

## Novel Image Segmentation Using Gaussian Mixture Models -- Application to Plant Phenotypic Analysis

Jinhai Cai, Mahmood Golzarian

Phenomics and Bioinformatics Research Centre  
School of Mathematics and Statistics  
University of South Australia  
Mawson Lakes, Adelaide, Australia  
Jinhai.Cai@unisa.edu.au  
Mahmood.golzarian@unisa.edu.au

Stan Miklavcic

Phenomics and Bioinformatics Research Centre  
School of Mathematics and Statistics  
University of South Australia  
Mawson Lakes, Adelaide, Australia  
Stan.Miklavcic@unisa.edu.au

**Abstract** — A novel algorithm is proposed for background estimation using statistical pattern recognition. Usually the segmentation of objects in images is achieved by identifying homogeneous regions in individual images or by finding motions of objects in videos. In this paper, we combine the advantages of these approaches for the estimation of background using only two images. The proposed algorithm uses the difference between images to obtain initial estimation of background and then to refine the estimation using statistical pattern recognition. Experimental results have shown that the proposed algorithm can achieve promising performance in terms of accuracy and speed.

**Keywords** - background estimation; object segmentation; Gaussian mixture models; statistical pattern recognition

### I. INTRODUCTION

Phenomics and bioinformatics are two indispensable aspects of molecular biology and genetics. Applied to plant biology, automated phenotypic analysis of plant images, captured as a function of growth conditions, can help us to obtain a large amount of information on the function of genes and to study the impacts of abiotic stress on plants. In order to do this, it is essential to automatically separate plants from the background in these images.

Background estimation for image segmentation has been an important research topic in image processing and computer vision for several decades. A wide range of computational vision problems could in principle make good use of segmented images [1], as evidenced by a range of applications from surveillance [2], plant physiology [3], food engineering [4] to medical science [5]. As a result of this extensive research, many different approaches and algorithms for image segmentation have been developed [6]. However, objects and backgrounds in the real world can be very complicated in terms of colors, textures and shapes. Therefore, image segmentation remains one of the most challenging topics in image processing and computer vision.

In general, algorithms for image segmentation can be divided into two categories: segmentation based on feature similarity [7] and segmentation based on motion [6]. Usually, algorithms for segmenting individual images belong to the first category. Most algorithms in this category focus

on modeling the consistent and similar features of an object for segmentation [7]. However, an object may have several distinct features within its boundary, which then become segmented into several regions. To solve this problem, some algorithms obtain models of objects by training, which are then applied to segment similar images. As these algorithms can only be applied to images containing objects similar to the training samples, the application of these algorithms is usually limited. Other algorithms require the human knowledge input via interactive schemes [8] for object segmentation in images. As a result, these algorithms can achieve impressive performance, but they cannot be used in a fully automatic manner. Usually, different objects have different features, which are most prominent at the boundaries between different objects. Therefore, it is logical to segment objects using boundary information. Level set methods, which focus on object boundaries, have been successfully applied in image segmentation, where the features of focus can be colors, textures, and shapes [6]. However, level set methods are sensitive to initialization and parameter selection.

In the second category, algorithms are based on motions, where motions can be estimated from multiple images. In this category, some algorithms directly segment objects in the optical flow space [6]. Others focus on the background estimation [9]. Segmentation in the optical flow space needs only two images, but significant errors may occur at object boundaries. Segmentation based on background subtraction performs well at object boundaries, but it is based on the assumption that all parts of the background will be unveiled at some time. Thus, background estimation usually needs a sequence of images.

In the application to plant phenomics, plants do not move. With their growth, more and more areas of the background will be covered by plants. Therefore, it is pointless to use images of plants in their late growth stages for background estimation. It is desirable for an automated algorithm to use only a few images of plants in their earliest growth stage to estimate the background. In this paper, we propose a method to estimate the background from two images using Gaussian mixture models and pattern recognition.

