COB-2019-1528
A FAST MULTILEVEL THRESHOLDING APPROACH FOR IMAGE SEGMENTATION USING GENETIC ALGORITHM AND AN APPLICATION IN CROP ROWS DETECTION

Pedro Ventura de Oliveira
Luciano Vieira Lima
Faculty of Electrical Engineering, Federal University of Uberlânda, Uberlândia – MG, Brazil
ventura.ufu@gmail.com, lucianovieiralimaster@gmail.com

Abstract. In this paper, an image segmentation method based on multilevel thresholding allowing a more efficient determination of the threshold points is proposed. Moreover, a practical application of this method for crop rows identification based on Progressive Probabilistic Hough Transform is presents, being an important initial step for the rows determination. The proposed thresholding combines a wavelet transform technique, for initially reduce the histogram length of an image converted to grayscale, with a genetic algorithm, responsible for finding the threshold points without an exhaustive search through all the possible values. The points, projected onto the original space of the histogram, are then set as the segmentation values. In addition, when used for crop rows detection, the method provide a refined threshold value, between the ones found in a multilevel segmentation, for the image binarization, which is an initial step in a protocol based on a Hough Transform method. Experiments using different images and thresholding approaches for comparative tests show the efficiency in speed and accuracy in the results for the proposed method. For the application in precision agriculture, the method show a better definition for the initial lines used to define the crop rows.

Keywords: image segmentation, genetic algorithm, thresholding, precision agriculture, crop rows detection

1. INTRODUCTION

Studies in the field of Precision Agriculture (PA) shows a relation between the increasing implementation of technologies associated with planting and farm profits by the application of these technologies (Schimmelpfennig, 2016). Imaging and mapping are important instruments on PA and largely applied in different tasks to obtain every kind of information that may influence the production of a plantation.

Nowadays agricultural market largely use Unmanned Aerial Vehicles (AUVSI, 2013) for mapping and imaging, being more efficient then satellites and airborne for imagery and orthophoto mosaics generation that may be useful in different application (Müllerova et al., 2016), such as in crop rows identification.

The Hough Transform (Hough, 1962) is used to find the location of linear objects (Sonka et al., 1993) being popular and practical. Methods based on this transform when properly combined with image segmentation and processing allow the identification of crop rows using the mentioned mosaics.

There are several algorithms proposed as solution for plantations rows detection that uses the Hough Transform (HT) as basis. Some of this methods supposes that the rows approximates by straight lines due to the way images are taken being the experiments conducted, e.g., with a camera mounted on a tractor (Søgaard and Olsen, 2003) or with images taken manually by a hand camera (Ji and Qi, 2011).

However, some plantation have a curved rows pattern following the natural geometry within the field. Being HT methods unable to perform a search for meaningful crop rows without identifying straight lines, algorithms based on it will mostly fail, when using images taken by a camera mounted on an UAV. Spite of that, before the properly use of a technique that can identify crop rows, the mosaic of the plantation must be prepared and processed.

Image segmentation is an important processing task for applications based on computer vision and further imaging analyses, making possible distinguish objects and regions of interest. There are different segmentation techniques, classified and categorized in different groups (Zaitoun and Aqel, 2015) englobing, e.g., k-means clustering, thresholding and split and merge methods.

The image thresholding is one of the most popular methods, used to separate a single object or more (depending if used for bi or multilevel) from the image background. There are several approaches for its application, such as Otsu's function (Otsu, 1979) and Kapur's function (Kapur, 1985) that gives a threshold point for a bi-level division, but that also can be extended for multilevel segmentation. However, in traditional methods such Otsu and Kapur, level extension brings

an increase of computation demand and, consequently, of execution time since they search the optimal thresholds calculating every possibility.

Therefore, Genetic Algorithms (GAs) turned to be a great and helpful strategy to solve image segmentation tasks based on multilevel thresholding with no need of an exhaustive search, due to its probabilistic characteristic, using these methods as objective functions and optimizing the process. The GAs were inspired by the Darwin's theory of evolution and are used for optimization and search problems constructing an algorithm that applies concepts of evolutionary operators such as population, selection, crossover, mutation and elitism (Bhandarkar and Hui Zhang, 1999).

