

Gaussian Mixture Models for Image-based Cereal Plant Canopy Analysis

Hamid Laga^a, Pankaj Kumar^a, Jinhai Cai^a, Stephan Haeefe^b, Raghu Anbalagan^b,
Nataliya Kovalchuk^b and Stanley J. Miklavcic^a

^a Phenomics and Bioinformatics Research Centre, University of South Australia

^b The Australian Centre for Plant Functional Genomics (ACPGF)

Email: hamid.laga@unisa.edu.au

Abstract: In this paper, we report our results of applying Gaussian Mixture Models (GMM) to the analysis of the canopy of cereal plants grown in competitive environments, such as large bins. We will particularly focus on the segmentation problem, i.e. separating the plant regions from the other image regions, such as soil, water pipes, and bin walls. We will show that GMMs, which require few training images, provide a flexible and efficient tool for high throughput segmentation at various growth stages and even in the presence of complex background. We discuss various implementation issues and provide results on a large scale experiment, where cereal plants of different genotypes are grown in large bins and subject to two different treatments (well watered and under drought stress).

Keywords: *Plant phenotyping, canopy coverage, plant growth analysis*

1 INTRODUCTION

Segmentation is a fundamental and a basic preprocessing task in many image analysis applications. It has been an important and active research topic in image processing for decades Freixenet et al. (2002); Piccardi (2004); Zhang et al. (2008). In plant image analysis, segmentation refers to the problems of (1) decomposing a plant or a canopy into various meaningful parts Chopin et al. (2015); Laga et al. (2014), such as individual leaves, stems, flowers, and spikes, and (2) separating the plants from the background, such as soil, weeds and other residues. This second problem is the focus of this paper. The performance of the segmentation step is critical to the success of many high level computation tasks, such as species recognition Laga et al. (2012, 2014), 3D reconstruction Laga and Miklavcic (2013); Yous et al. (2007), canopy coverage analysis, leaf surface area estimation, leaf counting, senescence analysis and biomass estimation.

Existing image segmentation methods can be roughly subdivided into two main categories. The first class of methods is based on motion analysis Piccardi (2004); Cremers et al. (2007); Kumar et al. (2005). These methods assume that the object of interest is moving while the background and the camera remain static. In applications to plant phenotyping, plants do not move with respect to their background and thus motion based techniques are often not applicable in this context. The second class of methods transforms the pixel data into indices that highlight color properties of the object of interest. When dealing with plant images, several methods in this category use vegetation indices Meyer and Neto (2008), e.g. excess in green index $ExG = 2g - r - b$, normalized difference index $NDI = \frac{G-R}{G+R}$, excess in green minus excess in red $ExG - ExR$, or a combination of multiple indices. Segmentation is then accomplished by learning the appropriate thresholds that separate the foreground pixels from the background ones. The learning can be either supervised, using some manually labelled data Golzarian (2011), unsupervised using some clustering methods Comaniciu and Meer (1999), or based on histogram analysis techniques, such as the Otsu threshold Otsu (1975). These methods have several limitations when applied to plant phenotyping in field or field-like environments. For instance, the background color and texture change significantly during the lifetime of a plant, due to changes in the environmental conditions, and the appearance of weeds and undesirable plants. Thus, thresholds learned for a specific imaging setup or growing conditions often do not generalize to other conditions.

Thresholding techniques can be seen as a particular case of the more general class of methods that formulate image segmentation as a pattern recognition problem. One, for instance, can consider learned thresholds as decision surfaces that divide the feature space into the various classes of objects or regions present in the image. In a more general case, one seeks to learn the statistics of the objects of interest, i.e. the probability distribution function that characterizes various classes of objects or regions present in the image. This class of methods is very popular since they allow to accommodate various lighting conditions and environmental changes, a situation that is typical in field and field-like plant phenotyping.

In this paper, we explore the usage of supervised machine learning techniques, particularly Gaussian Mixture Models, in the analysis of the canopy of cereal plants grown in field-like conditions (e.g. large bins). Traditionally, supervised techniques, which require manually labeled exemplar images for training, are not commonly used in high throughput plant phenotyping. In these applications, users, who are usually plant biologists, prefer fully automatic techniques in order to cope with the large amount of images that are acquired. We show in this paper that by manually labeling few images (11 images out of 1000) one can efficiently segment a large number of images with a significant gain in accuracy and performance. Finally, several interactive image segmentation techniques have been proposed in the literature, such as Graph Cuts Boykov and Jolly (2001), GrabCuts Rother et al. (2004) and LazySnapping Li et al. (2004). These methods achieve robust segmentation even in camouflage, when foreground and background color distributions are not well separated. These methods, however, require from the users to draw few strokes on every image to segment. Our approach however requires drawing few strokes only on few images. The learned statistical models are then applied to segment a significantly larger dataset. This makes our approach suitable for high-throughput phenotyping.

