Grading of Vegetative Cuttings Using Computer Vision

Sanjiv Singh & Mike Montemerlo

Robotics Institute
Carnegie Mellon University
Pittsburgh, PA 15213

July 1998

Abstract

This paper reports on recent progress in the development of system to group populations of vegetative cuttings. The system is required to assign a classification to cuttings such that they appear uniform after a growing period using single two-dimensional monochrome images. We have developed a fast segmentation technique that is able to measure plant features and a supervised learning scheme that learns a mapping from the features to a scalar classification. We report results based on segmentation of over 2000 geranium cuttings. The system is able to process images at 2 Hz and has an accuracy of over 90%. Both metrics exceed human performance.

1 Introduction

Propagation of many flowering plants in the floricultural industry is accomplished by transplanting cuttings from mature plants into growing media. See Figure 1 for three cultivars of geranium plants that are propagated in large quantities Subjected to a controlled regimen of lighting and nutrition, the cuttings take root and grow into plants.

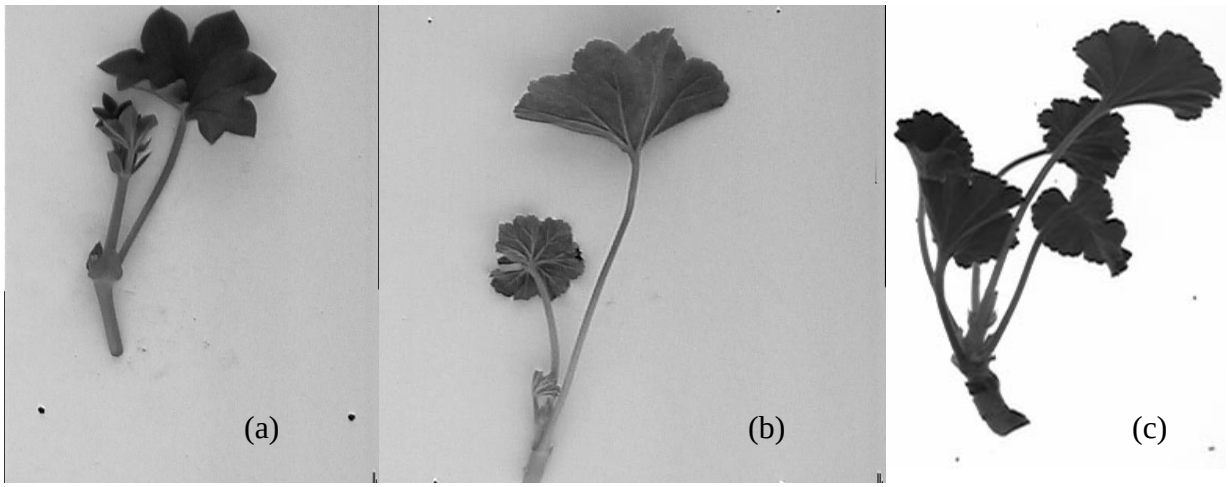

Figure 1 Three cultivars of geraniums (a) White Nicole (b) Sincerity (c) Kim.

Since it is not economically feasible to grow cuttings in singular containers, retrieve them individually, and then package them for shipment, it is necessary to assemble cuttings into multi-plant containers when they are planted so they will appear uniform at the end of the growing period. The need for uniformity arises because growing disparately sized plants next to each other results in improper lighting of some of the cuttings and can lead to stunted growth. More importantly, customers demand uniformly grown plants. Currently, in order to meet standards of uniformity it is necessary for workers to manually replace improperly grown cuttings as they are packaged for shipment.

One way to improve the quality of flowers shipped and to reduce costs is to improve the process of grouping (“grading” in the floricultural industry) the plants at the time they are

planted. For some species, the weight of a cutting is a good predictor of the size of the resulting grown plant, and automated systems have been built around such measurements. For other species, such as geraniums, weight is not a good predictor and the grading process must be done manually. Human graders can be taught the visual cues that distinguish categories of size. However, even though cuttings are graded into only three categories (small, medium and large), human grading is prone to error. A greenhouse that ships millions of flowering plants a year, finds it difficult to train people to remember the growth rates and differences between up to 100 different cultivars that might be processed. The job is also repetitive and boring; accuracy tends to degrade over time. These characteristics make the task suitable for automation.

This paper describes the development of a system that learns to grade cuttings based on a training set. That is, the system is presented with images of cuttings and their true grades. It learns to mimic this classification scheme. Recently concluded experiments are encouraging. We have tested our system with over 2000 geranium cuttings. Our system is able to process a cutting in 0.5 seconds, and has an accuracy of over 90%. Both metrics exceed human performance.