Nevertheless, despite GAs being a good option making thresholding process faster, is still possible speed it up by previously applying a Discrete Wavelet Transform (DWT) to the histogram of the desired image to be processed. A Wavelet Transform is a tool that can be used to analyze non-stationary frequencies (Daubechies, 1990), useful in applications such as image processing, filtering and compression.

Similar to Fourier Transform, both decompose the original signal to analyze the data, however, Fourier Transforms adapts better with stationary frequencies, not handling efficient and accurately with, e.g., histograms of images.

Thus, the DWT gives to the GA a histogram with a reduced number of grayscale levels leading the algorithm to an even faster response with guarantee of good accuracy of the results.

This paper have two proposes: the first is to define a multilevel thresholding approach combining a GA with a DWT applied to the histogram of an image, reducing its grayscale levels, for computation, and establishing a link with the original histogram for projecting the threshold values onto the original space. Moreover, we address an initial problem of sugarcane plantation rows identification on images sensed by a mounted RGB camera on a UAV: the proper binarization and skeletonization of the mosaic sections, for further processing a HT-based method, by using the multithresholding GA developed.

The next section describes how the Discrete Wavelet Transform works, the proposed GA approach and its application for crop rows determination. The subsequent section presents the data of experiments and its results followed by the conclusions section.

2. METHODS

The first step of the proposed algorithm is to compute the histogram of an image converted to grayscale. The histogram is indicated by $h(i)$, $i = 0, 1, …, L$-1, where L represents the total gray levels of an image and $h(i)$, the number of pixel with the gray level i. Typically, $L = 256$, representing the full scale.

Furthermore, a DWT reduce the scale of the histogram before perform of the GA. A comparison is made between results of the proposed method and two other algorithms: one without a GA and other using the proposed GA but without a DWT.

Finally, for the crop rows detection, the multilevel thresholding approach shows its utility as initial step for a better identification based on a HT method.

2.1 Discrete Wavelet Transform

When looking for a threshold value the more traditional methods search through all values from 0 to L-1 until it is possible to determinate which one best divides the foreground from background. That is why the expansion of Otsu's, Kapur's and some other methods overly consumes resources: for t desired thresholds, there are $L \times (L$-$1) \times … \times (L$-$t)$ positions to be checked before giving the best threshold values.

Although applying a GA makes possible to reduce the consumed resources and time, due to the probabilistic nature of the algorithm, $L = 256$ still gives a high range of possible thresholds. Therefore, to accelerate even more the process without significant losses in response accuracy the Discrete Wavelet Transform presents to be a reliable option.

Before searching the thresholds, a histogram reduction is perfomed by means of a Wavelet Transform method (Kim et al., 2003). The transform decompose the signal of the original histogram in two at a lower level: the detail signal and the trend or approximation signal, which gives the reduced dimension version of the histogram and contain the overall characteristics of the original histogram.

The discrete wavelet transform at a level r is performed with decimation operation 2^r after the convolution of the histogram $h(i)$.

$$h^r(j) = WT^r\left[h(i)\right], \; r \in h^r(j) + h^r_w(j) \tag{1}$$

Where $h^r_w(j)$ represents the detailed signal and $h^r(j)$ the approximation signal at a rth level.

Each approximation signal can be reduced at r+1 also by decomposing into trend and detail signals. Hence, for each level r there is a reduction of the histogram length given by $L^r = L/2^r$.

The reduction from the 256 levels accelerate the search and convergence of the GA, which takes place in the next step of the procedure for the image segmentation.

2.2 Proposed Genetic Algorithm

The GA starts with a set of random chromosomes, which represents the initial population. Each chromosome is a possible solution – a set of threshold values for the image segmentation – and their quality is measure by how close of the desired solution value they are, i.e., the quality depends of the chromosome fitness.

The evolutionary operators are applied to the population, their fitness are recalculated and the best chromosomes are storage. Then, a new generation is formed and the process repeats until a stop condition is reached, which means, in this case, until a predetermined number of generations being reached.

The storage of the best chromosomes helps to assure the increase of solutions quality on each new generation, making the result to converge over time.

Table I shows the major steps of the approach to determinate the threshold values.

