THE QUALITY OF 'GOLDEN DELICIOUS' APPLES BY COLOUR COMPUTER VISION

V. Leemans - H. Magein - M.-F. Destain

Faculté universitaire des Sciences agronomiques de Gembloux Passage des déportés, 2 B-5030 Gembloux Belgium

Abstract : A colour machine vision system was developed to form a basis for colour grading and defect inspection of 'Golden Delicious' apples. The criteria were based on European Union standards and took into account commercial practices which add subclasses to the basic categories. The system was able to grade correctly more than 90% of the apples for colour (94% by using three colorimetric parameters R,G,B or H,S,I and 91% by using the single canonical variate) and ensured good defect detection (russet, scab, fungi attack, etc.).

Keywords : Image analysis, Image segmentation, Automatic recognition, Fruits, Grading.

1. INTRODUCTION

The classification of fruits on the basis of external quality requires the measurement of colour and the detection of defects.

The first problem is generally investigated by colour machine vision. Most often, the whole information contained in the three RGB (red, green, blue) sensors of the cameras is not taken into account and only one or two channels are used to perform image analysis. Singh et al. (1992) evaluated peach maturity by using R and G values, as they estimate that B values were small with little variation from one peach colour class to another. Tao et al. (1995) classified potatoes (good or green) and apples (yellow or green) by using hue feature. They transformed thus threedimensional RGB space to one-dimensional H space.

The detection of defects is performed either by monochromatic or by colour cameras. By introducing the concept of topographic representation of apple images, Yang (1994) treated the detection of patch-like defects as one of catchment basin detection in apple grey-level landscape. A flooding algorithm was developed and completed by a snake algorithm (Yang and Marchand, 1995). Heineman et al. (1995) used the mean hue value to distinguish russeted and good apples.

The main objective of this study is to develop algorithms with potential to classify 'Golden Delicious' apples according to the European standards (classes Extra, I and II). Furthermore, we are interested in separating the apples into subclasses on the basis of colour as indicator of ripeness degree, as it is usually performed in commercial practices. The aim is also to detect defects. Measurements are performed with colour machine vision and it is made use of the inherent correlation that exists between red, green and blue channels.

2. THE COLOUR SPACES

Images are taken under diffuse illumination provided by two fluorescent tubes, positioned in the lower part of a horizontal reflector cylinder. The interior surface of the cylinder is painted flat white. A window at the top of the cylindrical chamber is made so that the 3-CCD camera can acquire an image of the upper part of the fruit. A dark conveyor belt is chosen to provide a high contrast between the fruit and the background.

The camera acquires colour images with three sensors R, G, B whose spectral response curves peak in the red, green and blue regions of the spectrum. The R, G, B signals may be transformed into HSI coordinate system. This representation is widely used since it is close to that of human perception : hue (H) is the colour frequency, saturation (S) measures the amount of grey in the colour and I is the luminance. Furthermore, the hue has a very important characteristic related to inspection applications because it is invariant to changes in light intensity.

It is obvious that colour systems carry more information than monochromatic cameras. In each case, luminance can be extracted. Furthermore, information can be obtained about the relative amount of spectral components (RGB space colour) or about the human perception of a colour (HSI space). On the other hand, a colour system requires processing three separate frames.

3. COLOUR DISCRIMINATION

Grading on colour basis was studied on four lots of twenty apples taken out of storage in May 1996 and belonging respectively to categories A++, A+, A and Ar : this nomenclature is used in Belgium to differentiate apples within categories I into four subclasses according to their ground colour, from the greenest (A++) to the yellowest (Ar). Four images of each apple were taken perpendicularly to the stalkcalyx axis. The first image was randomly chosen and the other with 90° increments. The RGB values were measured and the image treatment may be described as follows. The resolution was reduced by three in the two orthogonal directions of the plane : one fruit corresponds thus to 10.000 to 15.000 pixels. A median filter with a square window (5*5) was applied to remove unwanted noise (lenticels, ...). The background was suppressed by a threshold on the R channel. The outer part of the fruit which appears darker due to the fruit curvature was removed by erosion. This also allowed to suppress the stem. The HSI colour space was computed.



Figure 1 : Different levels for the variability.

In a first stage, the statistical characteristics of the fruits were analysed to judge the ability of colorimetric parameters to operate good discrimination. In the RGB or HSI space, the colour of a fruit is a multidimensional distribution and the colour dispersion is represented by the variances and covariances matrix.

