Automation of the brazil-nuts classification process using dynamic level set

Sonia Castelo Quispe*, July Diana Banda Tapia*, Mónika N. López Paredes*,
Dennis Barrios-Aranibar* and Raquel Patiño Escarcina*
*San Agustín National University
Research and Software Development Center,
Cátedra Concytec in TIC's
Arequipa-Perú
{scastelo2, inf.jdtapia, kaloni25, dennisbarrios, raquel.patino}@gmail.com

***Abstract*—The usual method for classification processes of brazil-nut is manual and present some drawbacks like slowness, subjectivity, and inconsistency. In this paper, the main objective is to automate the classification process by analysing digital images with multiple brazil-nuts. These images have been segmented using the Level Set method without reinitialization with a new stopping criteria based on the area of objects, which have been proposed in order to reach better results. The goal is to optimize the manual classification process that has been done until now by generating higher productivity. The efficiency achieved by the proposal is 97.63% for whole brazil-nuts and 84.09% for broken brazil-nuts for image taken at a distance of 40 cm and 99.55% and 98.44% for whole and broken brazil-nuts respectively for image taken at a distance of 30 cm, properly classified.**

***Keywords*-classification process; active contours; level set method; automation**

I. Introduction

Nowadays, one of the main economic activities in cities near the forest such as Madre de Dios in Peru, the north of Bolivia, Brazil, etc. is the production and commercialization of non-timber products like brazil-nuts. They are dry fruits and does not need special care for its cultivation. However, before to its exportation is necessary a classification process, which is done manually. One of the main classification criteria is given by the shape of the brazil-nut depending on if it is whole or broken [1].

There are researches based on external defects classification, like in tomatoes, oranges, pistacho-nuts, etc., which use computer vision and image analysis, because is a trend adopted by many companies in this sector. These have given support to this investigation. For example, in Spain was presented a research [2], showing the use of computer vision for damage classification in fruits, which was carried out by human inspection. They developed an approach using a computer vision system available to detect defects in the fruit peel and classify the type of damage. Others presented different methodologies to recognize and classify images [3], like the classifier based on color, texture and morphological features to recognize and classify horticultural products. Likewise, a proposed system allowed to evaluate the quality of beans based on size and color quantification of the samples. This analysis performed was based on their color intensity distribution. [4].

Also to reduce misclassification, was developed a computer vision framework to automatically classify the quality of corn tortillas according to five subclasses given by a sensory panel. Once the development of a feature selection algorithm is done, the most relevant features are selected for classification [5].

This means that the automation of processes involving the classification of products has been implemented by its high contribution to companies by reducing classification errors as well as accelerating the process; but the first step in many computer vision applications, is segmentation. The goal of it is to cluster pixels into salient image regions, which are called objects, and the rest of the image is known as background. It can be used for object recognition and occlusion boundary estimation without motion [6]. Then, a better segmentation process could improve the classification.

In this way, active contours is a technique of continuous contour which elastically wrap and enclose a target object by locking on its edges. In order to control it, is used a function of energy based on look for the maximum gradient value over each line of the image. After that, an algorithm tries to follow the maximum gradient around the image until the new initial point is found, and look for the new maximum gradient again [7]. Then, then Level-Set method introduced by Osher and Sethian [8] in 1988 had far-reaching impact in different applications, such as computational geometry, fluid dynamics, image processing, and computer vision [9].

In this sense, this paper propose to automate the brazil-nut classification process in whole and broken ones, based on the relation between mayor and minor axis of each brazil-nut. To segment the images was used the level set method with new stopping criteria based on the area of the objects.

The rest of the paper is organized as follow: section 2, describes the art state of classification techniques based on computer vision. In section 4, is the approach proposed. Section 5, shows the experiments and results. Finally, section 6 draws conclusions.

978-1-4577-2152-6/11/$26.00 ©2011 IEEE

II. State of the Art

Nowadays, several kinds of classification methods have been developed such as desicion tree induction, Bayesian networks, k-nearest neighbor classifier, case-based reasoning, genetic algorithms and fuzzy logic techniques, which are grouped in supervised and unsupervised [8].