Table 1. Proposed Genetic Algorithm for Multilevel Thresholding

Main steps of the proposed GA
1. Convert the input RGB image for grayscale
2. Compute and normalize the histogram of the image
3. Use the Discrete Wavelet Transform to reduce the length of the histogram
4. Generate the initial population with a selected number of desired thresholds
5. Calculate the fitness and store a set of best chromosomes, separately (elitism)
6. Generate the next population by performing the evolutionary operators (selection, crossover and mutation)
7. Compare a matching number of the worst chromosome on the generation with the previous elite. The chromosomes with a better fitness replace the worst fitness on the new population.
8. If the desired number of generation is not reached, the new population go to step 5, repeating the sequence
9. Expand the best chromosome to the original space of the histogram
10. Use the threshold values for the image segmentation

2.2.1 Fitness

To allow determining the optimal threshold values, the chromosomes fitness are computed using as cost function the Otsu's between-class variance. The best individuals, i.e., the chromosomes with best set of thresholds based on its fitness, can be find by maximizing this Otsu's function (Otsu, 1979). For a multilevel problem, the between-class variance can be expressed as (Merzban and Elbayoumi, 2018):

$$\sigma_b^2(t) = \sum_{k=0}^{K} w_k \times (\mu_k - \mu_T)^2 \tag{2}$$

Where K represents $t + 1$ clusters and $k \in \{0,1,\ldots,K\}$.

For a cluster C_k, the mean μ_k is expressed in terms of the probability of a gray level i occurrence, $P(i)$, and its cumulative probability w_k:

$$\mu_k = \sum_{i \in C_k} \frac{i \times P(i)}{w_k} \tag{3}$$

Where:

$$w_k = \sum_{i \in C_k} P(i) \tag{4}$$

The total mean of the image is given by:

$$\mu_T = \sum_{i=0}^{L-1} i \times P(i) \tag{5}$$

2.2.2 Chromosome

For the GA's parameters the chromosome is represented by a binary string, with length L^r, where the one-bits indicates threshold points; the fitness is computed using the position of these one-bits.

The image's original grayscale histogram would give a chromosome with $L = 256$, however, with the DWT, there is a space reduction leading the histogram to a length L^r. Thus, the one-bits indicates the threshold points in this lower space, being randomly distributed accordantly to the desired number of thresholds when starting the initial population. For instance, a desired solution with four thresholds stablishes that chromosomes must have four one-bit values in the string, filling the other positions with zero-bits, and every operation with this chromosomes must guarantee its integrity in length and number of thresholds.
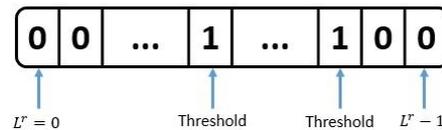

Figure 1. Chromosome representation. The one-bits represents the threshold positions. In the reduced histogram space, L^r represents the length of the chromosome.

2.2.3 Selection

After creating the initial population, the calculation of fitness for each individual allows the selection process to choose, by using a given strategy, the chromosomes for the crossover and mutation operators.

The tournament is the strategy applied on the proposed algorithm and guaranties a fair selection and variability of population selecting one between two random chromosomes from the current population based on chance. A winning rate is defined and, if a random number gives a result within this rate, the chromosome with better fitness is selected; otherwise, the one with worst fitness is selected.

The procedure repeats until a matching number of chromosomes with the same size of the original population is reached.

2.2.4 Crossover

After the tournament selection, a pair of individuals, the parent chromosomes, from this new population are selected and each threshold gene – one-bit element from the string – submitted over the crossover rate. If the chromosomes exchange a gene, another random threshold gene recombines in a way to ensure the maintenance of the chromosome integrity having always the same amount of threshold values.

In a simple manner, if "parent A" exchange a one-bit with a zero-bit from "parent B", in a given position, "parent B" randomly selects a one-bit from a different position and recombines it with a zero-bit from "parent A".

The offspring generated on this process goes through the mutation operator individually.
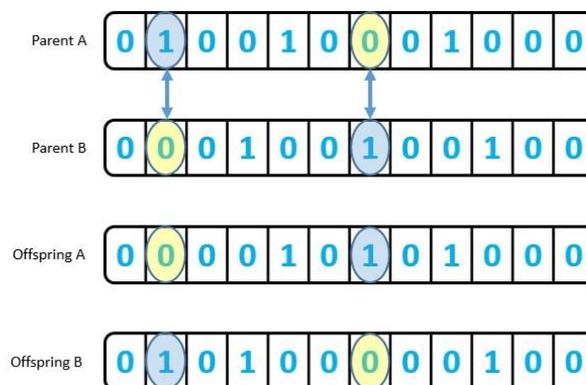

Figure 2. Crossover process.

