A VISION-BASED METHOD FOR AUTOMATIZING TEA SHOOTS DETECTION

Hai Vu, Thi-Lan Le, Thanh-Hai Tran

International Research Institute MICA
Hanoi University of Science and Technology
{hai.vu, thi-lan.le, thanh-hai.tran}@mica.edu.vn

Thuy Thi Nguyen

Faculty of Information Technology
Hanoi University of Agriculture
ntthuy@hua.edu.vn

ABSTRACT

Counting tender tea shoots in a sampled area is required before making a decision for plucking. However, it is a tedious task and requires a large amount of time. In this paper, we propose a vision-based method for automatically detecting and counting the number of tea shoots in an image acquired from a tea field. First, we build a parametric model of a tea-shoot's color distribution in order to roughly separate Regions-of-Interest (ROIs) of tea shoots from a complicated background. For each ROI, we then extract supportive (local) features with expectations that these features will only appear around an apical bud of tea shoots thanks to two measurements: the density of edge pixels and a statistic of gradient directions. Consequently, the extracted features are put into a mean shift cluster to locate the position of tea shoots. The proposed method is evaluated on a set of testing images with different species of tea plants and ages. The results show 86% correct tea shoots detected, whereas 25% of a false alarm rate exists. It offers an elegant way to build an assisting tool for tea harvesting.

1. INTRODUCTION

Tea harvesting is an operation in which the tender tea shoots (buds) are picked, which is generally termed as "plucking". In order to make a plucking decision, counting correctly the maturity of tea shoots is extremely important because it determines the quality and the quantity of tea production. However, it is a tedious task and takes large amount of time for tea producers. In this work, we show how computer vision can be used to improve the manual method. Our automatic system requires only the images acquired from a tea field in order to count tea shoots. The proposed method thus significantly reduces burdens to tea producers.

Computer-aided-systems in agriculture engineering have been a growing area in computer vision. Leafsnap [1] is a well-known application for automatic plant species identification. It utilizes image features such as the saturation-value of HSV color space, and curvature features representing the shape of the leaf. Works in [2, 3] combined practical techniques of morphological operations, shape-moments for classifying leaf images with a complicated background. Golzarian et al. in [4] evaluated several level set methods such as region-based and contour-based for separating a cereal plant from captured images. Regarding tea image processing techniques, Zhiyi et al. in [5] utilized twelve color features and six texture features to identify tea quality level. Wang et al. in [6] proposed a method building a 3-D model of tea shoots from 2-D captured images. The authors in [6] carefully extracted only one tea shoot image. They thus paid much attentions to both image processing and hand-based techniques for segmenting a tea shoot.

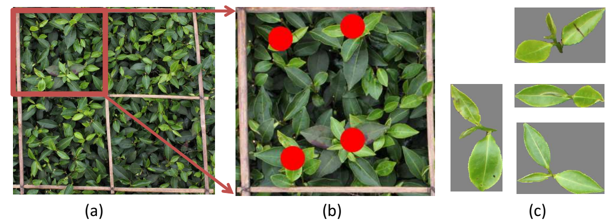

Fig. 1. (a) Our goal: estimating the density of tea shoots inside a sampled area (the square frame). (b) Position of tea shoots in red circles. (c) Illustrations of the tea shoots on (b).

As distinguished from [6], our proposed system is to automatically measure the density of tea shoots in a sampled area (per one meter square frame), as shown in Fig.1(a). To obtain this goal, we need to locate the position of tea shoots in the captured images (e.g, red markers in Fig.1(b)). It is an unwise selection if we straightforwardly utilize a segmentation-based approach such as [6] for separating tea-shoots from mother leaves[1]. Because segmenting the amount of tea shoots with high accuracy from complicated background objects (such as branches, mother leaves and other interferences) is very challenging task and requires huge computational time. In this work, we take into account informative descriptors that can easily and intuitively interpret tea shoots. We first obtain a specific the color model characterizing color distribution of tea shoots. Particularly, as shown in Fig. 1(c), a tea shoot

[1]Terms of tea plantation using in this paper is adopted from http://tea-plucking.blogspot.com

978-1-4799-2341-0/13/$31.00 ©2013 IEEE

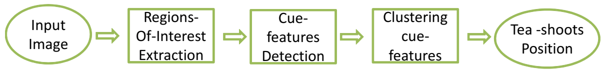

Fig. 2. The work-flow of the proposed method

consists of intersections between young leaves and an apical bud (actuve). These points can be extracted based on two measurements: the density of edge pixels and a statistic of gradient directions. They are valuable local features because the more cues we can detect, the more easily we can locate the position of tea shoots. It is a feasible scheme when the detected features can be clustered into a group. Center ones present the position of tea shoots. The proposed work-flow is illustrated in Fig. 2.

