COLOUR SEGMENTATION OF *CITRUS* FRUITS IMAGES FOR STEM LOCATION

L.A. Ruiz, E. Moltó, F. Juste,
Instituto Valenciano de Investigaciones Agrarias (I.V.I.A.),
Apartado Oficial, 46113 - Moncada (Valencia).

After mechanical harvesting, *Citrus* fruits with long stems may cause damage on adjacent fruits during transportation and storage. A first step to solve before measuring the stem and taking the decision of cutting it or not, is the orientation of fruits. For this purpose, an image analysis method to locate the stem insertion point based on colour segmentation is proposed and evaluated.

Four classes were defined for segmentation: *background*, *peel*, *stem-calyx* and *cut stem*, and a classifier based on bayesian decision rules was developed assuming that the independent variables (RGB) followed a normal distribution function, and that the classes had equal covariance matrix. The results point out an excellent performance of the classifier for the first two classes, and satisfactory for the other two. The location algorithm used information of areas and centroids to estimate the coordinates of the stem insertion point. From a total of 86 images tested, it was correctly estimated a 90,3% of times. The method resulted to be efficient for fruits without leaves appended to the stem.

1.- Introduction.

The development of the "Citrus" fruits picking robot "CITRUS", in the frame of the EUREKA program combining french and spanish efforts, brings some additional problems not present in manual harvesting, such as the presence of fruits with long stems, with leaves or without calyx after detachment from the tree. Long stems and leaves may cause damage on adjacent fruits, while the absence of calyx opens a way for possible infections during transportation and storage. This also means a loss of uniformity of the product not desirable for fresh market. Therefore, a system for cutting long stems and for detecting the absence of calyx before the fruit arrives to the packing houses would be convenient to preserve the quality of fruit mechanically harvested.

Traditional de-stemming systems are based on random rotation of fruits against cutting surfaces (Chen, 1993), but the contact with these surfaces may bruise the fruits. Other researchers have worked on image analysis techniques to orientate fruits and vegetables or to locate stems or other similar parts of plants. Most of the algorithms, which use a profile view of the fruit, are based on contour analysis and detection of sharp changes in contour layout. This is the case of a blueberries stem detection algorithm developed by Wolfe et al. (1984) using profile images. However, the sphericity of oranges hinders the possibility of a mechanical orientation along a conveyor belt, so the detection of the stem-calyx area using a colour vision system seems to be an adequate way to orientate the fruit from random views. Once the stem insertion point has been located, the orange can be situated at the correct position in order to determine the length of the stem and to take the decision of cutting or not.

Colour machine vision has been widely studied for classification tasks of agricultural products in the cases where grey level information was not sufficient to obtain accurate results. Slaughter and Harrell (1989) developed a classification model to discriminate oranges from natural background using digital colour images. Miller and Delwiche (1989) found a high correlation between colour maturity classification of peaches with a computer vision system and manually. Good results were also obtained in the segmentation into distinct colour regions and subsequent classification of tissue culture segments of potato plantlets (Alchanatis et al., 1992).

The objective of this study was to design a reliable image analysis method based on colour segmentation to locate the stem-calyx area of *Citrus* fruits randomly presented to the camera, in order to be able to orientate the fruit, as well as to classify it on the basis of presence or absence of stem and leaves.

2.- Image Acquisition.

Colour images were acquired with a CCD RGB video camera. Red, green and blue video signals were digitized by a colour frame grabber connected to a PC bus, to form colour images with 512 x 512 x 24-bit pixel resolution. Considering the spatial resolution provided by the digitizing board uselessly large for colour and profile analysis, the images were subdivided into 256 x 256 pixel subsets and stored as binary files. The vegetal sample used was composed of 86 oranges, var. *Salustiana*, manually collected from the field, but keeping the same proportion of long-stemmed oranges obtained with the CITRUS robot, as reported by Fornes et al., 1993, so that the results could be evaluated and referenced with respect to the robot performance. Since diffuse light is highly effective in eliminating shadows, specular reflection, and in preserving well-defined edges (Paulsen and McClure, 1986), an illumination chamber with indirect fluorescent light and diffusing material was built and used to take the images. A black background was used to increase the contrast with respect the samples.

3.- Methods.