2.2.5   Mutation

The mutation process scans the entire chromosome and ensure the initial number of thresholds by swaping the positions of a randomly chosen threshold when the mutation probability value fall within the mutation ratio. If a zero-bit gene mutates it becomes a one-bit, then, a random one-bit becomes a zero-bit and vice-versa.
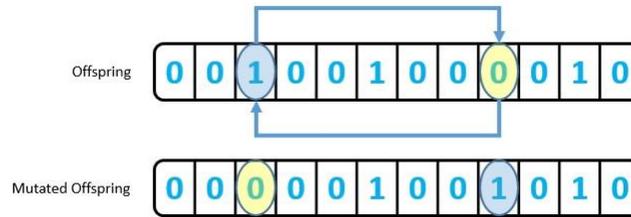

Figure 3. Mutation process.

For guarantee that good solutions are not lost in next generations, where a new set of chromosomes creates a new population, the elitism is performed taking a small group composed with the best solutions and carrying them on to the next generation by replacing the worst offspring.

2.2.6   Expantion of Thresholds to Original Space

The optimal thresholds t_i^r founded on the reduced histogram, i.e., $t_i^r \in [0, L^r]$, must be expanded to express their values in the original space. Each threshold t^i is multiplied by a factor 2^r, so that (Kim et al., 2003):

$$t_i = t_i^r \times 2^r, \text{ for } i = 1, 2, ..., k \text{ and } t_i \in \left[0, L-1\right] \tag{6}$$

Now, the real threshold values applied to the image can perform the segmentation and extracts the required information for a desired application.

2.3 Crop Rows Detection

The input data for this work are composed by images of sugarcane plantation captured by an UAV mounted RGB camera on an average height of 100 meters from the ground. These images afterwards composed a single image by means of an imaging mosaic technique.

To deal with the curved rows problem and still using a HT method, the mosaic is subdivided in small sections in which the segments within can be approximated by straight lines. Then, an approach subjecting each section to the transform, after a preprocessing procedure, is applied.

The preprocessing step consist of two parts: the image binarization, separating the desired pixels from the others, i.e., separating plants from soil and extraneous elements; and a skeletonization process, evidencing the plants and the crop rows direction.

For a plantation problem, a binarization process that simply performs a bi-level thresholding can lead to a bad threshold point for distinguish soil from plants, taking some mid-term pixels that should be backgrounded (e.g. leaves and shadows) to the plants portion of segmentation, constituting an unwanted noise. Therefore, a multilevel thresholding method, by dividing an image in different point, can simplify the extraction of the desired pixels that had better represents the vegetation through the selection of one of the threshold values, more accurate for the application needs.

The implementation of the presented approach, besides giving a fast segmentation for all sections, also allow to better determinate a point for dividing the plants from undesirable elements.

After the GA using DWT be performed over a section, the higher threshold value from the best chromosome is set as the binarization point. High thresholds values, represents grayscales closest to white which are, for the mosaic used, the scales that better represent the plants portions.

With the image binarized the skeletonization process takes place to the final preprocessing step. Figure 4-c shows the grayscale representation from one of the mosaic's sections (Fig. 4-b) after the binarization process. Hereafter, by means morphologic operations, the skeletonization of the image is performed being presented in Fig. 4-f.

These morphologic operations uses structuring elements that translate over the desired binary image to, e.g., enlarge a region or make it smaller. These structuring elements, normally with a few pixels dimension, act similar to probes of the binary image (Shapiro and Stockman, 2000). The skeletonization gives a one pixel thick image, which means that the pixels identified as plants are reduced to a simplified form with a minimal thickness (a skeleton form).
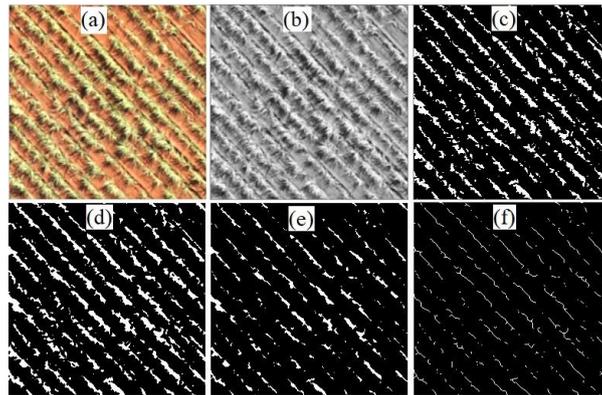
Figure 4. Algorithm steps: (a) Section from sugarcane plantation; (b) graycale converted image; (c) binarization using GA's higher threshold value; (d) morphological closing; (e) morphological erosion; (f) skeletonization.