## II. THE DIFFERENCE IMAGE

The Australian Plant Phenomics Facility (APPF) provides state-of-the-art plant growth environments and the latest technology in high throughput plant imaging for the automatic and non-destructive phenotypic experiments of plants. It can take thousands of plant images by different cameras from different angles, two orthogonal side views and a top view, in one day. Given the amount of data collected, it is desirable to be able to process these images automatically.

### A. Lens Distortion Correction

In the high throughput plant imaging facility, we have four cameras for fluorescent, near-infrared, infrared and visible spectral images, respectively. The degree of lens distortion of each camera is different. Therefore, it is necessary to correct for lens distortion before integrating these images for plant phenotype analysis. The idea behind the lens distortion correction is to convert curves in distorted images into straight lines in undistorted images. We used an algebraic approach [10] for distortion correction, and example results of which are given in Fig.1.

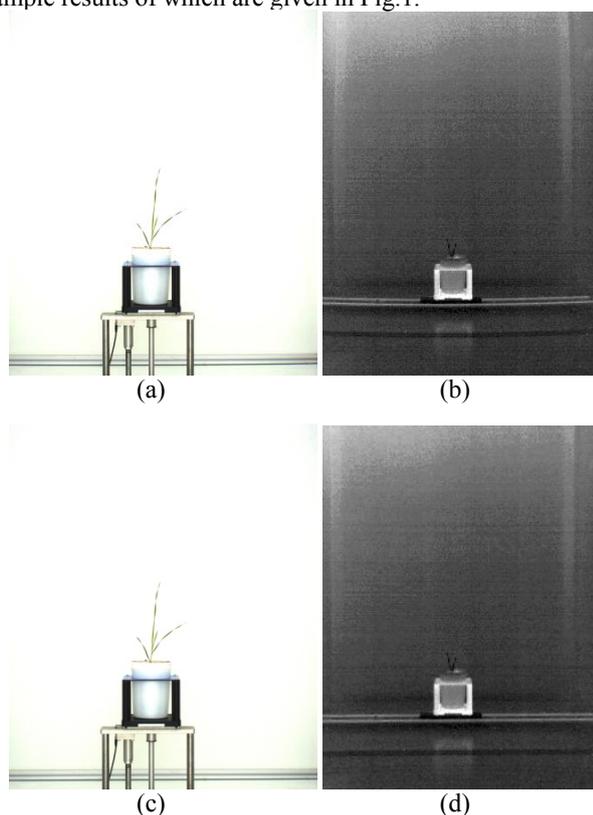


Figure 1. The correction of lens distortions. (a) & (b) original images; (c) & (d) images after the lens distortion correction.

### B. The Difference Image

It is well known that the performance of segmentation algorithms using individual images is unsatisfactory in many applications. Therefore, it is natural to use multiple images to improve the performance of segmentation. For example,

stereo algorithms use two rectified images to calculate the disparity of each pixel, and objects in images can be segmented based on their disparities as shown in Fig.2. The

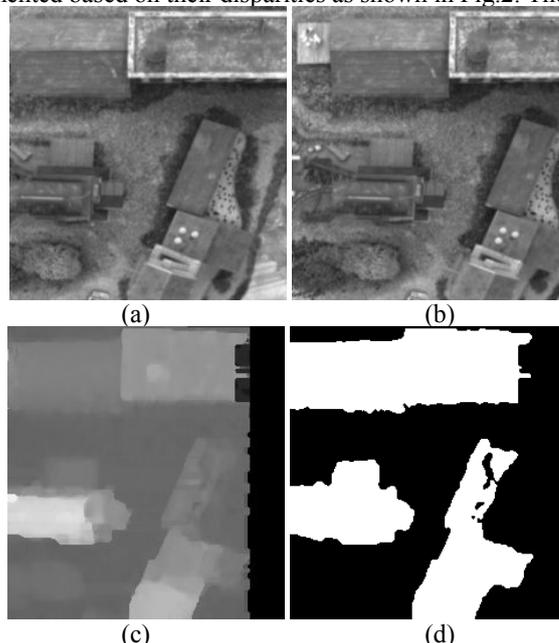


Figure 2. Segmentation in stereo vision. (a) & (b) original stereo images; (c) the disparity image; (d) the result of building segmentation.