1.1 The Research Problem

Assume that each specimen of a population can be represented by a *n*-vector of features, *A*. If it were possible to tag specimen with a unique identifier, assigning a category only after a large batch had been completely processed, the problem could be stated as one of grouping specimen into the largest clusters in space spanned by *A*. In some cases, as in the case of greenhouse automation, a classification is required immediately after a single presentation. The task, we will call *classification*, then, is to find a category *C*, given a vector *A*:

$$C \leftarrow F(A) \tag{1}$$

where F is an a priori mapping. Since a greenhouse might propagate many cultivars and growing rates will vary with location and season, it is necessary for an automated grading system to learn the mapping automatically from examples presented to it. This process of *learning*, can be stated as thus: Given groundtruth, m instances of A and the corresponding C, find a mapping F that minimizes.

$$\left\| C_i - F(A_i) \right\|, i = 1 \ldots m \quad (2)$$

Finding the mapping F is complicated by several factors. First, plant morphology and growth rates vary from cultivar to cultivar. The morphology of plants can be so complicated that measurements of the feature set, A, can be very noisy. For example, cuttings are three dimensional but are imaged in two dimensions. View angle can change the appearance significantly. Additionally, growth may be non-linearly related to some features of a cutting. For example, it is known that a physically small cutting with many leaves will grow into a larger plant than a bigger cutting with fewer leaves. The classification problem is simplified somewhat because the categories are not totally disconnected. In fact there is some adjacency between categories (there is a progression from small, medium, and large). This means that it maybe possible to use discrete-valued regression.

An important question is “what sort of a representation, A, should be used for the purposes of a basis set?” Given that we have limited the ourselves to two dimensional images, the only cues that can be developed are those of shape, length and area. Thus, the feature set A, is composed of measurements that are common to all cuttings in the population that was investigated. The feature set was biologically motivated to start. For example, it is known that cuttings with wider stems grow larger than those with narrower stems. We started with an intuitively motivated set of features and used statistical analysis to eliminate those that

showed low correlation to the ultimate grade. A finer level mapping from the feature set to the grade is developed automatically by the learning method.

The task of grading a cutting has two main phases: *segmentation* and *classification*. The segmentation phase takes a grayscale image of a cutting and separates the plant into *stem*, *petiole, leaf,* and *leaf bud*. The segmented image is used to generate nine measurements characterizing certain key features of the cutting. Finally, the classification phase feeds the nine measurements into classification algorithm and generates a grade for the cutting.

2 Relationship to Other Work

Grading is a common task in the agricultural industry. With an increasing need for automation, systems have been designed to grade produce such as fruit [1], vegetables [2] and shellfish [3]. Typically such systems are based on simple criteria such as weight or size that can be measured very quickly in a production line. Grading of flowering plants has been intractable until recently partly due to the lack of economical sensing and computing. Researchers earlier in the decade have looked at the task of grading plants automatically and concluded that it was not cost-effective [4] and accuracy was not on par with human performance [5]. Simple grading schemes that separate plants into “acceptable” and “cull” have been developed [6] and recently some researchers have developed grading techniques to mimic human operators in the production of roses [7] and begonias [8].

There is a large literature on the segmentation of images of rigid objects. However, the morphology of plants is so varied that little of this work is relevant. A smaller literature exists on the segmentation of plant images, and usually for very specific purposes. For example, Simonton has developed methods to identify locations on a stem where a gripper can be applied to pick up a plant cutting [9]. Other examples of segmentation techniques can be

found in [10] and [11]. Our earlier research developed a segmentation method that uses “erosion” and “dilation” of images based on knowledge of the size of certain plant features [12][13]. Although this method is commonly used in image processing, we found it to be both slow and brittle. Recently we have developed a more robust segmentation scheme that does not require parameter tuning when the plant morphology changes from one cultivar to another. Additionally, plant images are converted into an efficient data structure (similar to the one used by Simonton) before segmentation and this allows for very rapid processing. Our new method achieves similar or better accuracy that before, but is faster by an order of magnitude.

As formulated here, learning is stated as a problem of multi-dimensional function approximation.We have started with global linear methods such as global regression and compared the results obtained with methods that attempt to capture non-linear discrimination boundaries (locally weighted regression and neural networks). Our experience is that there is very little gain in accuracy with the non-linear methods, possibly because there is a lot of noise in the measurement of the feature set, A, and inconsistencies in the groundtruth data. With a sufficiently large data set (approximately 500-600 data points), the global methods produce sufficient accuracy.

3 Segmentation

The goal of the segmentation phase is to correctly categorize each pixel of the cutting image into one of four categories— *stem*, *petiole*, *leaf* and *leaf bud* (Figure 2).

A few notes are necessary on the segmentation process. First, since the images are binarized before they are processed, the structure of the plant must be decided based solely upon the shape of the cutting (Figure 3).The advantage of such a method is that the amount of

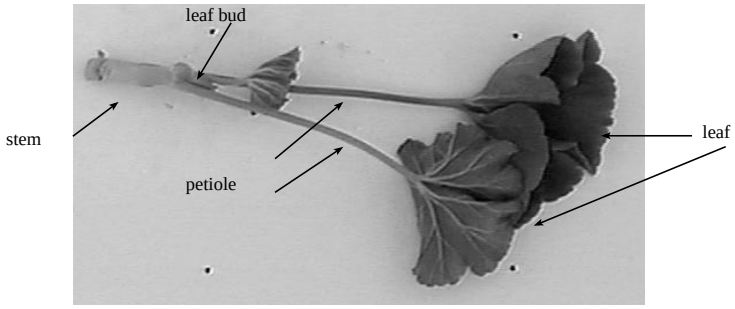

Figure 2 Parts of a cutting that must be identified automatically by the segmentation process.

