de la distancia d entre los hiperplanos h1 y h2	31
Figura 7: Descripción estructural del DT	32
Figura 8: Descripción estructural de RNA	35
Figura 9: Diseño del sistema computacional.....	38
Figura 10: Diagrama de metodología propuesta.....	40
Figura 11: Clasificación de grano de café por defecto.....	41
Figura 12: Características de los parámetros del grano de café	44
Figura 13: Sistema continuo de adquisición de imágenes: (1) vista frontal del sistema (2) vista interior del sistema.....	48
Figura 14: Clasificación manual de granos de café	48
Figura 15: (a) Imagen en RGB; (b) Imagen en escala de grises; (c) Imagen segmentada ...	49
Figura 16: Distribución de parámetros morfo geométricos de los granos de café por clase	50
Figura 17: Distribución de parámetros de color Cie L*a*b'' de granos de café por clase ...	51
Figura 18: Distribución de parámetros de color RGB de los granos de café por clase	52
Figura 19: Distribución de parámetros de color HSV de los granos de café por clase.....	52
Figura 20: Distribución de las relaciones entre colores RGB en granos de café por clase ..	53
Figura 21: Distribución de las diferencias normalizadas entre parámetros RGB, (a) diferencia normalizada de verdes y rojos (b) diferencia normalizada de verdes y azules	53
Figura 22: a) ACC y b) ERR al clasificar granos de café	57

ÍNDICE DE TABLAS

Tabla 1: Descripción general de grados de café arábica	21
Tabla 2: Métodos de selección de variables	26
Tabla 3: Ejemplo de matriz de confusión.....	36
Tabla 4: Distribución de granos de café por clase	41
Tabla 5: Parámetros morfo geométricos	44
Tabla 6: Característica de forma y color	45
Tabla 7: Selección de características de mayor relevancia	54
Tabla 8: Matrices de confusión para las técnicas de caracterización	55
Tabla 9: Matrices de confusión usando características de interés	56

RESUMEN

La clasificación de la calidad del café antes de ser tostado, una de las operaciones más importantes para definir su calidad y precio en el mercado, se realiza manualmente por personal entrenado en el reconocimiento de los defectos del café. Sin embargo, el carácter subjetivo, costo y tiempo que este involucra genera un campo de investigación importante para la aplicación de tecnologías como la visión artificial. El objetivo de este estudio fue evaluar la capacidad de identificación de defectos y clasificación de granos de café mediante un sistema de visión por computadora en el espacio red-green- blue (RGB). Para este fin se implementó un sistema de adquisición y análisis de imágenes, desarrollando una aplicación informática en Matlab 2015^a. Se compraron en el mercado local muestras de café verde, clasificando cada grano de acuerdo con la NTP 209.027 2001. Se adquirieron las imágenes de cada clase y se analizaron determinando en cada grano seis parámetros de forma, seis parámetros de color, en los espacios RGB y HSV, y dos índices o diferencias normalizadas. Se determinaron los parámetros con influencia estadística en la clasificación mediante software de análisis de datos WEKA y se implementaron tres modelos de clasificación máquinas de soporte vectorial (Support Vector Machine - SVM), arboles de decisión (Decision Tree - DT) y K-vecino más cercano (K-Nearest Neighbor). Los tres tipos de clasificador utilizados en la presente investigación mostraron precisión entre 89% y 92.3% lo cual prueba la posibilidad de implementar sistemas basados en imágenes RGB para clasificar granos de café.

ABSTRACT

The classification of quality on coffee before toasting, one of the most important operations to define its quality and price in the market, is done manually by personnel trained in the recognition of coffee defects. However, the subjective nature, the cost and the time that it involves generates an important research field for the application of technologies such as artificial vision. The objective of this study was to evaluate the ability to identify defects and classify coffee beans using a computer vision system in the red-green-blue (RGB) space. For this purpose, a system for acquisition and analysis of images was implemented, developing a computer application in Matlab 2015^a. Samples of green coffee were purchased on local market, each grain being classified according to NTP 209.027 2001. Images of each class were acquired and analyzed by determining in each grain six shape parameters, six color parameters, in the RGB and HSV spaces, and two normalized indices or differences. Statistical relevance of parameters was determined using the software for data analysis named WEKA and using these three models of classification, vector machines (SVM), decision trees and nearest K-neighbor (K-neighbor), were implemented. The three types of classifier used in the present investigation show accuracy between 89% and 92.3% which probe the possibility to implement systems based on RGB image to classify coffee beans.