In this sense, Neural Networks have been successfully applied to a variety of applications in industry. In the specific case of fruits classification, this process have been based on criterias such as size, color, shape, moprphological features, texture and defects. Another methods are based on a Backpropagation Neural Network (BPNN) [10], [3]. Another case where neural networks were use was in a classification system for beans [4], whose reach a performance of 90.6%. One of the main advantages of the Neural Networks is that do not use threshold values, but need a training process, which can be a disadvantage in real time systems. Similarly, a Fuzzy method [11] was used to classify rice grains; the inputs of this system were the area, perimeter, circularity and compactness. Compared con human inspection this system reached a 90% of correct results.

Likewise, a standard non-linear Bayesian discriminant analysis was used to determine the classification functions, in order to identify the type of the defect affecting the skin damage in citrus fruits, [12]. This system worked with spectral information about the defects with morphological estimations. Its success rate reached 86%.

In the case of the quadratic analysis, most of the time gave a more accurate classification, but not significantly better than the discriminant analysis, as was showed in a color classifier for symptomatic soybean seeds [13], where the classification accuracy for linear and quadric functions ranged from 67 to 81%. Another studies [14], [5] to determine a quantitative classification algorithm for fruit shape in kiwifruit and based on support vector machie were done.

However, on the common problems in all of these methods is the segmentation process, which allow to identify the objects in a image. In this sense, many segmentation algorithms have been developed [15], being the Sober Filter and Canny considered the most widely used such as edge detection method, however these techniques present low efficiency in high level processing, or when there are a high presence of noise.

As we can see, another common problem of the majority of segmentation algorithms is the presence of noise in the image, for which new approaches have been done. One of them are the active contour models, which have been recently studied by their advantages, as its robustness under high level of noise in a image, being able to segment more complex images such as medic images that the traditional segmentation methods can not do it. One way of implementation of active contour models is the Level Set method [16].

This method was introduced by Osher and Sethian and is based on the curve evolution equation with a zero level set in a given point of time t. The evolution function changes its topology through the time and allow to identify the shape of segmented object. However one of the main problems of this method is the stopping criteria, which is given by the iteration number defined by the user. In this sense, it is necesary to stablish a stopping criteria that dynamically determine the number of iterations, based on some properties of the objects such as the area, and can be applied in all kind of images.

III. Approach

Actually, to determine if the quality of a brazil-nut is a whole or broken one, people have to make a manual process by using visual inspection, which present drawbacks such as fatigue, slowness and others. This is the reason and importance of automate the classification process. Likewise, and for the segmentation process, was used the Level Set method without reinitialization; however, one of its main problems is the stopping criteria, which is given by the number of iterations. In the case of the brazil-nut segmentation without a stopping criteria many of the brazil-nuts are not segmented correctly, for this reason we have implemented a new stopping criteria based on the area of objects which had better results.

This stopping criteria is based on the area's variation of the brazil-nuts in the images, to reach a better segmentation that allows a feature extraction to stablish if a brazil-nut is whole or broken. In order to develop this stopping criteria, the area value is calculated for each object which have been already segmented by the evolution of Level Set function. Then, all values of the area of the objects are compared with the areas obtained in previous evolutions, to determine their variation; the evolution continues until the variation of areas from an evolution to another is very small. Additionally it is considered the difference between the mean area of brazil-nuts segmented and the global area of those that have not been segmented.

After the image have been segmented, next step is the feature extraction of each object that allows to classify the brazil-nuts. This process is done using an threshold value E, which allows to stablish if a brasil-nut is whole or broken; this value is calculated based on the major and minor axis as it is shown in the equation 1:

$$E = MajorAxis/MinorAxis \tag{1}$$

The following algorithm describes segmentation, feature extraction process and, classification process:

Algorithm 1 Classification Process using Dinamic Level Set

Given an image I;
Transformation of Intensities
Change Color Spaces using the Channel R
//Segmentation using the new approach of the Level set
Set: image I, zero LSF, k as the level of similarity between objects
while $(Area > k * Aver\_Areas) or (regions = 1)$ **do**
  Set the previous number of regions
  Get the number of regions with their areas and centres coordinates by applying the evolution Level Set
  if number regions > previous number regions **then**
    Look for similar center objects in the actual iteration with the previous one
    if Area of previous iter. object = Area of actual iter. object **then**
      Delete border of this object
      Return a new map of Level Set border
    end if
  end if
  Set the size of area without segmentation
  Set the average area, the array of areas.
end while
Binarize(I);
//Classification
for $i = 1 \rightarrow TotalRegions$ **do**
  Major axis= Image.region(i).major;
  Minor axis= Image.region(i).minor;
  Calculate the relation e=Major axis/Minor axis ;
  if e> threshold defined **then**
    It is whole
  else
    It is broken
  end if
  $i \leftarrow i + 1$
end for

IV. Experiments and Results

To evaluate the efficiency of the proposed classification method using new stopping criteria, it have been tested on images of brazil-nuts, taken at 30 cm and 40 cm focal distance, and the results have been compared using the traditional level set method without the stopping criteria, only based in the iteration number.