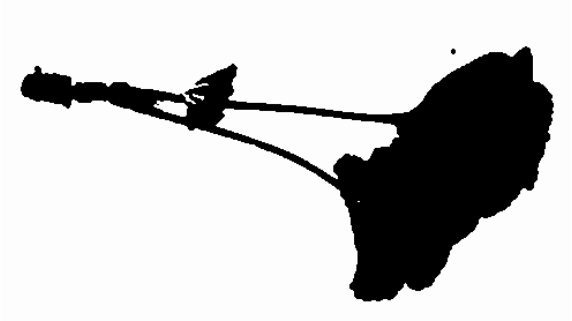

Figure 3 A binarized version of the image in Figure 3. The dark pixels represent the plant.

information to be processed is significantly reduced. However, occlusions become harder to identify correctly. Second, this segmentation algorithm is based primarily on the assumption that petioles separate the stem and the leaves of a cutting. This is true for nearly all images and in practice the segmentation fails only rarely. This assumption was demonstrated to work well in previous research [12][13].

Instead of operating directly on the image, the segmentation process converts the images into two compact graph structures roughly 5-10% the size of the original bit map. Consequently the amount of computation required is significantly reduced. Also, since the morphology of most cuttings is similar to a graph, common algorithms for graph manipulation can be used readily. The end result is a graph whose nodes represent various parts of the plant.

3.1 Conversion to a Dense Graph

After a binary image is obtained, it is converted into a *dense* graph. A dense graph is a representation of all the plant (as opposed to background) pixels (Figure 4).

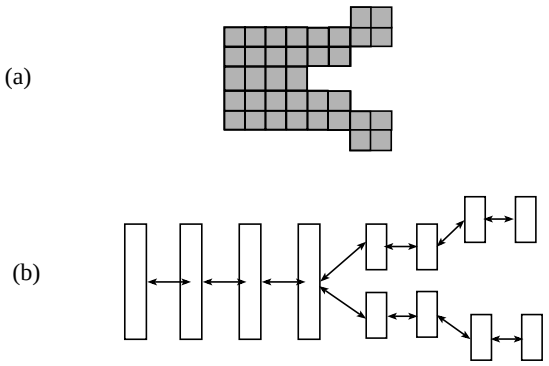

Figure 4 Creation of a dense graph from a binarized image (a) binarized image of a section of a cutting. Each cell represents a plant pixel (b) The dense graph; strips of pixels are connected into a graph

This graph is built by breaking the binary image into vertical strips, each one pixel wide. The picture is stored in an array in a manner similar to run-length encoding. Each strip has its starting row, ending row, and its column stored in the array. Also, each strip has a linked list of pointers pointing to its left and right neighbors.

The assumption that petioles separate stem and leaves implies that if the petioles can be identified and removed from the image, then the stem and the leaves will be left as separate regions. Petioles have a number of characteristics that make them simpler to identify than the stem or the leaves. They are long and thin, and consistent in width. Most importantly, they are significantly smaller in width than both the stem and the leaves. This difference shows up dramatically in a histogram of the lengths of the vertical strips in the dense graph. As shown in Figure 5, such histograms usually consist of two major peaks.

The first peak, which is largest, represents the large number of pixel strips that corresponds to petioles. The second peak represents the smaller number of pixel strips that corre-

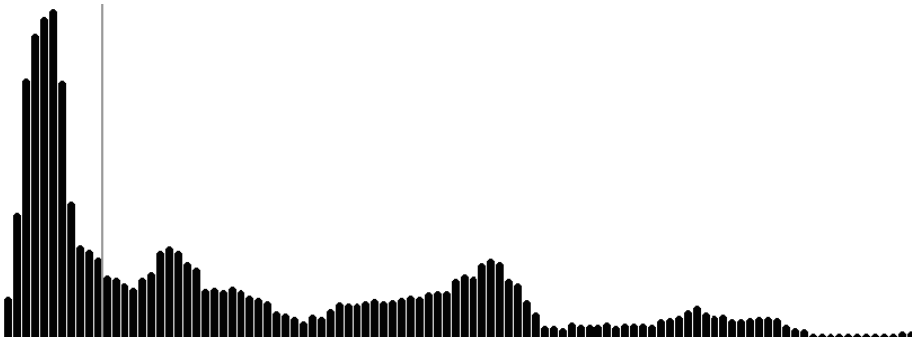

Figure 5 Histograms of lengths of the vertical strips. The left most peak corresponds to the petioles. Typically petioles have the thinnest and most consistent width.

sponds to the stem. Since the average petiole diameter and average stem diameter are usually quite different, the two peaks are distinct. Also, since the petioles are usually much longer than the stem, the petiole peak is usually stronger than the stem peak. The lack of consistency of width of the leaves keeps the pixel strips cutting through leaves from having any noticeable effect on the peaks of the histogram. Once the petiole and stem peaks have been identified, a threshold width is automatically picked. Pixel strips width widths smaller than the threshold are categorized as petiole strips.