We report in this paper our results in applying Gaussian Mixture Models (GMM) for the analysis of the canopy of cereal plants grown in large bins, i.e. field-like conditions. While GMMs are very popular among the image processing and computer vision communities, they are not commonly used in cereal plant phenotyping. Thus, our contribution is the demonstration of the usage of this powerful tool for solving a fundamental problem in cereal plant phenotyping, which is the segmentation of the canopy from the background. We particularly show that by training these models on few manually annotated images, one can successfully segment a large number of images. Thus, GMMs are suitable for high throughput cereal plants' canopy analysis.

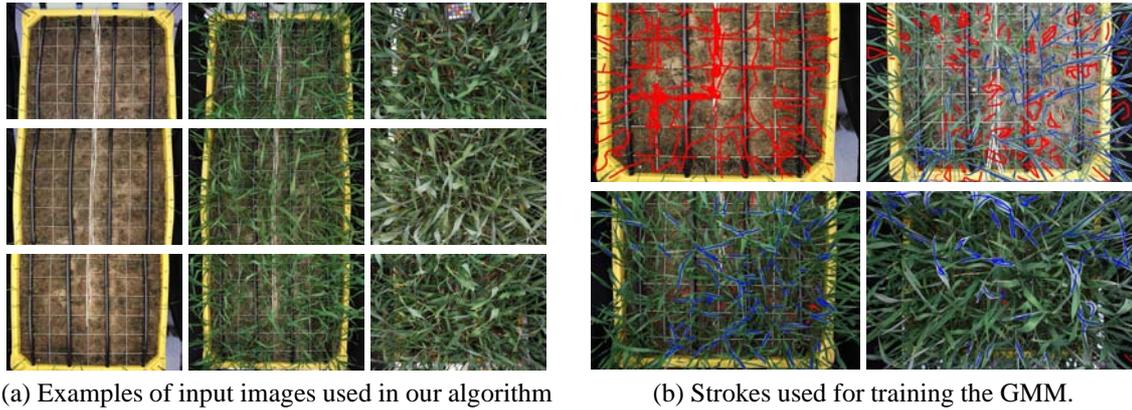


Figure 1. Examples of input images used in our algorithm.

2 APPROACH

We describe the proposed approach and discuss implementation-related choices. We will first describe the experimental environment and the imaging setup (Sec. 2.1). We then detail the segmentation procedure in Sec. 2.2. We conclude this section by detailing the type of features used in our implementation (Sec. 2.3).

2.1 Imaging system setup

We grow wheat plants in large bins inside a glasshouse. The bins are set in two rows, one for the well watered treatments and one for the drought treatments. We image each row of bins using a mobile overhead platform. The platform Anbalagan et al. (2013) contains three cameras set such that their combined field of views cover an entire bin. At the ground level, we allow approximately 30% overlap between the field of views of each pair of neighboring cameras. The height of the cameras can be manually adjusted into three different locations. During the first month (i.e. first growth stage), we set the cameras at the lowest position so that the small and tiny wheat plant leaves can be imaged at a sufficiently high resolution. When the plants grow, we move the cameras up, one step every month. This allows us to image the entire canopy at different growth stages.

In all our experiments, we used high resolution cameras (5184×3456 pixels). During the processing phase, we rescale the images to 25% of the original resolution for memory and speed purposes. Figure 1-(a) shows images taken by the three cameras, thus covering the entire bin, at three different points in time. From this figure, one can clearly see the variations that occur in colors of the plants themselves and in the background (due to different treatments, environmental conditions, and other factors). Thus, it is important to develop a segmentation procedure that takes into account not only the average colors but also the distribution of the colors over various genotypes, growth stages and environmental conditions.