Because of the notorious difference of colour between the peel and the stem-calyx area of *Citrus* fruits, the classification and stem location algorithm proposed is based on a colour segmentation of the images using the Red, Green and Blue colour space. Some preliminary trials revealed that this colour space was more efficient than the Hue, Saturation and Intensity space for this particular application. As showed in figure 1, a colour segmentation process was carried out and the classifier obtained would be the basis for the stem location algorithm. Both, the segmentation and the algorithm were finally evaluated.

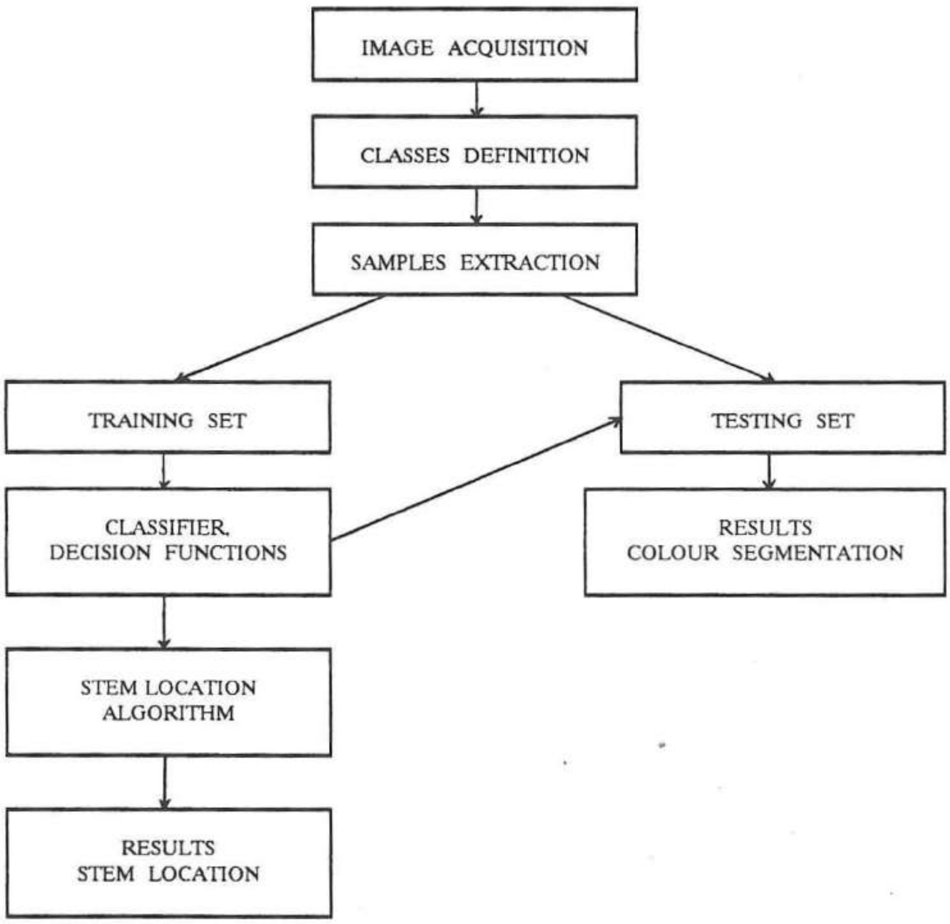

Figure 1.- Methodology followed to develop and evaluate the stem location algorithm.

The class definition was made attending to our objectives and to the ease of differentiation between potential classes. Thus, four classes were established: *background, peel, stem-calyx* and *cut stem*. The last class represents those parts of the stem that have been torn during the harvesting process, showing internal vegetal tissue without chlorophyll, and helping at the estimation of the actual size of the stem. Differentiation of classes on the RGB colour space is represented on the two scatterplots in figure 2. The first three classes are manifested on the plane formed by the components red and green, while the blue component is useful to evidence the *cut stem* class.

The extraction of samples was made interactively by a human operator, selecting groups of pixels from the images visualized on a video monitor, then gathering the red, green and blue values and assigning them a class, storing these data into a file for later analysis. In this way, two sets of pixels were selected: the *training set*, that would make the basis for the development of the classifier, and the *testing set* to evaluate its performance. The former was composed of 15877 samples or pixels extracted from 10 different images, chosen trying to be representative of each class. For the testing set, 50427 samples taken from 24 different images were used.