A number of issues complicate the histogramming process. First, if the petioles of the cutting are very short, the petiole peak may be smaller than the stem peak. As the petiole peak gets smaller, the chances of picking the wrong threshold increases. For this reason, we used the ratio of the petiole peak to the stem peak as a measure of segmentation confidence. This can be used to signal images that were segmented incorrectly and reject them. A second problem is due to changing plant orientation. Since the pixel strips are all vertical, petioles may appear wider depending on their orientation. Local linear regression is used to correct the width of each. This tightens both the stem and petiole peaks, and increases segmentation confidence.

3.2 Creation of a sparse graph

Now a more compact (sparse) graph can be created. A node of the sparse graph represents a list of contiguous pixel strips of the same type (e.g. all stem pixels).

First, the sparse graph is created by adding nodes to represent clusters of like pixels in the dense graph. Once the petioles have been identified, the corresponding nodes can be labeled as such. Traversing the graph from the start and stopping until petiole nodes are reached identifies the stem nodes. The remaining nodes are labeled as leaf nodes. Some of the nodes are misclassified because their length is small enough to be confused as a petiole. A few heuristics are used on a second traversal of the sparse graph to collapse the misclassified nodes. The final step of the segmentation algorithm is identifying the leaf bud, if it exists. Any stem nodes of the sparse graph that occur to the right of the left most petiole in the image are changed into a leaf bud node. The process of constructing a sparse graph is shown in Figure 6.

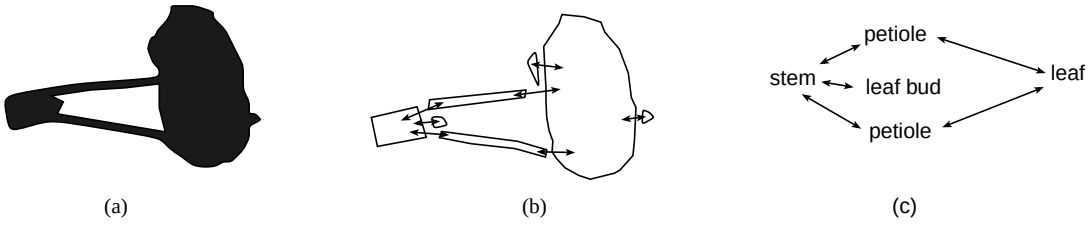

Figure 6 A sparse graph created from a dense graph (a) the initial image (b) the sparse graph in which node represents a distinct part of the cutting (c) sparse graph after second traversal.

A successful segmentation is shown in Figure 7. Experimentation with 2000 images of 8 cultivars has shown that this algorithm works well with over a range of plant morphology. Some features such as the length of the petioles and leaf count are estimated accurately even in the presence of occlusions. Estimation of other features such as leaf area and stem length can vary significantly due to occlusions.

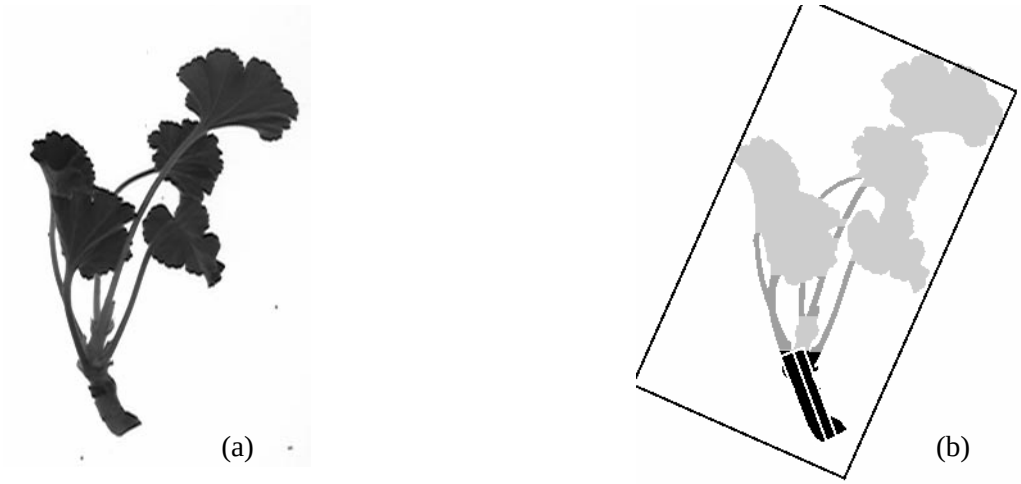

Figure 7 A successful segmentation (a) the original image (b) Stem, leaf bud, petioles and leaves have been correctly identified. A bounding box around the cutting has been computed.