major advantage of motion based segmentation is that it is robust to shades and other interferences. In our application, different plants are imaged against the same background, and the same plant can change its appearance due to its growth. Therefore, it is very difficult to estimate the optical flow, but it is easy to find the difference between two images. Inspired by motion based segmentation, we propose to use the difference image for background estimation and image segmentation.

In the dataset available, each plant was imaged every second day by cameras fixed in stations. During the imaging, it is inevitable to have some vibrations. As a consequence, there are some displacements in the images. In order to remove the effects of camera vibrations, the optical flow is used as it is the fastest method to find matched feature point pairs and its performance is very good if the camera's rotation is small. Let  $\{x_i, y_i\}$  denote the position of a point in the reference image and it matches a point in another image, where its location is  $\{x'_i, y'_i\}$ . We can estimate the affine transformation

$$\begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x'_i \\ y'_i \\ 1 \end{bmatrix} \quad (1)$$

where  $m_{11}, m_{12}, \dots, m_{23}$  are elements of the transformation matrix. In this way, we align all images to the reference image. Fig.3(a) shows the difference image without lens distortion correction or alignment. Clearly,

differences in some regions, particularly at object boundaries, are salient. In contrast, differences in calibrated images as shown in Fig.3(b) are much small at object boundaries. This is mainly due to the fact that calibration removes the effects of camera vibrations. Fig.3(c) shows the whole difference image.

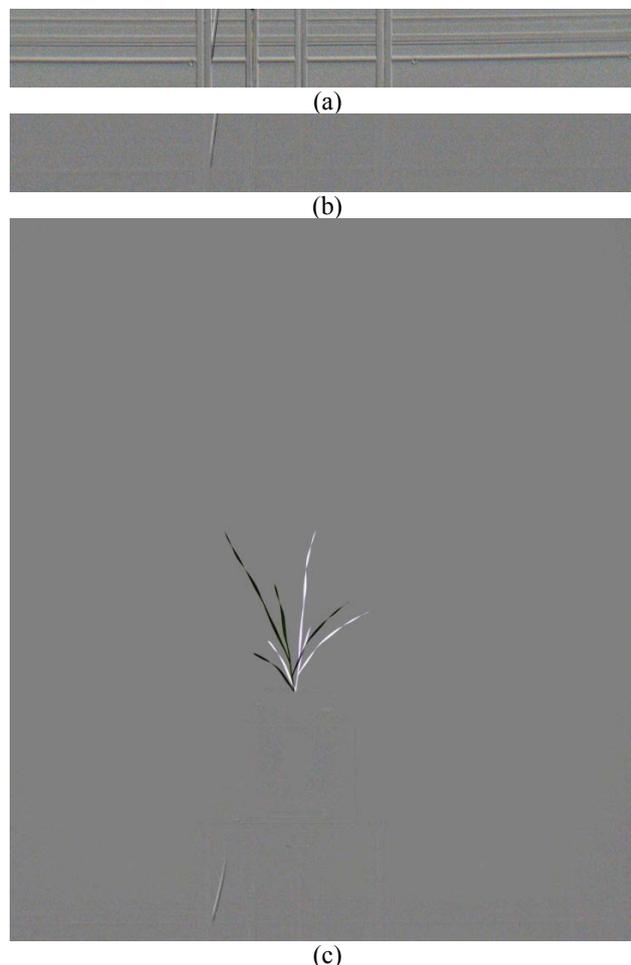


Figure 3. Difference images: (a) a section of the difference image without calibration; (b) a section of the difference image with calibration; (c) the whole difference image.