2. METHODOLOGY FOR FEATURE EXTRACTION AND DETECTION OF TEA SHOOTS

2.1. Extracting Regions-of-Interest of tea shoots

To extract ROIs of tea shoots, our approach is based on the fact that the colors of young leaves and the apical buds usually are light-green, unlike the dark-green of background of mother leaves. We learn a color model of tea shoots by adapting the skin color segmentation approach in [7]. Following comparisons in [7], the normalized color space is preferred because it provides robustness to illumination variation.

In this context, we create a training data including 300 tea shoot images which are manually segmented from different species of tea plants and ages (e.g, Fig.1(c)). Given a color pixel [R,G,B] at $[i,j]$ of a tea shoot image, the feature vector $X(i,j) = [r\ g]^T$ is extracted by: $r = \frac{R}{R+G+B}$ and $g = \frac{G}{R+G+B}$. Fig. 3(a) shows the cumulative histogram obtained from the $[r\ g]$ values of the tea shoot colour pixels in training data. The histogram dimensions are 100 x 100 bins. Fig. 3(b) is a top-view of Fig. 3(a) on a logarithmic scale whose nine contours visualize density levels. This visualization confirms that the distribution of $[r\ g]$ appears to be compatible with a unimodal elliptical Gaussian joint p.d.f model. Therefore, the color distribution of tea shoots can be parameterized using mean vectors $m_s = [m_r\ m_g]^T$ and the covariance matrix:

$$C_s = \begin{bmatrix} \sigma_r^2 & \sigma_{rg} \\ \sigma_{rg} & \sigma_g^2 \end{bmatrix} \quad (1)$$

The distance between a pixel $X(i,j) = [r\ g]^T$ and the mean color of the distribution is given by a Mahalanobis distance:

$$\lambda(i,j)^2 = [X(i,j) - m_s]^T C_s^{-1} [X(i,j) - m_s] \quad (2)$$

The probability of a pixel to be a tea shoot pixel (denoted by L_s) thus is given by:

$$P(X(i,j)|L_s) = (2\pi)^{-1} |C_s^{-\frac{1}{2}}| exp[\frac{-\lambda^2(i,j)}{2}] \quad (3)$$

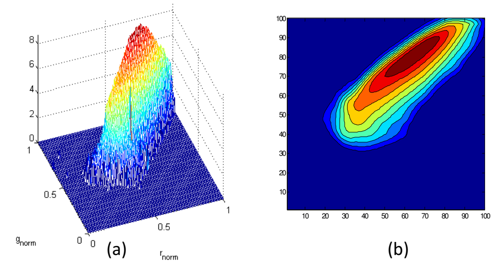

Fig. 3. (a) The color distribution of tea shoot; (b) Top view of the histogram (a) with nine levels of density

The ROIs are segmented on test images by calculating $\lambda(i,j)$ for every pixel and comparing its values to a threshold τ, which is heuristically selected. A value of 1 is assigned to pixel (i,j) if $\lambda(i,j) \leq \tau$ and a value of 0 otherwise. The result of thresholding operator is a binary image that is subjected to a median filter and top-hat transformation [8] in order to connect scatted pixels. These implementations are illustrated in Fig. 4. A quarter of an input image is shown in Fig. 4(a). Fig. 4(b) shows the image representing the distance $\lambda(i,j)$. Using the pre-selected threshold τ, ROIs are extracted as shown in Fig. 4(c). A median filter with a kernel size of $[5 \times 5]$ pixels and a structuring element a disc with diameter of 5 pixels for top-hat transformation are applied on Fig. 4(c). These regions therefore are separated as blobs (their centroid are marked in yellow circles as shown in Fig. 5(a)).