This segmentation algorithm has a number of advantages over the previous method reported in [12][13]. Previously we removed petioles through a process of erosion and dilation. Besides being computationally expensive, erosion and dilation requires a hard-coded threshold value to be provided for each variety. The new algorithm requires only a difference in average width between the stem and petioles of the cutting. Erosion and dilation also smooth the image and sometimes features such as small leaves can be lost or merged with other features. Figure 8 compares the results of segmentation using the two methods.Note

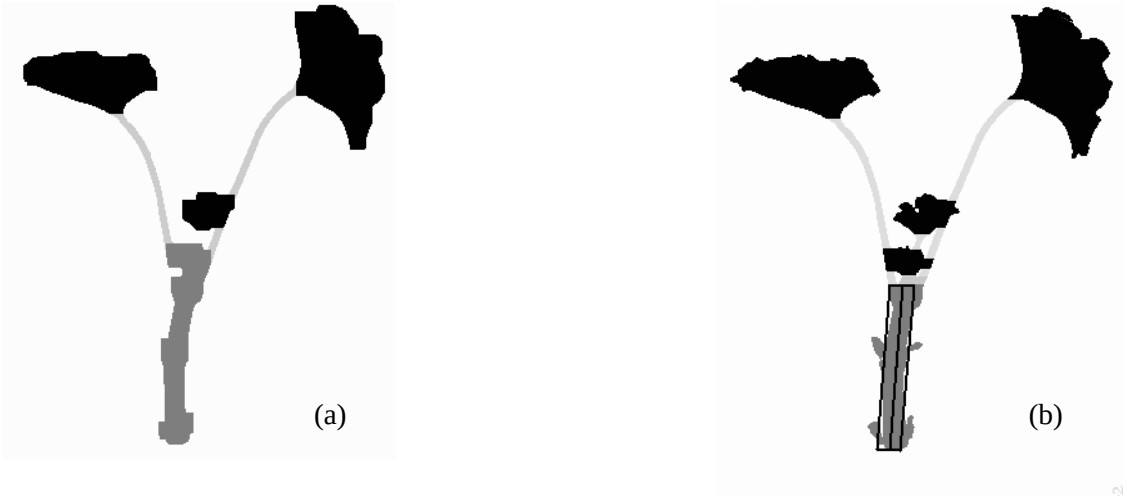

Figure 8 Segmentation of a cutting using two methods (a) erosion/dilation (b) histogramming. One of the leaves is lost with the first method due to erosion and dilation operation.

that the erosion/dilation method loses one of the leaves and the shape of features is smoothed out. In comparison, the second method correctly estimates the number of leaves and preserves details of features on the cutting. In addition, the new method poses less stringent requirements on the lighting of the cuttings. This provides an additional time savings because a simple global thresholding scheme [14] can be used rather than a more complicated adaptive tresholding that was necessary earlier. Finally, this algorithm is approximately an order of magnitude faster, segmenting a 480 x640 image in an average of 0.4 seconds on a Pentium 166 Mhz processor.

Segmentation fails in about 2% of the cases we have tried (see Figure 9 for an example).

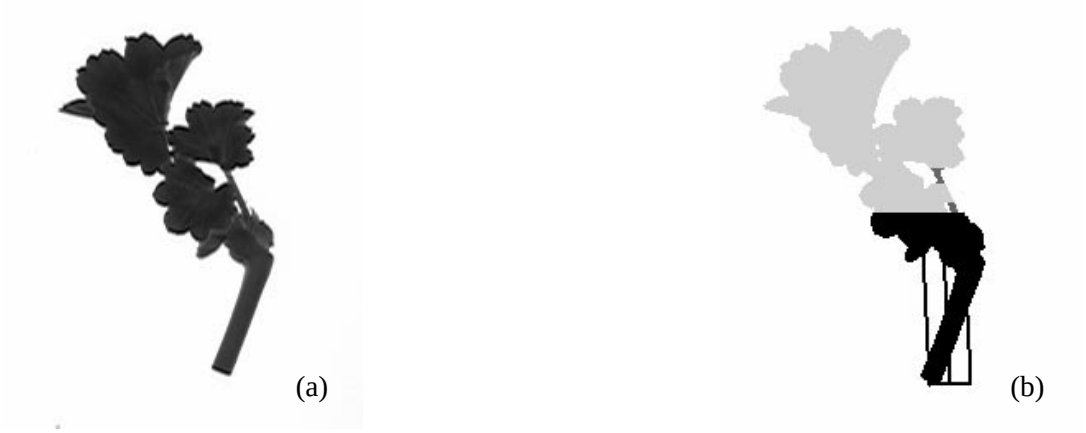

Figure 9 Error in segmentation due to occluded petioles. The corresponding histogram is not easy to bisect.

The failure is chiefly due to near-total occlusion of the petioles. In such cases it is not possible to pick a correct threshold between peaks corresponding to stem and petioles. Given that the overall error rate is low, rather than develop special algorithms to deal with these special cases, we feel that it is sufficient to signal a failure in a production system which rejects such cuttings.