NOTA DE ACCESO

No se puede acceder al texto completo pues contiene datos confidenciales

REFERENCIAS BIBLIOGRÁFICAS

- Alzate Ocampo, N. (2015). Importancia del gen sh3 y su relación con la resistencia a la enfermedad roya del cafeto *hemileia vastatrix* bajo la modalidad de compilación bibliográfica. Retrieved from <http://repository.unad.edu.co/handle/10596/3385>
- Arakeri, M., & Lakshmana. (2016). Computer Vision Based Fruit Grading System for Quality Evaluation of Tomato in Agriculture industry. *Procedia Computer Science*, 79, 426–433.
- Araújo, S., Pessota, J., & Kim, H. (2015). Beans quality inspection using correlation-based granulometry. *Engineering Applications of Artificial Intelligence*, 40, 84–94.
- Aroca-Santos, R., Cancilla, J. C., Pariente, E. S., & Torrecilla, J. S. (2016). Neural networks applied to characterize blends containing refined and extra virgin olive oils. *Talanta*, 161, 304–308.
- Arvy, M.-P., Gallouin, F., Ubillos, M. Á. M., & Montalbán, J. M. (2007). *Especias, aromatizantes y condimentos* (Vol. 77). Mundi-Prensa. Retrieved from <http://www.sidalc.net/cgi-bin/wxis.exe/?IsisScript=AGRISSA.xis&method=post&formato=2&cantidad=1&expresion=mfn=004350>
- Ayitenfsu, B. (2014). Method of Coffee Bean Defect Detection (Vol. Vol. 3-Issue 2 (February-2014)). Presented at the International Journal of Engineering Research and Technology, IJERT. Retrieved from <https://www.ijert.org/view-pdf/8387/method-of-coffee-bean-defect-detection>

- Benalia, S., Cubero, S., Prats-Montalbán, J., Bernardi, B., Zimbalatti, G., & Blasco, J. (2016). Computer vision for automatic quality inspection of dried figs (*Ficus carica* L.) in real-time. *Computers and Electronics in Agriculture*, *120*, 17–25.
- Bicho, N., Leitão, A. E., Ramalho, J., & Lidon, F. (2014). Application of colour parameters for assessing the quality of Arabica and Robusta green coffee. *ResearchGate*, *26*(1), 9–17.
- Borrelli, R. C., Visconti, A., Mennella, C., Anese, M., & Fogliano, V. (2002). Chemical Characterization and Antioxidant Properties of Coffee Melanoidins. *Journal of Agricultural and Food Chemistry*, *50*(22), 6527–6533. <https://doi.org/10.1021/jf025686o>
- Condori, R. H. M., Humari, J. H. C., Portugal-Zambrano, C. E., Gutiérrez-Cáceres, J. C., & Beltrán-Castañón, C. A. (2014). Automatic classification of physical defects in green coffee beans using CGLCM and SVM. In *2014 XL Latin American Computing Conference (CLEI)* (pp. 1–9).
- Craig, A., Franca, A. ., & Oliveira, L. (2012). Evaluation of the potential of FTIR and chemometrics for separation between defective and non-defective coffees. *Food Chemistry*, *132*(3), 1368–1374.
- Cruz, J., Silupu, W., & Lopez, L. (2016). Effect of different combinations of size and shape parameters in the percentage error of classification of structural elements in vegetal tissue of the pumpkin *Cucurbita pepo* L. using probabilistic neural networks. *Revista Facultad de Ingeniería*, *0*(78), 30–37.
- Dan, L., Sun, D., & Zeng, X. (2013, October). Recent Advances in Wavelength Selection Techniques for Hyperspectral Image Processing in the Food Industry. Retrieved September 6, 2017, from