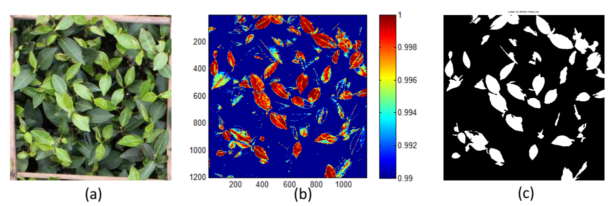

Fig. 4. (a) A quarter of an input image. (b) Distance λ calculated on (a). (c) The extracted ROIs after thresholding.

2.2. Extraction of the supportive features and Detection of tea shoots

The extracted ROIs roughly locate the young leaves of tea shoots (see yellow markers on Fig. 5(a)). However deciding which young leave(s) belong to a certain tea shoot is easily misunderstood because adjacent leaves on the captured image (even only considering young leaves) are relatively closed. To solve this issue, we take into account supportive (local) features, which are intersections between young leaves and apical buds. They uniquely present the appearance of a tea shoot. Therefore, the more we detect the supportive features, the more we can easy separate tea shoots.

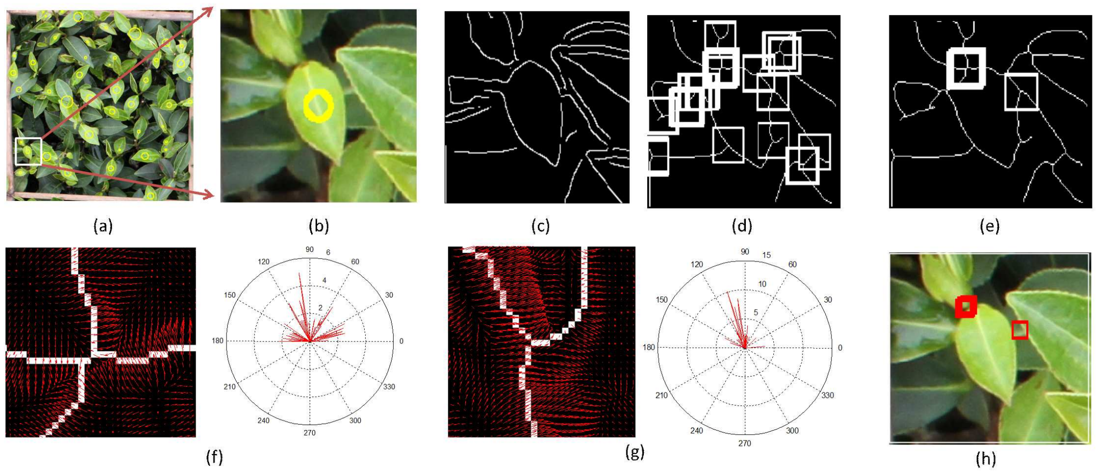

Fig. 5. Supportive features extractions. (a). Yellow circles which are centroid of the blobs are overlaid on the input image. The size of these markers is proportional to their blob size. (b) A zoom-in of window w around a centroid. (c). Edge detection results with $\sigma = 5$. (d) The features with high density edge pixels. (f) A Supportive feature (left panel) and its directional histogram (right panel) versus (g) A non-supportive feature and its directional histogram. (e) Supportive features are pruned thanks to distribution of edge orientations. (h) Supportive features are plotted on the original patch ((b)).

Given a centroid of the ROI, a set of supportive features $F = \{f_k | k = 1..N_{point}\}$ is able to search through a window patch $w = [m \times n]$ pixels around the centroid, whose size is equal two times boundary box of the ROI. At pixel $[i, j]$ in w, a feature f_k presents the intersections of leaves and bud when it ensures two criteria:

- High density of edge pixels around $[i, j]$.
- Uniformity of distribution of edge orientations.

Figure 5(b) shows a patch w for finding supportive features. To obtain this, we firstly apply a Canny edge detector [9] on the patch w, as shown in Fig. 5(c). Because intersections are points where strong streams (skeleton) converge, we then apply morphological operators including dilate and thinning [8] to get a skeleton version, as shown in Fig. 5(d). The fact that we need to remain only structure of w, a large σ value is used in the edge detector; and a high diameter of disk element is selected for the dilate operator. For the first criteria, we count edge pixels in a neighbor region with centroid at $[i, j]$. The features with high edge pixels are shown in small white rectangles in Fig. 5(d). As shown, pruning these features is required because they are located even around adjacent leaves or on the corners of a leaf. Therefore, the second criteria is applied. For each edge pixel p in the white small rectangles detected in Fig. 5(d), we extract its gradient vector $D(p) = \{dx, dy\}$. The amplitude $Amp(p)$ and direction of gradient vectors $\theta(p)$ are extracted. To express edge directional features, a polar histogram H is built with K bins (predefined with K = 256, $\triangle\theta = 360/K$ =1.4°): $H(\alpha_i) = \frac{N(\alpha_i)}{SN}$