Database: It has 2335 whole and 563 broken samples of brazil-nuts distributed in 65 digital images where 30 correspond to whole brazil-nuts, 10 to the broken ones, and 25 images that contain both, broken and whole brazil-nuts. Images were taken in groups of 40 to 80 brazil-nuts as it is showed in Figure 1. These images were taken with a Canon G-9 Digital camera, aperture value of F8.0, shutter speed of 1/125 seconds, and ISO 400 in a close environment and controlled white illumination. these images were taken at a distance of 40 and 30 cm. This is show in the Figure 1.

Figure 1. Original image

Pre-processing and Segmentation To the image pre-processing, the intensity transformation functions used were the Gamma Transformation and the Contrast Stretching Transformation; the first one allows us to curve the grayscale components either to brighter or darker intensity, likewise, Contrast-stretching Transformation increase the contrast and is used as an input to the gamma function. The color space used the Channel R to enhance the brazil-nuts. See figures 2 and 3.

Figure 2. Transformation of Intensities

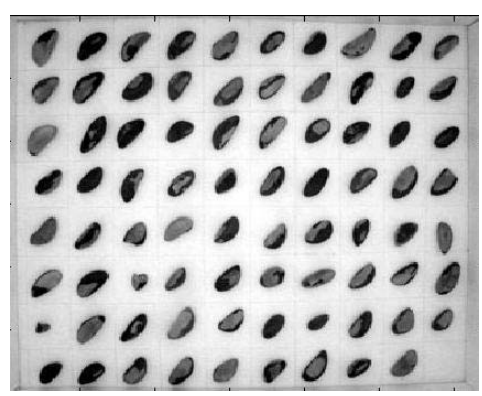

Figure 3. Color Spaces-Channel R

The segmentation process was carried out efficiently using the Level Set Method without reinitialization and using the new stopping criteria described in the previous section. Figure 4 and 5 shows the segmentation process

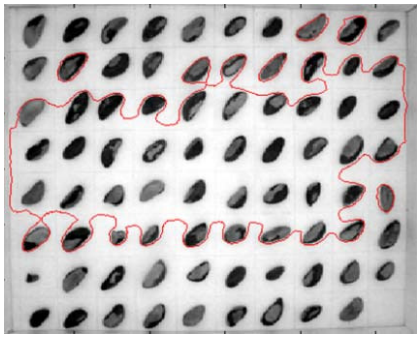

Figure 4. During the ejecution

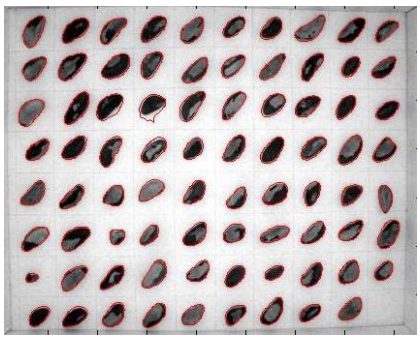

Figure 5. After Segmentation

Feature Extraction and Classification Process After the image have been segmented the next stept is the classification process; for which a binarization and features extraction process will be done.

- Binarization: to convert an image of up to 256 gray levels to a black and white image. We choose a threshold global value, and classify all pixels with values above this threshold as white, and bellow as black. This process is made with the function bwlabel(in Matlab) which allows regions identification. After the binarization result we identified several regions. Each region works independently, so calculations are made for each region. In this case, between 40 and 45 regions will be recognized, depending on the input image.
- Measures Extraction: the classification process is done using an umbral value E, in order to get an accurate quantity of good sorting as much of wholes as of brokens brasil-nuts. For this, two values are calculated: major axis and minor axis of the brasil-nuts, and a ratio between these is gotten:

$$E = MajorAxis/MinorAxis \tag{2}$$

To stablish the threshold E for the classification process, many experiments have been done, as show Table II and, table I. The best threshold found was 1.669 and 2.03 for images whit 40 cm and 30 cm of focal distance as shows the figure 6 and, 7.