3.3 Measurement

The next step is to generate the feature set, *A*. We have used nine features to represent cuttings: *stem length, stem caliber (width), total petiole length, average petiole caliber, leaf area,* t*otal plant area, leaf number, bounding box length,* and *bounding box width*.

Stem length and stem caliber are computed using a robust rectangle fitting procedure. First, a line is fit using the midpoints of all of the stem pixel strips using robust linear regression. This line is the center line of the stem. Next, the median perpendicular distance from this center line to the outermost stem points is computed. This provides the stem width.

Average petiole caliber is computed based on the "corrected" petioles widths described in the segmentation section. Total petiole length is determined by computing the total distance from the end of the stem to each leaf region that is connected to by a petiole. Leaf area and total plant area are computed in square pixels and then converted into square centimeters. Two issues make leaf number the most difficult measurement to compute correctly. First, since leaves often overlap, it is difficult to determine the total number of leaves in a large leaf blob. Also, sometimes there are very small developing leaves near the end of the stem which should be ignored in the leaf count. We use a heuristic based on the number of leaf nodes and the number of petioles to estimate the number of leaves.

Calculation of bounding box length and bounding box width is similar to stem length and width, except the process is performed on the entire plant. Robust line fitting is done on the outermost pixels of the cutting. The maximum perpendicular distance from this center line is chosen as the bounding box width. The maximum parallel length of the plant is the bounding box length. Unlike standard bounding boxes, this rectangle is oriented in the same direction as the cutting.

4 Learning

The last part of the grading process is the classification of a cutting into one of three categories (small, medium and large). We have formulated the problem as a supervised learning process. In the ideal case, the training process would use accurate measurements of the features along with accurate groundtruth. In fact neither is possible. The segmentation process can be in error due to occlusions of various kinds. Groundtruth data, that is the assignment of the correct grade to a cutting after it has been grown are also not always accurate. Lastly, even under perfect propagation conditions, plants do not grow consistently. Figure 10 shows the distribution of 9 features for a set of 600 cuttings (cultivar *kim*) separated by the final grade assigned by the expert grader.

As can be seen, there is a significant overlap between grades for many of the features (e.g. petiole caliber) while others (e.g. bounding box length) are more separated. The extent of the overlap changes from variety to variety and thus the system must learn on each variety separately.

We have used statistical analysis to examine the distribution and determine the extent to which features are good predictors. We have also examined different methods of classification that most accurately predict the groundtruth data that we have. Given the spread of the distribution for each class, we have found that the best prediction comes from a global regression method. Methods that provide a non-linear mapping between the feature set and the output, such as artificial neural nets and locally linear prediction were tried as well but the difference in the results were not statistically significant. We have used global regression because it operates relatively well in the presence of noise, is extremely simple to program, fast to execute and requires no tuning. Given a feature vector A composed of measurements