$$\text{where } N(\alpha_i) = \sum_{p \in \Theta} \log(Amp(p)) \text{ and } SN = \sum_{i=1}^{K} N(\alpha_i)$$
$$\Theta = \{p | \alpha_i - \frac{\triangle\theta}{2} \leq \theta(p) < \alpha_i + \frac{\triangle\theta}{2}\} \tag{4}$$

The pattern of histogram H is uniform and spread every directions when a supportive feature belongs intersections, as shown in Fig. 5(f), whereas it is a non-supportive feature when a polar histogram is distributed in only a dominant direction, as shown in Fig. 5(g). Result of the pruning is shown in Fig. 5(e). Only a few supportive features are remaining. These supportive features are plotted on the original patch w, as shown in Fig. 5(h).

The above procedures are repeated for all of the ROIs. The supportive features on a whole input image are shown in Fig. 6(a). The positions of the supportive features are put to a mean-shift cluster [10]. The positions of supportive features and centroid of clusters are plotted on Fig. 6(b). For selecting the bandwidth value of the mean-shift, an estimation is roughly based on a factor between the total area of the extracted ROIs and the size of the image.

3. EXPERIMENTAL RESULTS AND DISCUSSIONS

The data set of tea images were collected in the mountainous areas in the North of Vietnam with the support of technical staffs in the SAM project [2]. The collected data includes fourth

[2]http://en.nomafsi.com.vn/newsdetail.aspx?cate=103&msgId=407

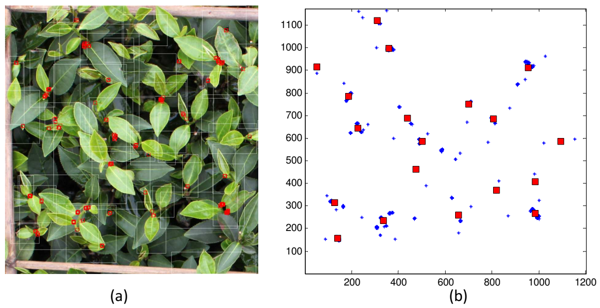

Fig. 6. (a) Red-points mark supportive features detected within windows w (white rectangle). (b) For clustering: The supportive features are presented in blue. Red squares are center of the clusters. See corresponding results of tea shoots detections on Fig. 7(a)

Table 1. Results of tea shoots detection in a testing set

# Image	TP	FP	FN	Sens. (%)	F.A.R.(%)
1	4	0	1	100	20
2	9	1	3	90	25
3	19	3	2	90	13
4	5	2	1	71	17
5	10	2	4	83	29
6	11	2	4	85	27
7	6	0	3	100	33
8	9	2	4	82	31
9	13	2	5	87	28
10	4	1	1	80	20
Avg.				86	25

species of tea plants, and two different ages. We then manually extracted 300 tea shoot images to learn their color model, as explained in Sec. 2.1. For evaluating the proposed method, we utilize a testing set including 10 input images. To make ground-truth data, technical staffs were required to manually mark positions of tea shoots on the captured images.

According to the proposed techniques, the procedures are set up and implemented by Matlab scripts on a PC Xeon 3.2 GHz, 4 GB RAM. It requires only 5 sec. to generate positions of tea shoots detected. Fig. 7(a) shows an example of testing images (Image # 3 in Table 1). It confirms that even with relatively high number of tea shoots in an input image, the proposed method still obtains reasonable results. Fig.7(b) also confirmed that the proposed method is successful with different species of tea plant. To evaluate the performance, we measure two criteria below:

$$\begin{aligned} Sensitivity &= \frac{TP}{TP+FN} \text{ and} \\ FalseAlarmRate(F.A.R.) &= \frac{FP}{TP+FP} \end{aligned} \quad (5)$$