https://www.researchgate.net/publication/260524256_Recent_Advances_in_Wavelength_Selection_Techniques_for_Hyperspectral_Image_Processing_in_the_Food_Industry

de Oliveira, E., Leme, D., Barbosa, B., Rodarte, M., & Pereira, R. (2016). A computer vision system for coffee beans classification based on computational intelligence techniques. *Journal of Food Engineering*, 171, 22–27.

Du, C., & Sun, D. (2004). Recent developments in the applications of image processing techniques for food quality evaluation. *Trends in Food Science & Technology*, 15(5), 230–249.

Edwards-Murphy, F., Magno, M., Whelan, P. M., O'Halloran, J., & Popovici, E. M. (2016). b+WSN: Smart beehive with preliminary decision tree analysis for agriculture and honey bee health monitoring. *Computers and Electronics in Agriculture*, 124, 211–219.

ElMasry, G., Cubero, S., Moltó, E., & Blasco, J. (2012). In-line sorting of irregular potatoes by using automated computer-based machine vision system. *Journal of Food Engineering*, 112(1–2), 60–68.

Gonzalez, R. ., & Woods, R. . (2007). Image processing. *Digital Image Processing*, 2. Retrieved from http://debut.cis.nctu.edu.tw/Courses/courseOld/IP/ip2005Night/IP_slides/Chapter1.pdf

Gotteland, M., & de Pablo V, S. (2007). ALGUNAS VERDADES SOBRE EL CAFÉ. *Revista Chilena de Nutrición*, 34(2), 105–115. <https://doi.org/10.4067/S0717-75182007000200002>

- Guerra Burgos, J. O., & Welchez Arita, J. A. (2013). Evaluación de la efectividad de cuatro fungicidas biológicos en el control del hongo de la roya de café *Hemileia vastatrix*. Retrieved from <http://bdigital.zamorano.edu/handle/11036/1758>
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. . (2009). The WEKA data mining software: an update. *ACM SIGKDD Explorations Newsletter*, 11(1), 10–18.
- Hu, J., Li, D., Duan, Q., Han, Y., Chen, G., & Si, X. (2012). Fish species classification by color, texture and multi-class support vector machine using computer vision. *Computers and Electronics in Agriculture*, 88, 133–140.
- International Coffee Organization. (2017). International Coffee Organization. Retrieved October 5, 2017, from http://www.ico.org/new_historical.asp?section=Statistics
- Kim, H. Y., Maruta, R. H., Huanca, D. R., & Salcedo, W. J. (2013). Correlation-based multi-shape granulometry with application in porous silicon nanomaterial characterization. *Journal of Porous Materials*, 20(2), 375–385.
- Kim, K. (2016). A hybrid classification algorithm by subspace partitioning through semi-supervised decision tree. *Pattern Recognition*, 60, 157–163.
- Kılıç, K., Boyacı, I., Köksel, H., & Küsmenoğlu, İ. (2007). A classification system for beans using computer vision system and artificial neural networks. *Journal of Food Engineering*, 78(3), 897–904.
- Kurtulmuş, F., Alibaş, İ., & Kavdır, I. (2016). Classification of pepper seeds using machine vision based on neural network. *International Journal of Agricultural and Biological Engineering*, 9(1), 51–62.