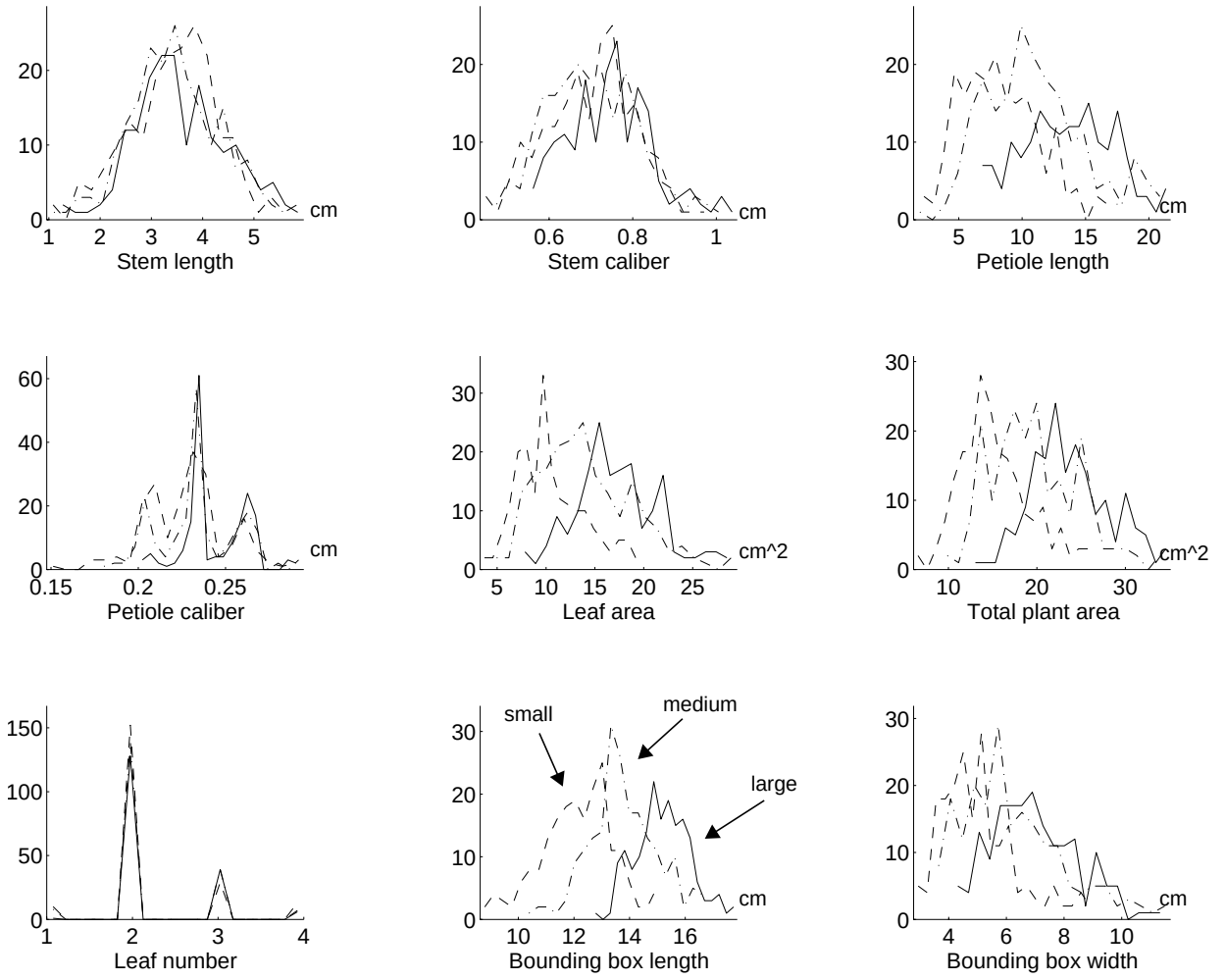

Figure 10 Histograms of nine features for a set of 600 cuttings separated by the final classification (small, medium and large).

such as shown in Figure 10, we write an over-constrained set of equations, one for each vector in the training set:

$$A_1 k_1 + \ldots + A_n k_n = C \quad \textbf{(3)}$$

The unknowns, K_i, are found by solving the normal equations:

$$K = (A^{\mathrm{T}} A)^{-1} A^{\mathrm{T}} C \quad \textbf{(4)}$$

Once K_i have been determined, the classification for a specimen is done by determining the feature vector A and solving (3).

Our investigation shows that with better ground truth (more consistent labeling), the data become more consistent, that is there is less overlap in the distributions. In this case, locally linear regression produces somewhat higher accuracy. Also, some of the features can be discarded because they do not improve the prediction.

5 Results

Our early results were based on ground truth collected in a very approximate manner. First, we imaged a large number (approximately 500-800 of each of 3 cultivars) of cuttings using a system as shown in Figure 11.

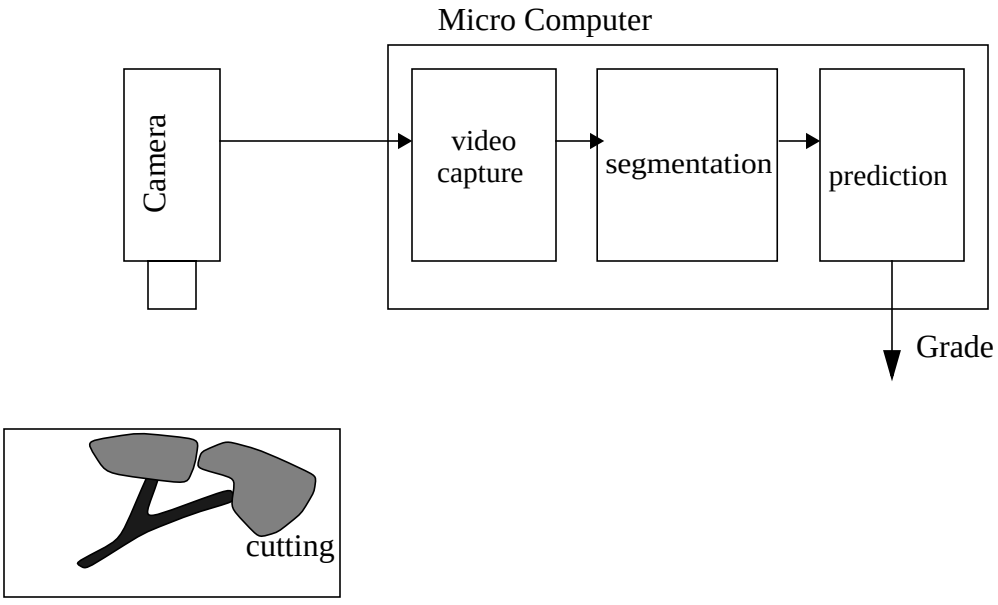

Figure 11 Integrated System for Grading. The system is able to grade cuttings at 2 Hz using a Pentium 166 processor, with an accuracy greater than 90%.