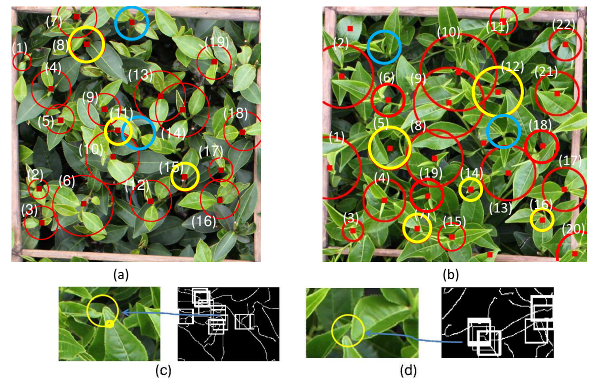

Fig. 7. (a)-(b)Two examples with different species of tea. In each panel, red circles mark true tea shoots detected. Red points at the center of circles are corresponding clusters' center in Fig. 6(b) ; yellow circles mark the wrong ones; light-blue circles mark missed detection. Each detected shoot is numbered for counting. (c)-(d) Two examples of the wrong detection cases. Correspondents of the supportive features are shown in right panels.

where TP (True Positive) is the number of true tea shoots detected; FP (False Positive) is the number of wrong tea shoots detected, and FN (False negative) is the number of lost tea shoots. Table 1 shows the evaluation results of 10 testing images. On averagely, we obtained 86% correct tea shoots detected, whereas a 25% of false alarm rate occurred. A major reason of false cases is that the supportive features were wrong detections when positions of neighbor young leaves, which belong different tea shoots, are two closed, as shown in Fig. 7(c-d). These failures may be solved by combining another image feature such as specific colors of apical bud.

4. CONCLUSIONS

In this paper, we presented an automatic system for counting tea shoots. Our proposed method focused on extracting the supportive features rather than that separates tea shoots from the background. To obtain these features, we utilized the density of edge pixels and statistics of edge orientations. These features are unique characters of tea shoots that our intuition can easy interpret. Our experimental results with different species of tea plants and ages confirmed the feasibility of the proposed method. It offers elegant means for developing an assistant tool which supports tea harvesting decisions.

5. ACKNOWLEDGEMENT

The authors thank the CUI Programme with the Hanoi Agricultural University for the support for this work.

6. REFERENCES

[1] N. Kuma, P.N. Belhumeur, and et al., "Leafsnap: A computer vision system for automatic plant species identification," in *Proceedings of the 12th European Conference on Computer Vision*, 2012, vol. LNCS 7584, pp. 502–516.

[2] G. Cerutti, L. Tougne, J. Mille, A. Vacavant, and D. Coquin, "Guiding active contours for tree leaf segmentation and identification," in *Cross-language Evaluation Forum*, Sep 2011.

[3] X.F. Wang, D.S. Huang, J.X. Du, H. Xu, and L. Heutte, "Classification of plant leaf images with complicated background," *Applied Mathematics and Computation*, vol. 205, pp. 916–926, 2008.

[4] R.M. Golzarian, J. Cai, R. Frick, and S. Miklavcic, "Segmentation of cereal plant images using level set methods a comparative study," *Int. Journal of Information and Electronics Engineering*, pp. 72–78, 2011.

[5] H. Zhiyi, C. Quansheng, and C. Jianrong, "Identification of green tea (camellia sinensis) quality level using computer vision and pattern recognition," in *Proceedings of the 2012 International Conference on Biological and Biomedical Sciences*, 2012, Advances in Biomedical Engineering, pp. 20–28.

[6] J. Wang, X. Zeng, and J. Liu, "Three-dimensional modeling of tea-shoots using images and models," *Sensors*, vol. 11, pp. 3803–3815, 2011.

[7] J.C. Terrillon and S. Akamatsu, "Comparative performance of different chrominance spaces for color segmentation and detection of human faces in complex scene images," in *Proceedings of the Vision Interface 1999,*. IEEE, 1999, pp. 180–187.

[8] R. Gonzalar and R. Woods, *Digital Image Processing*, Pearson/Prentice Hall.

[9] J. Canny, "A computational approach to edge detection," *IEEE T-PAMI*, vol. 8, pp. 679–698, 1986.

[10] K. Funkunaga and L.D. Hosteler, "The estimation of the gradient of a density function, with applications in pattern recognition," *IEEE Transactions on Information Theory*, pp. 32–40, 1975.