- León, K., Mery, D., Pedreschi, F., & León, J. (2006). Color measurement in L*a*b* units from RGB digital images. *Food Research International*, 39(10). Retrieved from <http://www.sciencedirect.com/science/article/pii/S0963996906000470>
- Lopez, J., Cobos, M., & Aguilera, E. (2011). Computer-based detection and classification of flaws in citrus fruits. *Neural Computing & Applications*, 20(7), 975–981.
- Lu, G., & Fei, B. (2014). Medical hyperspectral imaging: a review. *Journal of Biomedical Optics*, 19.
- Mancha Agresti, P. ., Franca, A. ., Oliveira, L., & Augusti, R. (2008). Discrimination between defective and non-defective Brazilian coffee beans by their volatile profile. *Food Chemistry*, 106(2), 787–796.
- Mery, D., Chanona-Pérez, J. J., Soto, A., Aguilera, J. M., Cipriano, A., Veléz-Rivera, N., ... Gutiérrez-López, G. F. (2010). Quality classification of corn tortillas using computer vision. *Journal of Food Engineering*, 101(4), 357–364.
- Minassie, H. (2008, January). *Image Analysis for Ethiopian Coffee Classification* (Thesis). Addis Ababa University. Retrieved from <http://10.6.20.92:80/jspui/handle/123456789/9415>
- Montoya Holguín, C., Cortés Osorio, J., & Chaves Osorio, J. (2014). Sistema automático de reconocimiento de frutas basado en visión por computador. *Automatic Recognition System of Fruits Based Computer Vision.*, 22(4), 504–516.

- Pedreschi, F., Mery, D., Bungler, A., & Yañez, V. (2011). Computer Vision Classification of Potato Chips by Color. *Journal of Food Process Engineering*, 34(5), 1714–1728.
- Puerto, D. A., Gila, D. M. M., García, J. G., & Ortega, J. G. (2015). Sorting Olive Batches for the Milling Process Using Image Processing. *Sensors*, 15(7), 15738–15754.
- Rojo Jiménez, E., Pérez-Urria Carril, E., Rojo Jiménez, E., & Pérez-Urria Carril, E. (2014). Café I (G. Coffea). *REDUCA Biología*, 7(2), 113–132.
- Romero-Acero, A., Marín-Cano, A., & Jiménez-Builes, J. (2015). Sistema de clasificación por visión artificial de mangos tipo Tommy. *Classification System for Artificial Vision Type Tommy Mango.*, 14(1), 21–31.
- Santos, J., Sarraguça, M., Rangel, A., & Lopes, J. (2012). Evaluation of green coffee beans quality using near infrared spectroscopy: A quantitative approach. *Food Chemistry*, 135(3), 1828–1835.
- SCAA. (2014). 2014 Portrait Country: Peru | The Specialty Coffee Chronicle. Retrieved December 12, 2016, from <http://www.scaa.org/chronicle/2013/12/30/2014-portrait-country-peru/>
- Silva, L., & Lizcano, S. (2012). Evaluación del estado de maduración de la piña en su variedad perolera mediante técnicas de visión artificial. *Iteckne*, 9(1), 31–41.
- Szczypiński, P., Klepaczko, A., & Zapotoczny, P. (2015). Identifying barley varieties by computer vision. *Computers and Electronics in Agriculture*, 110, 1–8.
- Taghizadeh, M., Gowen, A. ., & O'Donnell, C. . (2011). Comparison of hyperspectral imaging with conventional RGB imaging for quality evaluation of Agaricus

- bisporus mushrooms. *Biosystems Engineering*. Retrieved from <http://www.sciencedirect.com/science/article/pii/S1537511010002163>
- Tolessa, K., Rademaker, M., De Baets, B., & Boeckx, P. (2016). Prediction of specialty coffee cup quality based on near infrared spectra of green coffee beans. *Talanta*, *150*, 367–374.
- Visa, S., Ramsay, B., Ralescu, A., & Knaap, E. (2011). *Confusion Matrix-based Feature Selection*.
- Yang, L., Yang, F., & NOGUCHI, N. (2011). Apple Internal Quality Classification Using X-ray and SVM. *IFAC Proceedings Volumes*, (1). Retrieved from <http://www.sciencedirect.com/science/article/pii/S1474667016458994>
- Zhang, Y., & Wu, L. (2012). Classification of Fruits Using Computer Vision and a Multiclass Support Vector Machine. *ResearchGate*, *12*(9), 12489–505.