First, for reducing possible internal holes and similar structures on the rows regions that can distort the results generating ramifications or gaps in the actual rows, a morphological closing is performed. The next step is reducing pixels around possible rows regions giving more regularity and separating these rows by a second operation called erosion.

The closing operation starts by dilating the foreground object and then eroding it, regressing the object to its original size, but with less holes and possible gaps that should not exist, in first place. The second erosion reduce the thickness and eliminate some noises letting the image prepared for the skeletonization.

Theses previous operations are presented, respectively, on Fig. 4-d and Fig. 4-e and the skeletonization operation is presented on Fig. 4-f.

After preprocessing an image, a HT method takes place defining straight lines over the skeleton produced. For this step is performed the Progressive Probabilistic Hough Transform (PPHT), a probabilistic HT technique, capable of detecting lines with less computation and improved performance, when compared with the standard HT (Matas et al., 2000).

The PPHT applied over the skeleton gives the desired points for the crop lines designing over the original image.

3. EXPERIMENTS AND RESULTS

The multilevel thresholding GA was implemented on Python and make use of some functions from the OpenCV library.

Two well-known images on grayscale were used: *Lena* (512x512) and *Peppers* (512x512); and they had a three levels dimension reduction ($r = 3$) of their histogram length, dropping from $L=256$ gray levels to $L^3=32$ ($L/2^3$), using the Daubechies wavelet "db1".

The total population, P, depends of the chromosome length, L, and reduction level. Considering this, $P = 40$ and the stop condition is equal to the number of generation N where, for both images, $N = 100$.

Crossover and mutation rates are 70% and 5%, respectively, with a 5% elitism and tournament win rate of 70%.

For a comparative analysis, the traditional extended Otsu's method (exhaustive search over all combinations) and the same proposed GA without the DWT were tested. Due to the higher complexity of the chromosomes for these two methods ($L=256$), for achieve a better convergence of results a higher number of individuals is necessary, being $P=100$.

Figure 5. Test images in grayscale: Lenna (left) and Peppers (right).

Table 2 shows the performance for each method applied on the two images for 3 and 4 thresholds.

Table 2. Comparative results for *Lena* and Peppers images applying 3 and 4 thresholds.

Image	Method	Threshold values	CPU time (s)	Threshold values	CPU time (s)
Lena	Otsu	79 – 124 – 170	448.988	75 – 115 – 145 – 175	>>> 600.000
	GA without DWT	79 – 125 – 170	16.258	75 – 113 – 145 – 180	18.081
	GA + DWT (proposed method)	80 – 128 – 168	3.174	72 – 112 – 144 – 176	3.871
Peppers	Otsu	63 – 119 – 168	398.423	56 – 90 – 128 – 170	>>> 600.000
	GA without DWT	64 – 120 – 168	15.798	50 – 89 – 129 – 171	16.796
	GA + DWT (proposed method)	64 – 120 – 168	3.115	56 – 96 – 128 – 168	3.650

The proposed method shows good accuracy for the threshold values found with great reduction of time consumption, even comparing with the GA without DWT.

For the application on crop rows identification, the proposed method is applied using 4 thresholds before the binarization, that uses the higher threshold value found by the GA as its threshold point. From the usage of the PPHT, Fig. 6 shows the skeleton and a section of the mosaic marking the initial – and not refined – lines in two cases: by previously applying the developed method and by just binarazing the image using Otsu's method.
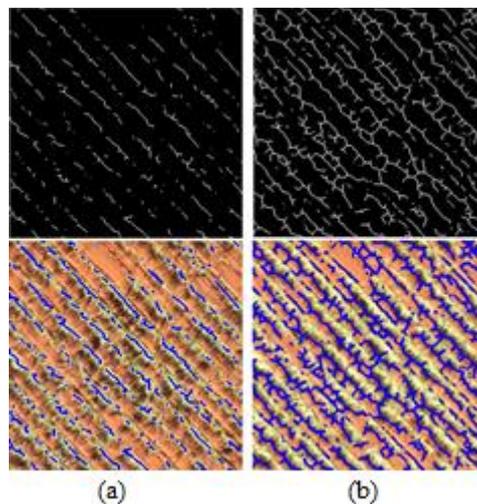

Figure 6. Skeleton and lines from PPHT applied on the same image. The left images (a) used the threshold given by the proposed multithreshold method and the right images (b) performed a direct Otsu's binarization.