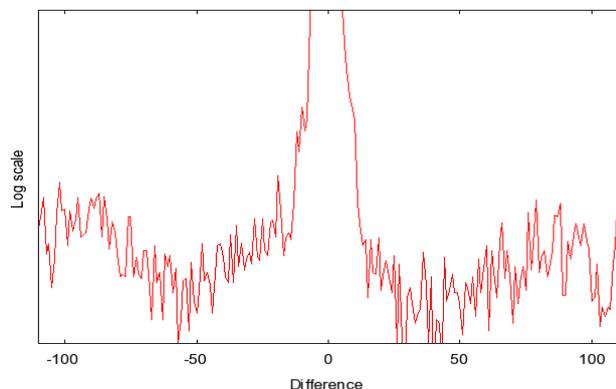


Figure 4. The histogram of the difference image.

### III. LEARNING THE BACKGROUND AND FOREGROUND

From Fig.3(c) we find that the difference image can be divided into three distinct parts. In the first part, the differences are close to zero. Since we know that the backgrounds of the two images are almost identical, the difference between them is very small after the calibration. Consequently, we can conclude that the region with very small image difference is likely to be the background. This part appears as the medium-gray region in the difference image. In the second part, the differences are negative. The plant appearing in the reference image will likely appear in the negative part of the difference image. In the third part, the differences are positive. The plant appearing in the second image will likely appear in the positive part of the difference image. The negative and positive parts are represented by the dark gray regions and light gray regions, respectively, in the difference image. The character of three parts in the difference image can easily be quantified by histograms as shown in Fig.4, where the histogram of the difference image has three blocks and each block corresponds to one part in the difference image.

#### A. Gaussian Mixture Models for the Difference Image

Gaussian mixture models (GMMs) have been applied successfully to approximate density functions in many applications. We will use GMMs for the separation of the foreground from the background. In the difference image, the feature vector of a pixel is  $\chi = \{\chi_j\}$ , where  $\chi_j$  is the feature in the  $j$ th channel where  $j = 1, 2$  or  $3$ . So, we have a density function  $p(\chi | \vartheta)$  that is governed by the set of parameters  $\vartheta$ . We assume that each channel is independent, so we have

$$p(\chi | \vartheta) = \prod_{j=1}^N p(\chi_j | \vartheta_j), \quad (2)$$

where  $\vartheta_j$  is the set of parameters for channel  $j$  and  $N$  is the number of channels. For each channel, there are three distinct parts in the difference image as shown in Fig.3(c) and Fig.4. Accordingly, we assume that the density function for each channel is in the form of the Gaussian mixture model and the number of mixtures is 3. We have

$$p(\chi_j | \vartheta_j) = \sum_{i=1}^M \alpha_i p_i(\chi_j | \vartheta_{ji}) \quad (3)$$

where  $M=3$ ,  $\sum_{i=1}^M \alpha_i = 1$ ,  $p_i$  is a Gaussian function and  $\vartheta_{ji}$  includes the mean and the variance. The parameters,  $\alpha_i$  and  $\vartheta_{ji}$  can be estimated by using the Expectation-Maximization (EM) algorithm.

#### B. Learn the background and foreground

As each Gaussian function corresponds to one part in the difference image, we can use Bayes's rule to separate each part in the difference image:

$$\Gamma(\mathcal{X} | C_i) = \prod_{j=1}^N \alpha_i p_i(\mathcal{X}_j | \vartheta_{ji}), \quad (4)$$

where  $\Gamma(\mathcal{X} | C_i)$  is the likelihood of the feature  $\mathcal{X}$  belonging to class  $C_i$ , and  $C_i$  is a class label of the background part, plant part in the reference image and plant part in another image. The classification can be performed by

$$C = \arg \max_i \Gamma(\mathcal{X} | C_i).$$

In this way, we can obtain the initial segmented plants and the background.