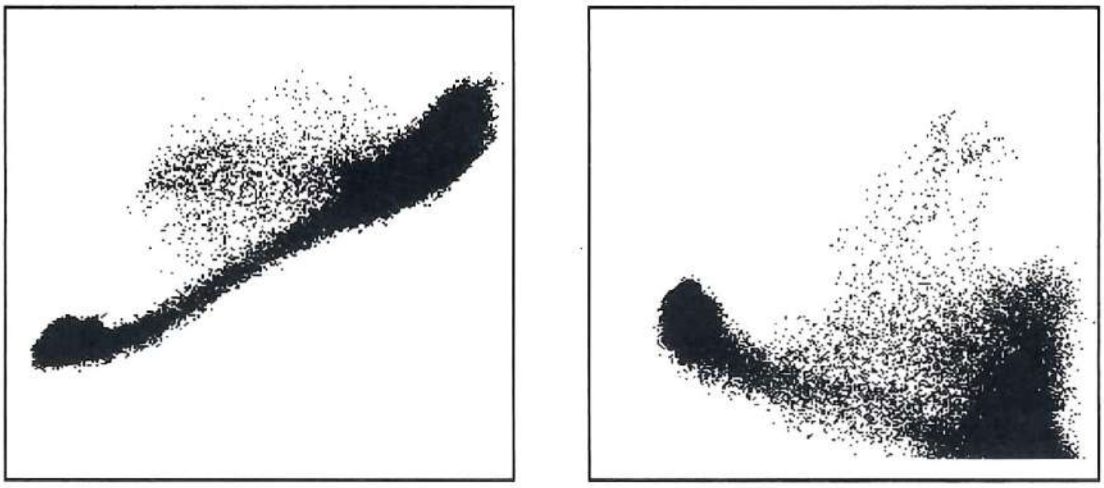

Figure 2.- Distribution of pixel values, from an orange fruit colour image, projected on the planes formed by the colour components red-green (left), and red-blue (right).

The covariance matrix of the red, green and blue variables were estimated from the training set of data to develop a classifier by the application of bayesian decision rules. The bayesian classifier is based on the computation of the *a posteriori* probabilities that a sample belongs to every class (Duda and Hart, 1973). In our case, those are the probabilities that each pixel belongs to *background*, *peel*, *stem* or *cut stem*, attending to the values of red, green and blue components for each observation. After the results obtained from an estimation of the area occupied by each class made over several images, the *a priori* probabilities entered for the four classes were 0.4, 0,4, 0,18 and 0,02 respectively. Two assumptions were made to simplify the process:

- The independent variables (red, green, blue) follow a normal distribution function for each class.
- They have equal covariance matrix.

Therefore, four linear decision functions minimizing the classification error, linear combination of the independent variables, were computed (table 1) and implemented in a segmentation algorithm that assigned each point in the RGB space to the class for which the decision function value was higher, so obtaining images segmented into four classes (figure 4). The stem location algorithm started from these images and followed two steps:

1- The detection of destemmed oranges was achieved computing the total area of the pixel classes *stem* and *cut stem*, and applying the condition:

If $(Area_{stem} + Area_{cut\ stem}) < L$, then neither stem nor calyx are present. In this case, the stem is positioned at the other side of the fruit, or it has been removed during the harvesting operation. "L" is an empirical limit of the stem-calyx area under which the orange is considered without stem. In general, this limit will depend on the resolution of the image and on the distance lens-object.

2- The estimation of the insertion point of the stem, if present, was made by computing the centroid of only the part of the stem-calyx area contained inside the contour of the peel class.

Table 1 - Coefficients of the decision functions used for colour segmentation.

	Peel	Stem-Calyx	Cut Stem	Background
Red	.7029	-.0178	.2089	.1581
Green	-.0396	.3629	.3727	-.0092
Blue	-.0062	-.0696	.0708	.1423
Constant	-66.6955	-23.4591	-69.2199	-9.9078

4.- Results.

Discriminant functions obtained with the colour segmentation procedure were effective to divide the colour space into the selected classes. Figure 3 shows the projection of some randomly chosen pixels from the training set on a plane defined by two discriminant functions.