When performing binarization using Otsu's method directly, a great number of pixels considered as noise (or background) are set as plants. The closing and erosion operation are inefficient to separate the rows and, consequently, the skeleton produced gives a poor match with the actual plants segments, as shown in Fig. 6-b.

From the multilevel thresholding, the higher threshold value is 200 while Otsu's method set a value of 148. The pixels have a higher density around the grayscales 125~180 with no relevant distance between the peaks of the gray levels that represents plants and peaks of the other elements. The bi-level method is not capable of properly represents a feature when its pixels close together on a small region with other pixels from background, being why results, in this case, have poor accuracy. However, a multilevel variant have a better opportunity to distinguish objects in this situation.

By using the cluster of grayscales from 200 up to 255, the algorithm is using the refined pixels that had better represent plants. Taking values down to 148, soil, some shadows, etc. are classified as plants, inducing the PPHT to a bad tracking of lines that could constitute a crop row.

Although being just an initial draw of the lines, Fig. 6-a gives a better representation of the crop rows by setting a different threshold point. Figure 7 shows the values chosen by thresholding the image histogram by the two methods in discusion.

Pedro Ventura de Oliveira and Luciano Vieira Lima
A Fast Multilevel Thresholding Approach for Image Segmentation Using Genetic Algorithm and an Application in Crop Rows Detection
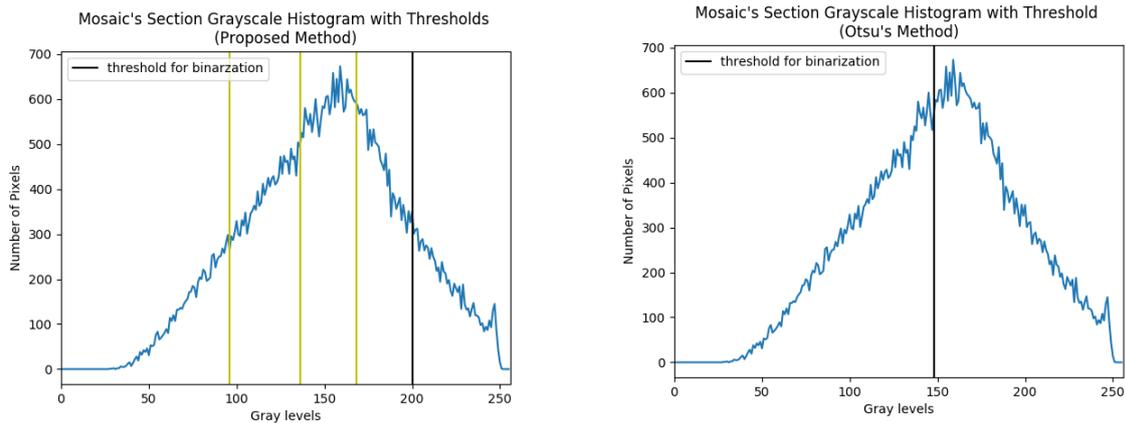
Figure 7. Grayscale histograms showing the thresholds values for the proposed method and for the Otsu's method. The proposed method set the gray level 200 as threshold value for binarization while Otsu sets the level 148.

4. CONCLUSIONS

In this paper, an approach of a fast multilevel thresholding for image segmentation is presented. Using a DWT it is possible to reduce the histogram length of a grayscale image, being capable to restore the original signal resolution and, in the lower resolution level, calculate the thresholds values. The approach combines a GA with the DWT speeding up the processing task of the optimal thresholds. The proposed method directly provides the segmentation values due its chromosome representation where each bit indicates a histogram position. After finding the result at the reduced histogram space, they can easily be expanded to the original space giving the real threshold values.