2.2 Segmentation procedure

The first stage of the processing is to segment the plant regions, hereinafter call *foreground*, from the non-plant image regions, called *background*. This problem, known as background subtraction, has been extensively investigated in the literature Kumar et al. (2003); Piccardi (2004). Images of cereal plants such as wheat and barley, however, raise different types of challenges. First the plant leaves are very narrow. Thus, small segmentation errors will lead to significant errors in high level measurements such as leaf surface area, canopy coverage or biomass, which are very critical to plant biologists. Second, the background may contain moisture, weeds and other undesirable plants with green colors, making their distinction from the cereal plants very challenging. Finally, cereal plants at various growth stages, under various treatments and stresses, and under various lighting conditions, exhibit color variations (e.g. appearance of senescence) that are very hard to capture with thresholding techniques based on vegetation indices, see Meyer and Neto (2008) for example.

In this paper, we follow a statistical approach. We learn in a supervised manner the foreground and background using Gaussian Mixture Models (GMM). Let $\{FG, BG\}$ be the two states (foreground, background) of an

image pixel \mathbf{s} . The distribution \mathbf{P} of \mathbf{s} may be modeled as a weighted sum of m Gaussians:

$$\mathbf{P}(\mathbf{s}|\Phi) = \sum_{k=1}^m \alpha_k \mathbf{P}(h(\mathbf{s})|\theta_k), \quad (1)$$

where $\Phi = \{\alpha_1, \dots, \alpha_m, \theta_1, \dots, \theta_m\}$, $\theta_k = \{\mu_k, \Sigma_k\}$ are the parameters of the k -th Gaussian, α_k are the mixing coefficients that sum to one, m is the number of Gaussians in the model, and $h(\mathbf{s})$ are some image features computed at the pixel \mathbf{s} . We classify a pixel \mathbf{s} as foreground if the ratio

$$L(\mathbf{s}) = \frac{f(\mathbf{s}|\Phi_f)}{g(\mathbf{s}|\Phi_g)}$$

is larger than one. Φ_f and Φ_g are the parameters of f and g , respectively.

To learn the background model of Equation 1, we first select few images taken in similar conditions as in the system usage, sampled over the entire growth period, manually draw strokes on them, using interactive tools, to specify foreground and background pixels, and fit a GMM with Gaussians of diagonal covariance matrices to the background and to the foreground. Note that the proposed algorithm and its training procedure do not require a set of accurately segmented images for training. Instead, we only require few strokes on each image by the user. We use pure red strokes to specify the background and pure blue color for specifying the foreground, i.e. plant pixels. The choice of these two colors is motivated by the fact that they are the only colors that were not present in the captured images and thus it simplifies the automatic reading of the labels during the training procedure.

In practice, we used as few as ten images for training and spent approximately 30 minutes to label with strokes all the training images, using MS Paint software, see Figure 1-(b) for few examples. The number of images automatically processed at runtime is larger than 10K. In our experiments, we have found that one GMM f , with $m = 2$ for the foreground and one GMM g , with $m = 3$ for the background provide good segmentation results for the type of features (i.e. the function h) that we used in our implementation.

Finally, note that the training stage should be executed whenever a drastic change in the experiment setup or acquisition conditions (e.g. background type, lighting conditions) occurs. Also the segmentation procedure may produce some background noise, which we automatically eliminate using morphological operations.

2.3 Image features

The function h of Equation (1) returns at each pixel \mathbf{s} a feature vector of dimension d . We have tested with several features, however we found that color-based image features provide better performance over edge-based or generic ones, such as Zernike moments or wavelet coefficients. In particular, at each pixel \mathbf{s} , we compute $h(\mathbf{s}) = (r_s, g_s, b_s, vg_s, ug_s)$, where (r_s, g_s, b_s) are the red, green and blue values of the pixel \mathbf{s} normalized to the range $[0, 1]$ and vg_s and ug_s are vegetation indices that measure the greenness of the pixel \mathbf{s} :

$$\begin{aligned} vg_s &= \frac{2g_s - r_s - b_s}{r_s + g_s + b_s} \text{ if } (r_s + g_s + b_s) \neq 0 \text{ and } 0 \text{ otherwise.} \\ ug_s &= \frac{g_s - r_s}{r_s + g_s} \text{ if } (r_s + g_s) \neq 0 \text{ and } 0 \text{ otherwise.} \end{aligned} \quad (2)$$

We observed however that at the early growth stage, the vegetation indices were sufficient for efficient segmentation of the canopy. At the later stages, adding the three color channels to the features computed at each pixel provided better segmentation results.