The cuttings were then planted and grown for 3-4 weeks. At this point a human grader recorded grades of the grown plants. Note that the grades are much more evident when the plants are grown than they are when they are in the form of cuttings. We then assigned grades to each set of measurements, one set corresponding to a single cutting. The accuracy

of prediction was tested using a method called *n*-fold cross validation. The data set is divided into *n* batches. For each of these batches, the system is trained on the remaining *n*-1 batches and tested with the "held out" batch. The measure of goodness is the average error over all the data tested in this way. Unfortunately, the accuracy accrued with this method of testing was low (70% - 75%). In later experimentation we discovered that a significant source of the error had to do with inconsistently assigned ground-truth data. That is, in some cases, the prediction was correct in absolute terms but the assignment of the ground truth by the human grader was itself incorrect.

Finally, we conducted a definitive test. We used a training set of approximately 600 cuttings. For each grade and each feature, any data set outside a 3 sigma range was removed. The resultant set was used to classify 450 new cuttings which were then planted and grown in a greenhouse for three weeks. After the growing period, each cutting was graded carefully under standards higher than those normally used in the production cycle. The results can be summarized in the matrix below.

	Small (Actual)	Medium (Actual)	Large (Actual)	Total
Small (Graded)	60	4	0	64
Medium (Graded)	5	147	156	156
Large (Graded)	7	14	209	230
Total	72	165	213	450
Accuracy	**83.33%**	**89.09%**	**98.12%**	**92.44%**

These accuracy numbers are for the full data set of 450. Nine of these cuttings were incorrectly segmented but the grading was conducted anyway. These cases can be easily detected and rejected before prediction. A revised accuracy measure if the cuttings with the segmentation errors are removed is 93.6%.

6 Conclusions and Future Work

A large portion of our work to date has been in establishment of a protocol for conducting experiments and in the segmentation of images. We have worked to improve segmentation accuracy without compromising generality and hope to extend the work to other commonly propagated species. We have found careful assignment of groundtruth improved classification significantly, even with the use of a simple regression method. In the future we plan to test methods of classification that require smaller training sets.

References

[1] Calpe, J.; Pla, F.; Monfort, J.; Diaz, P.; Boada, J.C.; “Robust low-cost vision system for fruit grading,” Proceedings of 8th Mediterranean Electrotechnical Conference on Industrial Applications in Power Systems, Computer Science and Telecommunications (MELECON 96); Part: vol.3; Bari, Italy; May 1996; pp. 1710-13.

[2] Heinemann, P.H.; Pathare, N.P.; Morrow, C.T., “An automated inspection station for machine-vision grading of potatoes, “Machine Vision and Applications; Mach. Vis. Appl. (USA); vol.9, no.1; Springer-Verlag; 1996; pp.14-19.

[3] Kassler, M.; Corke, P.I.; Wong, P.C.; “Automatic grading and packing of prawns,” Computers and Electronics in Agriculture, vol.9, no.4; Dec. 1993; pp. 319-33.

[4] Muttiah, R. S, Miles, G. E., “System engineering of grading plants using machine vision,” American Society of Agricultural Engineers, Paper No. 88-1543, 1988.

[5] Sistler, F. E, “Grading agricultural products with machine vision,” Grading agricultural products with machine vision Proceedings. IROS '90. IEEE International Workshop on Intelligent Robots and Systems '90, July 1990; pp. 255-61.

[6] Kranzler, G. A., Rigney, M. P., “Machine vision grading for tree seedlings,” Proceedings of the 11th International Congress on Agricultural Engineering.” pp 1883-1888, A.A. Balkema, Dubin, Ireland, September 1989.

[7] V. Steinmetz, M.J. Delwiche, D.K. Giles and R. Evans. 1994. “Sorting cut roses with machine vision,” Transactions of the ASAE. Vol. 37(4):1347-1353.

[8] Timmermans, A. J. M, Borm, T. J. A, Meinders, M. B. J., “Color vision for online sorting of begonias based on learning techniques”, In Proc. SPIE East 96, Optics in Agriculture, Forestry, and Biological Processing II. Boston, Nov 1996.

[9] Simonton, W. and Pease, J., “Orientation Independent Machine Vision Classification of Plant Parts,” *Journal Agricultural Engineering Research*, Number 54, pp. 231-243, 1993.

[10] Shiraishi, M.; Sumiya, H.; “Plant identification from leaves using quasi-sensor fusion”, Transactions of the ASME. Journal of Manufacturing Science and Engineering; Trans. ASME, J. Manuf. Sci. Eng. (USA); vol.118, no.3; ASME; Aug. 1996; pp. 382-7.

[11] Peleg, K., Cohen, O, Ziv M., and Kimmel, E., “Machine identification of buds in images of plant shoots,” Machine Vision And Applications, v. 6, p224, 1993.

[12] Ji, Q. and Singh, S., “Early results in the grading of vegetative cuttings using computer vision”, Technical Report, Robotics Institute, Carnegie Mellon University, CMU-RI-TR-96-22, May, 1996.

[13] Ji, Qiang and Singh, S., “Automated visual grading of vegetative cuttings,” In Proc. SPIE East 96, Optics in Agriculture, Forestry, and Biological Processing II. Boston, Nov 1996.

[14] Otsu, N., “A threshold selection method from gray-scale histograms,” IEEE Transactions on System, Man, and Cybernetics, Vol. SMC-9, 1979, p62-66.