Comparing the pure application of the developed GA with the whole approach used, results shown that even introducing additional steps with the decomposition of the histogram and the later expansion of the results, the proposed method proved to be faster due to the less complex analysis of the chromosomes on every step of the GA. Furthermore, the approach using the GA with a wavelet transform proved to be efficient by giving good accuracy in comparison to Otsu's method.

Between many applications of a multilevel threshold-based algorithm, the approach proved useful on an initial step for a method of crop rows detection, representing its relevance by determining a best threshold value for binarizing a plantation image before applying a HT-based algorithm.

The algorithm provides a threshold value that can better identify pixels that represents plants working as a refinement of the binarization point and more properly giving the crops positions for the desired application.

Although it is not the focus of this article, the GA presented may be useful for other PA applications (e.g. creating regions of interest, objects or defect fast detection, etc.) and this will be better explored in future works.

5. ACKNOWLEDGEMENTS

The authors kindly thanks the company Sensix (http://sensix.com.br/) for providing the mosaic from a sugarcane plantation, used in tests for the development of this paper.

6. REFERENCES

AUVSI. (2013). The economic impact of unmanned aircraft systems integration in the united states march. Economic report, Association for Unmanned Vehicle Systems International, Washington DC. 22 Mar. 2019 < https://www.auvsi.org/our-impact/economic-report>

Bhandarkar, S., & Zhang, H. (1999). Image segmentation using evolutionary computation. IEEE Transactions on Evolutionary Computation, 3(1), 1-21. doi:10.1109/4235.752917

Daubechies, I. (1990). The wavelet transform, time-frequency localization and signal analysis. IEEE Transactions on Information Theory, 36(5), 961-1005. doi:10.1109/18.57199

Hough, P.V.C. (1962). METHOD AND MEANS FOR RECOGNIZING COMPLEX PATTERNS. United States Patent 3.069.654.

Ji, R., & Qi, L. (2011). Crop-row detection algorithm based on Random Hough Transformation. Mathematical and Computer Modelling, 54(3-4), 1016-1020. doi:10.1016/j.mcm.2010.11.030

Kapur, J., Sahoo, P., & Wong, A. (1985). A new method for gray-level picture thresholding using the entropy of the histogram. Computer Vision, Graphics, and Image Processing, 29(1), 140. doi:10.1016/s0734-189x(85)90156-2

Kim, B., Shim, J., & Park, D. (2003). Fast image segmentation based on multi-resolution analysis and wavelets. *Pattern Recognition Letters,24*(16), 2995-3006.

Matas, J., Galambos, C., & Kittler, J. (2000). Robust Detection of Lines Using the Progressive Probabilistic Hough Transform. Computer Vision and Image Understanding, 78(1), 119-137. doi:10.1006/cviu.1999.0831

Merzban, M. H., & Elbayoumi, M. (2019). Efficient solution of Otsu multilevel image thresholding: A comparative study. Expert Systems with Applications, 116, 299-309. doi:10.1016/j.eswa.2018.09.008

Müllerová, J., Brůna, J., Dvořák, P., Bartaloš, T., & Vítková, M. (2016). Does The Data Resolution/origin Matter? Satellite, Airborne And Uav Imagery To Tackle Plant Invasions. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences,XLI-B7*, 903-908.

Otsu, N. (1979). A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics,9*(1), 62-66.

Schimmelpfennig, D. (2016) Farm Profits and Adoption of Precision Agriculture, ERR-217, U.S. Department of Agriculture, Economic Research Service, < www.ers.usda.gov/publications/err-economic-research-report/err217>

Shapiro, L. G., & Stockman, G. C. (2001). Computer vision. Upper Saddle River, NJ: Prentice Hall.

Søgaard, H., & Olsen, H. (2003). Determination of crop rows by image analysis without segmentation. Computers and Electronics in Agriculture, 38(2), 141-158. doi:10.1016/s0168-1699(02)00140-0

Sonka, M., Hlavac, V., & Boyle, R. (1993). Image processing, analysis, and machine vision. London: Chapman & Hall Computing.

Zaitoun, N. M., & Aqel, M. J. (2015). Survey on Image Segmentation Techniques. *Procedia Computer Science,65*, 797-806. doi:10.1016/j.procs.2015.09.027

7. RESPONSIBILITY NOTICE

The authors are the only responsible for the printed material included in this paper.