3 RESULTS

We have implemented the approach described in the previous sections in Matlab and tested it for the analysis of images of cereal plants grown in the large bins. The bins have been imaged at a frequency of one image per day (except on weekends and public holidays) over a period of three months (September to November 2013). Figure 1-(a) shows a sequence of images of the same bin taken at three different growth stages.

To segment the plants from the background we trained two Gaussian mixture models, one for the foreground and another for the background. To train the models, we have manually labelled few foreground and background pixels in only ten images, see Figure 1-(b) for few examples. We use a GMM of two Gaussians for the foreground and of three Gaussians for the background.

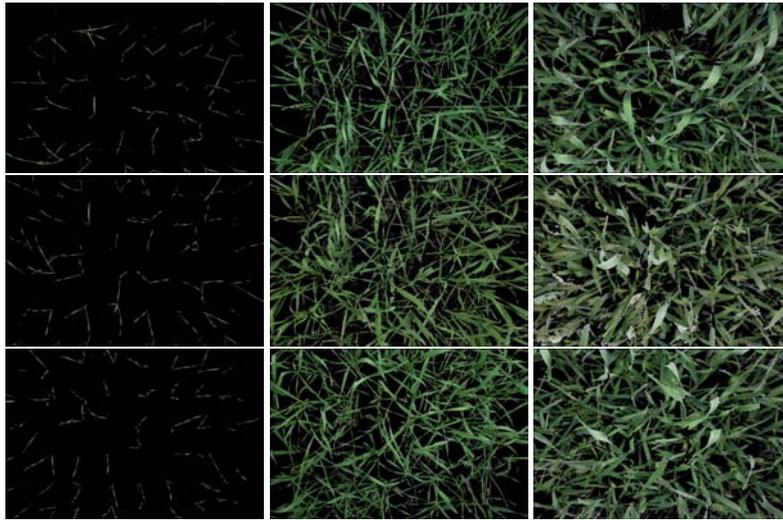


Figure 2. Segmentation of the images of Figure 1-(a).

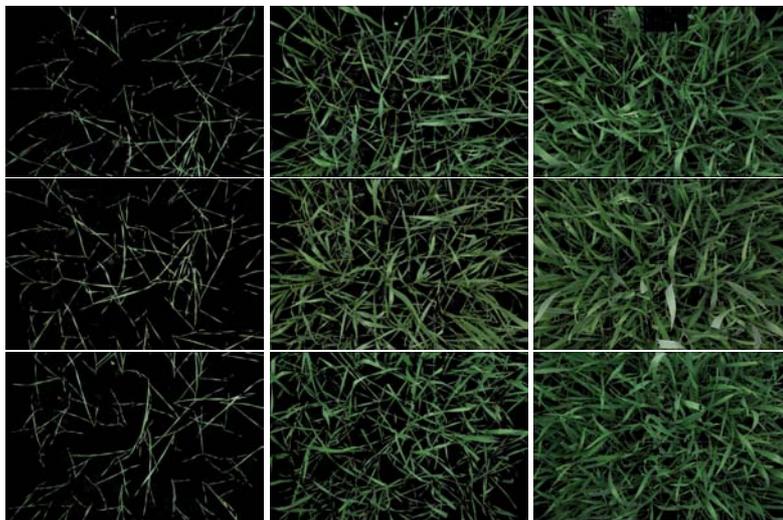


Figure 3. Additional segmentation results.

Figure 2 shows the results of segmenting the images in Figure 1-(a). Three other examples are shown in Figure 3. By visual inspection of the results, we can clearly see that the method presented in this paper is able to correctly segment the canopy of wheat plants in the presence of complex background features.

To evaluate the performance of the proposed method, we have manually segmented 24 images, recorded the canopy area in each image, and used them as ground truth. We ran our algorithm on the same set of images and compared the estimated canopy area with the manually measured one, see Figure 4 for a few examples of such images and their segmentations. Figure 5 shows a scatter plot of the estimated areas versus the manually measured one for each image. If the estimated values are exactly equal to the ground truth then all the points would be on the diagonal line of equation $y = x$ (illustrated in black in the figures). The R-squared¹ with respect to the diagonal line is 0.9232. This shows that the automatically estimated canopy area is very close to the manually measured one.

We also perform linear regression analysis on the estimated measurements. We fit to the data a straight line that minimizes the sum of squared distances between the fitted line and all the data points, see the line in blue

¹The closer the R-squared value is to one the better is the fitting.