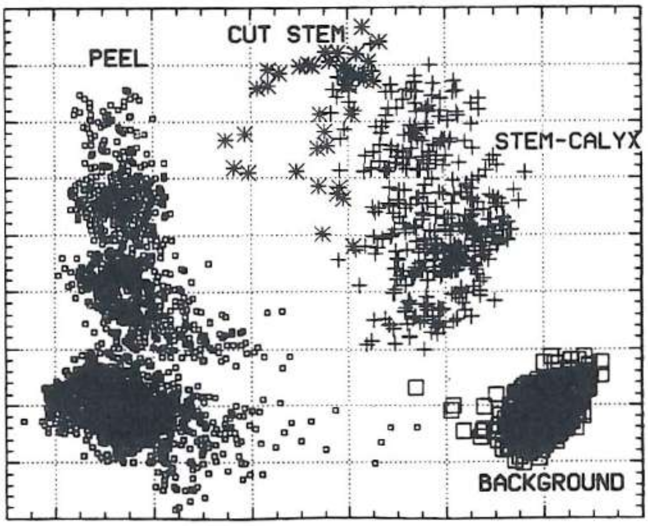

Figure 3.- Representation of a group of pixels, randomly chosen from the training set, on the plane defined by two discriminant functions.

The classification matrix obtained after applying the colour segmentation model on the testing set (table 2), points out an excellent performance of the classifier for the background and the orange peel (100% and 99,97% of correctly classified pixels, respectively), a good performance for stem, calyx and leaves (93,37%), and acceptable for the cut stem (87,30%) class. It also shows that some pixels belonging to the stem-calyx class (most of them shadowed areas of leaves) are classified as background, and darkest pixels of the class *cut stem* are confused with the calyx.

Table 2 - Classification matrix of the colour segmentation process performed over the testing set of samples. (Rows: observed classifications; columns: predicted classifications).

CLASS	% correct	peel	stem-calyx	cut stem	background
peel	99.97	21395	1	0	5
stem-calyx	93.37	6	7099	2	496
cut stem	87.30	10	54	440	0
background	100.00	0	0	0	20919
TOTAL	98.86	21411	7154	442	21420

The stem location algorithm was tested using 86 images, obtaining the following results:

- The presence or absence of calyx was correctly detected the 100% of cases. A 16% of the 86 fruits tested did not have calyx.
- Image coordinates of the stem insertion point were well estimated a 90,3% of times, what represents a 8% of fails considering the total of samples tested. It is remarkable that most of the fruits where the algorithm failed presented some leaves attached to the stem, but the presence of leaves was always detected.

Some examples of the algorithm performance are showed in figure 4.

5.- Conclusions.

- The colour segmentation of *Citrus* fruits images by mean of a bayesian classifier, resulted to be effective for discrimination of the four classes defined in this study. The described lightning conditions helped to it reducing the variability into each class.
- The stem insertion point location algorithm based on this segmentation technique and using some additional information about areas and centroids of the classes, was suitable to determine presence or absence of stem and to locate it. However, its performance decreased considerably for fruits with leaves and branches attached to the stem.

- Fruit used for this research did not present serious damage or injuries but, thinking in a de-stemming system before the fruits reach the packing houses, further research should be done in order to discriminate between the stem-calyx region and some external damage with similar aspect.