There is some chance that plant regions in one image overlap with these in another image. Since plants share similar colors, the differences of two images in these overlap regions are small. Consequently, the overlapped plant regions might be classified as the background if we used the difference image alone. To avoid this problem, we utilize the two original images as well as the difference image for the learning of background and foreground. As the total area of overlap is very small, its impact on the background estimation is negligible. Therefore, we need only the initially segmented regions in the original images to learn the background using the Gaussian mixture model. Similarly, we can use the initially segmented plant regions in the original images to learn the foreground.

For the estimation of background from two images, there is no information on the background in regions blocked by plants in both images. These blocked regions cause the same problem in the estimation of digital terrain models (DTMs) [11], where the invisible parts are approximated by neighboring visible regions. We use the same approach to approximate the background in regions of plant overlap.

#### IV. EXPERIMENTAL RESULTS

Background estimation and plant segmentation are the first step in a phenotyping process. They are essential components in high-throughput image analyses. To evaluate the performance of the proposed algorithm, images from the APPF Plant Accelerator<sup>®</sup>, which is equipped with LemnaTec's<sup>®</sup> cutting edge imaging and conveyor belt technologies, were used in our experiments.

In our experiments, we estimated the background from two images as shown in Fig.6(a) and (b). The estimated background is given in Fig.6(c). The overall performance of the algorithm is excellent with the exception of a few minor errors at edges of leaves as shown in Fig.6(d). It is inevitable that there will be some errors in the estimation as there is no clear-cut boundary between the background and plants in the histogram as shown in Fig.4. Fortunately, most of these minor errors occur in strips of one pixel width, thus they can be removed easily by a median filter.

Once we have the estimated background, we can perform the segmentation of plants easily. Fig.7(a) and (b) show the segmented plants from Fig.6(a) and (b), respectively. The algorithm performs very well in most cases as shown in Fig.7(c) and (d). However, there are some minor errors in plant segmentation as shown in Fig.7(e), where a small area of the pot rim covered by dust has been classified as part of

the plant. However, in the post-processing stage, leaf detection and tip detection should be able to remove these errors.

It is interesting to compare the performance of plant segmentation using the proposed algorithm to that of level-set methods. We first manually cropped a rectangular region around a plant out its whole image as shown in Fig.8(a). We evaluated several level-set methods using this image. We

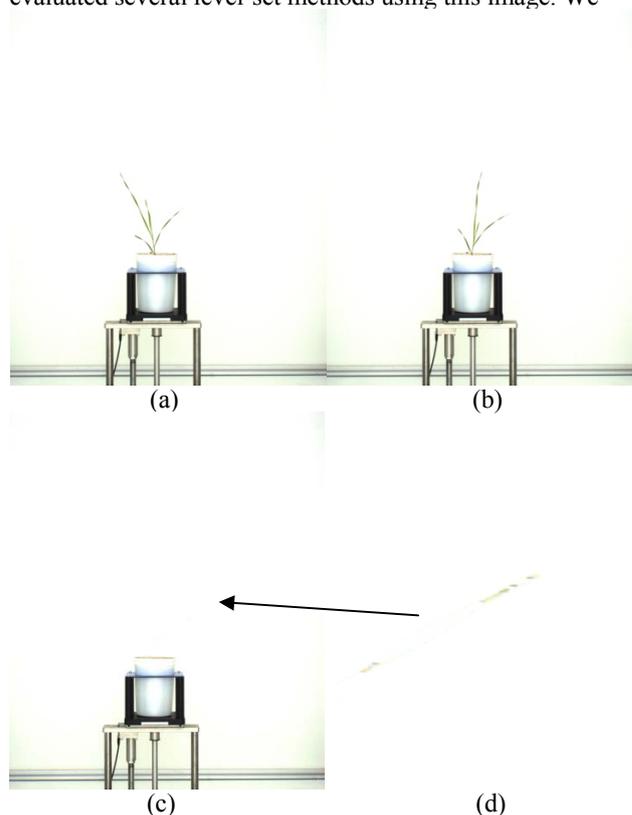
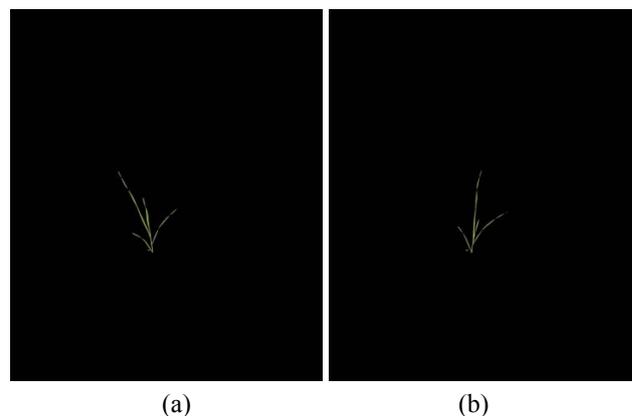