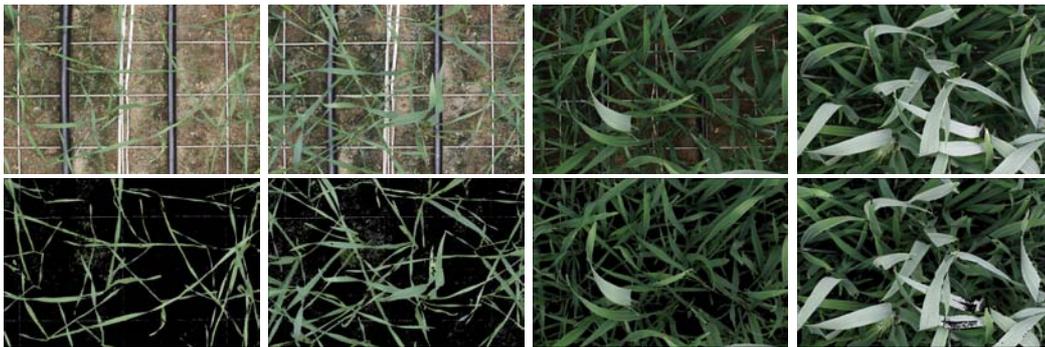


Figure 4. A sample of the images used to evaluate the accuracy of the segmentation. The top row shows the original images while the bottom row shows the results of the automatic segmentation.

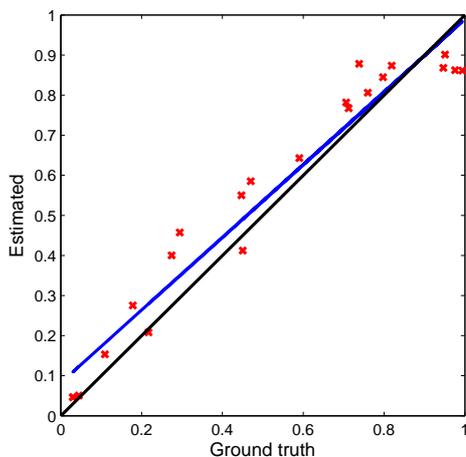


Figure 5. Comparison of the estimated canopy area with the ground truth. Each red cross corresponds to one image. The larger the error is the more the red crosses deviate from the black diagonal line of equation $y = x$. The R -squared with respect to the diagonal black line is 0.9232, and with respect to the fitted blue line is 0.9452.

in Figure 5. We obtain an R -squared value of 0.9452. This again shows that the proposed method provides accurate estimates of the canopy area.

Finally, we measure the precision and recall of the proposed method. The recall here refers to the ratio between the pixels that have been correctly classified as foreground (true positives) to the correct number of foreground pixels. The precision refers to the ratio between true positives to the total number of pixels that are classified as foreground. On average we achieved a precision of 81.21%, with a standard deviation of 15.06%, and an average recall of 86.51%, with a standard deviation of 9.44%.

Limitations. The approach presented above has two main limitations. First, results of the segmentation depends heavily on the type of images that have been used in the training stage. In our implementation, we have only used healthy plants. Thus, the learned Gaussian Mixture Models are not able to segment correctly plants that are at the senescence stage. We believe that this problem can be solved by using during training stages images of plants that are at the senescence stage.

The second limitation is the choice of the number of Gaussians both for the background as well as for the foreground. In our experiments, we have empirically chosen three Gaussians for the background and two for the foreground. Different setups may require a different number of Gaussians, particularly for the background. Finding an automated criteria for the choice of the number of Gaussians will be the subject of future research.

4 CONCLUSIONS

In this paper we have demonstrated the usage of Gaussian Mixture models for segmentation of various types of cereal plant images. We have particularly focused on canopy analysis in large bins. We have shown that with few interactions, i.e. manual labeling of few pixels in few images, one is able to subsequently segment automatically a large number of images (more than 1K in the case of bin images). This clearly suggests that Gaussian mixture models, trained with carefully chosen features, provide an efficient tool for high throughput

cereal plant phenotyping. The approach presented in this paper has some limitations that we aim to address in the future. One important issue is the choice of the number of Gaussians in the foreground and background models. For the foreground, one can for instance perform a systematic study of the distribution of the plant colors, at different growth stages and under various treatments and conditions, in order to identify potential clusters in the color space. The background, however, is particularly challenging since it can drastically change from an experiment to another. Finally, although we have only tested our framework on cereal plants, we believe the methodology applies to a wide range of image-based plant phenotyping problems.