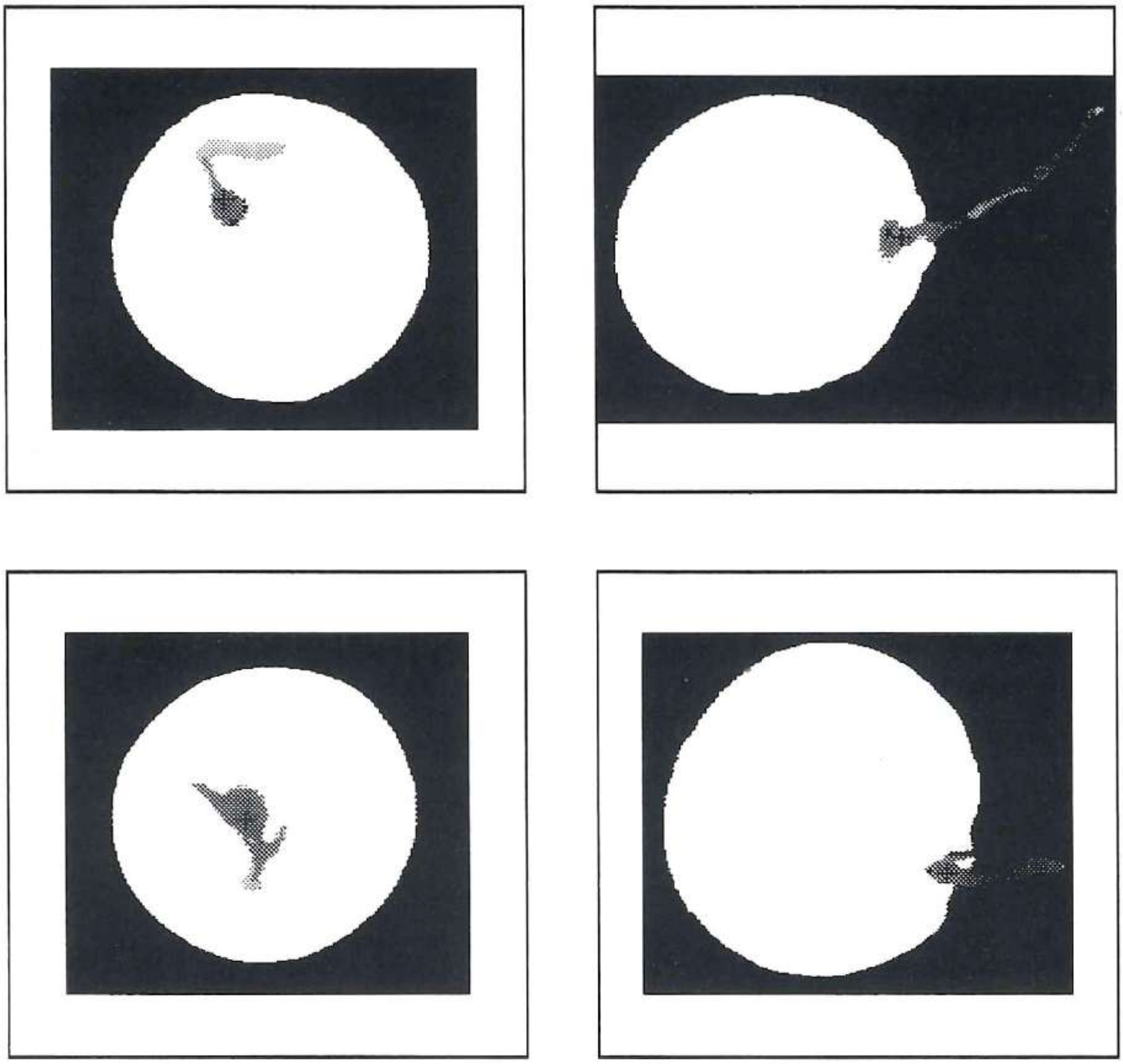

Figure 4.- Four colour segmented images of *Citrus* fruits with indication, by mean of a cross, of the insertion point estimated by the location algorithm. Each class is represented by a different grey level.

References.

Alchanatis, V., Peleg K., Ziv M. *Classification of tissue culture segments by colour machine*. Journal of Agricultural Engineering Research, (1993) 55, 299-311.

Chen, P. *Mechanical destemming of oranges.* IV International Symposium on Fruit, Nut and Vegetable

Production Engineering. 1993, Valencia-Zaragoza, Spain.

Duda, R.O. and Hart, P.E. *Pattern classification and scene analysis.* 1973, NY, John Wiley & Sons, Inc.

Fornes, I., Juste, F., Santamarina, C., Bimbo, B. *Potential damage to Citrus fruits in mechanical harvesting.* IV International Symposium on Fruit, Nut and Vegetable Production Engineering, 1993, Valencia-Zaragoza, Spain.

Miller, B.K., and Delwiche M.J. *A Color Vision System for Peach Grading.* Transactions of the ASAE, 1989, 32: 1484-1490.

Paulsen M.R., McClure W.F. *Illumination for computer vision systems.* Transactions of the ASAE, 1986, 29: 1398-1404.

Slaughter, D.C. and Harrell R.C. *Discriminating Fruit for Robotic Harvesting Using Color in Natural Outdoor Scenes.* Transactions of the ASAE, 1989, 32: 757-763.

Wolfe, R.R. and Sandler, W.E. *An algorithm for stem detection using digital image analysis.* Transactions of the ASAE, 1985, 28: 641-644.