Figure 6. The background estimation: (a) and (b) original images after the correcting for lens distortion; (c) the estimated background; and (d) the magnified area in the background image (c), which shows some minor errors in background estimation.



(a)

(b)

Continue ... ..

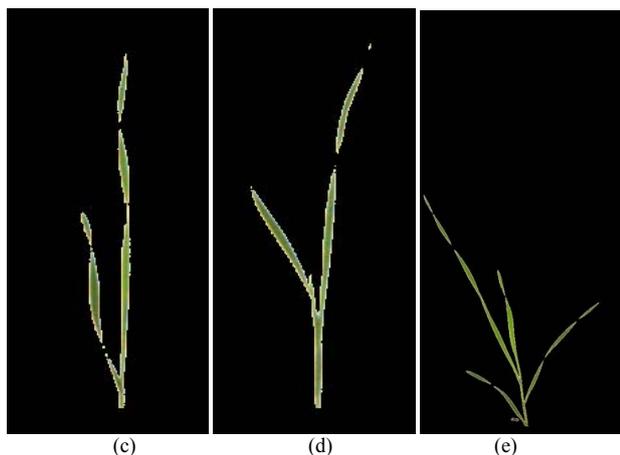


Figure 7. The segmentation of plants: (a) and (b) are plants segmented from Fig.6(a) and (b), respectively; (c) and (d) are two other plants segmented by the proposed method. (e) is the same plant in (a) but magnified to show the error at the bottom of the image

found that [12] and [13] can achieve good results as shown in Fig.8(b) and (c), respectively. The result of [13] on this plant is very similar to the result shown in Fig.7(d). It is also interesting to apply different approaches to whole images for evaluation. However, [12] and [13] do not perform well when whole images are used for segmentation [14]. Although the initial contours cover the plant only, the evolved contours extend to whole images as shown in Fig.8(d) and (e).

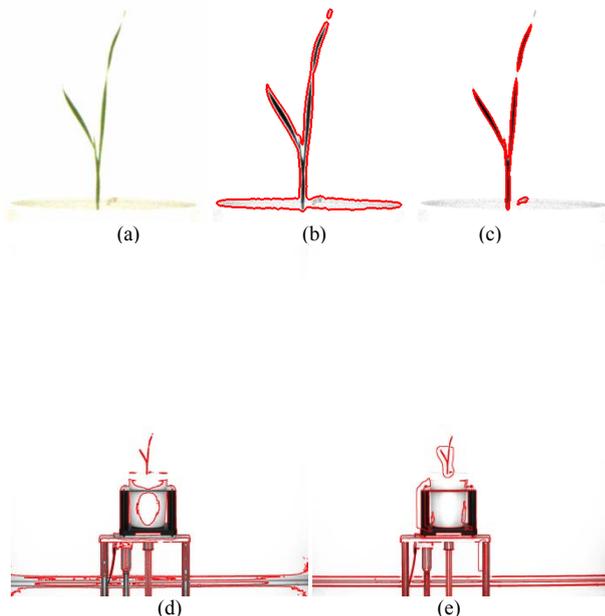


Figure 8. Level-set segmentation. (a) the plant in the original image; (b) the result by [12]; (c) the result by [13]; (d) the result of the whole image by [12]; (e) the result of the whole image by [13].