REFERENCES

- Anbalagan, R., N. Kovalchuk, B. Parent, A. Kovalchuk, M. Okamoto, R. Whitford, and S. M. Haefele (2013). A phenotyping platform for transgenic wheat: method and initial results. *20th International Congress on Modelling and Simulation*.
- Boykov, Y. Y. and M.-P. Jolly (2001). Interactive graph cuts for optimal boundary & region segmentation of objects in nd images. In *Proceedings. Eighth IEEE International Conference on Computer Vision*, Volume 1, pp. 105–112. IEEE.
- Chopin, J., H. Laga, C. Y. Huang, S. Heuer, and S. J. Miklavcic (2015). Rootanalyzer: A cross-section image analysis tool for automated characterization of root cells and tissues. *PLoS one* 10(9), e0137655.
- Comaniciu, D. and P. Meer (1999). Mean shift analysis and applications. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, Volume 2, pp. 1197–1203. IEEE.
- Cremers, D., M. Rousson, and R. Deriche (2007). A review of statistical approaches to level set segmentation: integrating color, texture, motion and shape. *International journal of computer vision* 72(2), 195–215.
- Freixenet, J., X. Muñoz, D. Raba, J. Martí, and X. Cufí (2002). Yet another survey on image segmentation: Region and boundary information integration. In *Computer Vision ECCV 2002*, pp. 408–422. Springer.
- Golzarian, M. R. (2011). Novel image segmentation based on machine learning and its application to plant analysis. *International Journal of Information and Electronics Engineering* 1.
- Kumar, P., S. Ranganath, and W. Huang (2003). Queue based fast background modelling and fast hysteresis thresholding for better foreground segmentation. In *Joint Conference of the Fourth International Conference on Information, Communications and Signal Processing, 2003 and Fourth Pacific Rim Conference on Multimedia*, Volume 2, pp. 743–747.
- Kumar, P., S. Ranganath, W. Huang, and K. Sengupta (2005). Framework for real time behavior interpretation from traffic video. *IEEE Transactions on Intelligent Transportation Systems* 6(1), 43–53.
- Laga, H., S. Kurtek, A. Srivastava, M. Golzarian, and S. J. Miklavcic (2012). A riemannian elastic metric for shape-based plant leaf classification. In *Digital Image Computing Techniques and Applications (DICTA), 2012 International Conference on*, pp. 1–7. IEEE.
- Laga, H., S. Kurtek, A. Srivastava, and S. J. Miklavcic (2014). Landmark-free statistical analysis of the shape of plant leaves. *Journal of theoretical biology* 363, 41–52.
- Laga, H. and S. J. Miklavcic (2013). Curve-based stereo matching for 3d modeling of plants. pp. 524–530.
- Laga, H., F. Shahinnia, and D. Fleury (2014). Image-based plant stornata phenotyping. In *Control Automation Robotics & Vision (ICARCV), 2014 13th International Conference on*, pp. 217–222. IEEE.
- Li, Y., J. Sun, C.-K. Tang, and H.-Y. Shum (2004). Lazy snapping. In *ACM Transactions on Graphics (TOG)*, Volume 23, pp. 303–308. ACM.
- Meyer, G. E. and J. C. Neto (2008). Verification of color vegetation indices for automated crop imaging applications. *Computers and Electronics in Agriculture* 63(2), 282–293.
- Otsu, N. (1975). A threshold selection method from gray-level histograms. *Automatica* 11(285-296), 23–27.
- Piccardi, M. (2004). Background subtraction techniques: a review. In *Systems, man and cybernetics, 2004 IEEE international conference on*, Volume 4, pp. 3099–3104. IEEE.
- Rother, C., V. Kolmogorov, and A. Blake (2004). Grabcut: Interactive foreground extraction using iterated graph cuts. In *ACM Transactions on Graphics (TOG)*, Volume 23, pp. 309–314. ACM.
- Yous, S., H. Laga, M. Kidode, and K. Chihara (2007). Gpu-based shape from silhouettes. In *Proceedings of the 5th international conference on Computer graphics and interactive techniques in Australia and Southeast Asia*, pp. 71–77. ACM.
- Zhang, H., J. E. Fritts, and S. A. Goldman (2008). Image segmentation evaluation: A survey of unsupervised methods. *computer vision and image understanding* 110(2), 260–280.