Table I
PERCENTAGE OF BRAZIL-NUTS CORRECTLY RECOGNIZED OF $30cm$.

Threshold	Whole Nuts(%)	Broken Nuts(%)
1.78	88.54	54.02
1.83	90.50	62.70
1.88	92.86	72.67
1.93	94.93	81.03
1.98	97.44	90.68
2.03	99.55	98.44
2.08	96.91	92.60
2.13	92.35	83.48
2.18	88.05	76.39
2.23	82.55	68.90
2.28	74.36	60.13

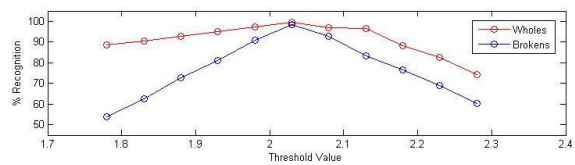

Figure 6. Percentage of brazil-nuts recognized correctly (photos taken at a distance of $30cm$)

Table II
PERCENTAGE OF BRAZIL-NUTS CORRECTLY RECOGNIZED OF $40cm$.

Threshold	Whole Nuts(%)	Broken Nuts(%)
1.419	81.82	30.97
1.469	84.31	42.18
1.519	86.74	52.51
1.569	91.01	69.32
1.619	91.01	88.50
1.669	97.63	84.09
1.719	90.12	76.52
1.769	82.53	64.82
1.819	73.60	54.94
1.869	66.29	48.85
1.919	57.36	43.02

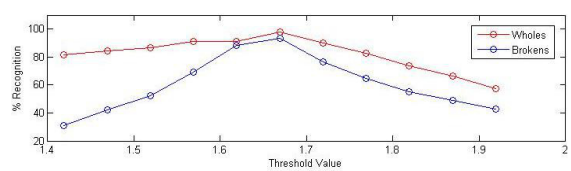

Figure 7. Percentage of brazil-nuts recognized correctly (photos taken at a distance of 40cm)

Finally, after having performed the experiments using the traditional level set method without the stopping criteria, the threshold is 1.40 as show the figure 8, where the segmentation achieve an accuracy of 96.32% and 87.65% for whole and broken brazil-nuts respectively as show Table III.

In contrast, the segmentation process using the new stopping criteria in the level set method the accuracy have been improved reaching a 97.63% and 84.09% for whole and broken brazil-nuts respectively, for images taken at a distance of 40 cm. and 99.55% and 98.44% for whole and broken brazil-nuts respectively for image taken at a distance of 30 cm as it is shown in Table IV.

Table III
PERCENTAGE OF BRAZIL-NUTS CORRECTLY RECOGNIZED WITHOUT STOPPING CRITERIA ($40cm$)

Umbral	Whole Nuts(%)	Broken Nuts(%)
1.15	76.37	4.56
1.20	78.93	17.63
1.25	83.75	40.12
1.30	88.72	60.79
1.35	95.39	85.11
1.40	96.32	87.65
1.45	89.26	75.11
1.50	81.48	63.64
1.55	74.58	56.05
1.60	66.31	49.03
1.65	57.83	43.46

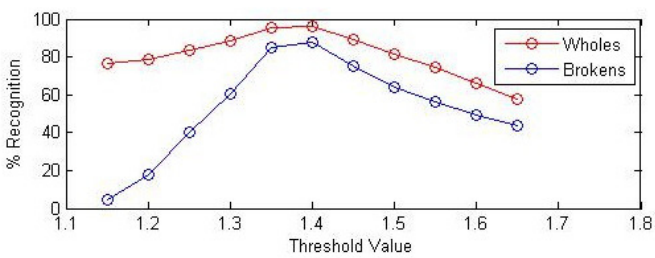

Figure 8. Percentage of brazil-nuts recognized correctly without stopping criteria ($40cm$)

Table IV
RESULTS OF CLASSIFICATION

Distancia	$40cm$		$30cm$	
	Total	Accuracy	Total	Accuracy
Brokens	301	97.63%	262	99.55%
Wholes	1165	93.13%	1170	98.44%

In table V is showed the results for segmentation process using the new stopping criteria for pictures taken at 40 cm and 30 cm.