## V. CONCLUSIONS AND DISCUSSIONS

We present a novel method for background estimation and image segmentation using two images only. The

proposed method uses the difference image for the initial segmentation and uses GMMs for further refinement. The method has been evaluated on a plant image dataset and promising performance has been achieved.

We have conducted a simple performance comparison between the proposed method and some level-set methods. In the near future, we will investigate the possibility of integrating our region-based method with edge-based methods, such as the level-set method, for image segmentation.

## ACKNOWLEDGMENT

The authors would like to thank the APPF Plant Accelerator for providing them with the image data for this research.

## REFERENCES

- [1] Pedro F. Felzenszwalb, and Daniel P. Huttenlocher, "Efficient Graph-Based Image Segmentation," *Int. J. Comput. Vision*, vol.59, no.2, pp.167-181, 2004.
- [2] J.M. Ferryman, S.J. Maybank, and A.D. Worrall, "Visual Surveillance for Moving Vehicles," *Int. J. Comput. Vision*, vol.37, no.2, pp.187-197, 2000.
- [3] B. Berger, B. Parent, and M. Tester, "High-throughput shoot imaging to study drought responses," *Journal of Experimental Botany*, vol.61, no.13, pp.3519-3528, 2010.
- [4] D. Mery and F. Pedreschib, "Segmentation of colour food images using a robust algorithm," *Journal of Food Engineering*, vol.66, no.3, pp.353-360, 2005.
- [5] L. He, Z. Peng, B. Everding, X. Wang, C.Y. Han, K.L. Weiss and W.G. Wee, "A comparative study of deformable contour methods on medical image segmentation," *Image and Vision Computing*, vol.26, no.2, pp.141-163, 2008.
- [6] D. Cremers, M. Rousson, and R. Deriche, "A Review of Statistical Approaches to Level Set Segmentation: Integrating Color, Texture, Motion and Shape," *International Journal of Computer Vision*, vol.72, no.2, pp. 195-215, 2007.
- [7] J. Chen, T.N. Pappas, A. Mojsilovic, B.E. Rogowitz, "Adaptive perceptual color-texture image segmentation," *IEEE Trans. on Image Processing*, vol.14, no.10, pp.1524-1536, 2005.
- [8] J. Ning, L. Zhang, D. Zhang, and C. Wu, "Interactive image segmentation by maximal similarity based region merging," *Pattern Recognition*, vol. 43, no. 2, pp.445-456, 2010.
- [9] V. Reddy, C. Sanderson, B.C. Lovell, "An efficient and robust sequential algorithm for background estimation in video surveillance," *Proc. Of IEEE International Conference on Image Processing*, pp.1109-1112, 2009.
- [10] L. Alvarez, L. Gomez, and J.R. Sendra, "An algebraic approach to lens distortion by line rectification," *Journal of Mathematical Imaging and Vision*, vol.35, no.1, pp.36-50, 2009.
- [11] H.S. Lee and N.H. Younan, "DTM extraction of Lidar returns via adaptive processing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 9, pp.2063-2069, September 2003.
- [12] O. Bernard, D. Friboulet, P. Thevenaz, M. Unser, "Variational B-Spline Level-Set: A Linear Filtering Approach for Fast Deformable Model Evolution," *IEEE Transactions on Image Processing*, vol.18, no.6, pp.1179-1191, 2009.
- [13] C. Li, C.Y. Kao, J.C. Gore, Z. Ding, "Minimization of Region-Scalable Fitting Energy for Image Segmentation," *IEEE Transactions on Image Processing*, vol.17, no.10, pp.1940-1949, 2008.
- [14] M. Golzarian, J. Cai, R. Frick and S. Miklavcic, "A comparative evaluation of level set algorithms with applications for the segmentation of narrow leaf plant images," *International Conference on Signal Acquisition and Processing*, 2011.