Table V
RESULTS OF SEGMENTATION WITH NEW STOPPING CRITERIA

Distancia	$40cm$		$30cm$	
	Total	Accuracy	Total	Accuracy
Brazil-Nuts	1466	90.18%	1432	97.91%

V. CONCLUSIONS

This paper have shown that it is possible to automate the classification process of Brazil-nuts, based on the relation between the major and minor axis of each one, reaching a 97.63% and 84.09% of whole and broken Brazil-nuts respectively, with a threshold of 1.669 in images taken at a distance of $40cm$. In the case of the images taken at a distance of $30cm$. whose threshold was 2.03 and the accuracy was of 99.55% and 98.44% for whole and broken Brazil-nuts respectively.

Likewise, establish a stopping criteria based on the area of objects, without have a subjective number of iterations in the evolution function of the level set method; being one of its main advantages the flexibility to work with all kind of images with uniform objects; having a high performance in the segmentation process, rating an 94.05 percent.

REFERENCES

[1] L. A. Gómez, "Evaluación económica de la extracción de castaña (bertholletia excelsa h.b.k) , departamento de madre de dios," *Programa de Ordenamiento Ambiental*, vol. 1, 2006.

[2] J. Lopez, E. Aguilera, and M. Cobos, "Defect detection and classification in citrus using computer vision," in *Neural Information Processing*. Springer, 2009, pp. 11–18.

[3] D. Savakar and B. Anami, "Recognition and classification of food grains, fruits and flowers using machine vision," *International Journal of Food Engineering*, vol. 5, no. 4, p. 14, 2009.

[4] K. Kilic, I. Boyaci, H. Koksel, and I. Kusmenoglu, "A classification system for beans using computer vision system and artificial neural networks," *Journal of Food Engineering*, vol. 78, no. 3, pp. 897–904, 2007.

[5] D. Mery, J. Chanona Pérez, A. Soto, J. Aguilera, A. Cipriano, N. Veléz-Rivera, I. Arzate-Vázquez, and G. Gutiérrez-López, "Quality classification of corn tortillas using computer vision," *Journal of Food Engineering*, 2010.

[6] A. Blake, M. Isard *et al.*, *Active contours*. Springer London, 1998, vol. 8.

[7] A. Mansouri, T. Chomaud, and J. Konrad, "A comparative evaluation of algorithms for fast computation of level set pdes with applications to motion segmentation," in *Image Processing, 2001. Proceedings. 2001 International Conference on*, vol. 3. IEEE, 2001, pp. 636–639.

[8] S. Osher and J. Sethian, "Fronts propagating with curvature-dependent speed: algorithms based on hamilton-jacobi formulations," *Journal of computational physics*, vol. 79, no. 1, pp. 12–49, 1988.

[9] C. Li, C. Xu, C. Gui, and M. Fox, "Distance regularized level set evolution and its application to image segmentation," *Image Processing, IEEE Transactions on*, vol. 19, pp. 3243–3254, 2010.

[10] Y. Al Ohali, "Computer vision based date fruit grading system: Design and implementation," *Journal of King Saud University-Computer and Information Sciences*, vol. 23, pp. 29–36, 2011.

[11] N. A. Siriluk Sansomboonsuk, "The appropriate algorithms of image analysis for rice kernel quality evalution," in *The 20th Conference of Mechanical Engineering Network of Thailand*, 2006.

[12] J. Blasco, N. Aleixos, J. Gómez-Sanchís, and E. Moltó, "Recognition and classification of external skin damage in citrus fruits using multispectral data and morphological features," *Biosystems engineering*, vol. 103, pp. 137–145, 2009.

[13] I. Ahmad, J. Reid, M. Paulsen, and J. Sinclair, "Color classifier for symptomatic soybean seeds using image processing," *Plant disease*, vol. 83, pp. 320–327, 1999.

[14] M. G. Majid Rashidi, "Classification of fruit shape in kiwifruit using the analysis of geometrical attributes," *World Journal of Agricultural Sciences 3*, vol. 6, pp. 735–740, 2007.

[15] R. E. W. Rafael C. Gonzalez, *Digital Image Processing*. Prentice Hall, 2008.

[16] X. Xie, "Level set based segmentation using local feature distribution," *Pattern Recognition, International Conference on*, vol. 0, pp. 2780–2783, 2010.