Several levels from an image to the population have to be considered to analyse the colour variability (Fig. 1). Each of them is characterized by its mean value and by its dispersion matrix noted Σ . $\Sigma_{imiapple}$, $\Sigma_{apple|cat}$, $\Sigma_{cat|pop}$ respectively represent the $\Sigma_{pi|im}$, variability of the pixels within an image, the images within a fruit, the fruits within a category or the categories within the whole population. To ensure a good discrimination, the variance of the parameter must increase as the level becomes higher. Within this scope, R, H, S, I are good parameters, since σ_R^2 , σ_H^2 , σ_S^2 , σ_I^2 increase from the apple|cat level to the cat|pop one. As examples, the numerical values of $\Sigma_{apple|cat}$ and $\Sigma_{cat|pop}$ are given in Table 1.

Table 1 : Dispersion matrix - RGB and HSI colour spaces

Matrix name	RGB			HSI			
mean $\Sigma_{app cat}$	(70 53 45	53 51 43	45 43 47)	1.18 -2.21 -1.91	-2.21 16.8 7.44	-1.91 7.44 37.9	
$\Sigma_{cat pop}$	(446 184 60	184 87 32	60 32 20	$\begin{pmatrix} 13\\ -46\\ -36 \end{pmatrix}$	-46 178 139	-36) 139 124)	



Figure 2 : Mean and standard deviation for RGB and HSI colour parameters of the apples within the categories.

It can be seen that σ_R^2 equals 70 when considering the apple within the categories and reaches 446 for the categories in the whole population.

The validity of R, H, S and 1 as colour indicators appears also when looking at Fig. 2 which shows the evolution of the RGB and HSI parameters according to the categories. It may be noticed that the evolution of R, H, S and I is monotone, as the colour of the fruit moves from green (A++) to yellow (Ar). During fruit ripeness, the red colour increases as concentration in chlorophyll **a** decreases. This pigment is characterized by an absorption at a wavelength of 670 nm, namely in the R channel frequency response spectrum. The hue value H, which designates the colour, decreases.

In a second stage, linear discriminant functions were computed to sort the fruit in the different categories A++ to Ar, according to the classification performed by human operators. Either one, two or three parameters defined in both colour spaces were taken into account. The percentage of apples correctly classified is given in Figure 3.



Figure 3 : Amount of apples correctly classified with one, two or three parameters.

When considering one of the three basic parameters of the RGB space, it appears that R is the best. This confirms the validity of the statistical analysis presented above. The H and S considered individually are more efficient than R for fruit grading, while I is somewhat worse. When considering two parameters, the precision increases in the classification, the best combinations are RB and HI which are obtained by associating low correlated By using simultaneously the three parameters. parameters R, G, B or H, S, I, the percentage of fruit correctly classified reaches 94%, whatever the chosen colour space.

As the R, G, B values are not independant, a canonical analysis was performed to make easier the classification (Dagnelie, 1975). The canonical variables were computed as the eigenvectors of the matrix A

 $\mathbf{A} = \mathbf{F}\mathbf{E}^{-1},$

where F is the factorial sum of square and products of deviates matrix, representing the variability between the categories, while E is the residual sum of square and products of deviates matrix, representing the variability within the categories. The first canonical variable CV which maximize the ratio of the variance between categories to the variance within the categories is used to classify the fruits. It appears that 91 % of the apples are correctly graded using this single value. Even if the accuracy is somewhat smaller than that obtained with 3 parameters (RGB or HSI), this result is very interesting. Indeed, in automatic operations, it would be easier to change the setting of a machine by adjusting one parameter rather than three.

4. DEFECT SEGMENTATION

The defect segmentation is based on the following principle. In a first step, a model describes the normal colour of healthy apple tissues, taking into account the existing variability. In a second step, each pixel of a fruit is analysed : if the difference between its colour and that of the model is too high, the pixel is considered as defect.

Eighty fruits with different ripeness degrees were used to establish the model. The resolution was reduced as in the colour study and one pixel correspond to about 0.28 mm². A principal component analysis was performed on the whole set of data. A new coordinate system (X, Y, Z) was thus found, in which all colour parameters were uncorrelated. These new features are linear combination of the old features and are eigenvectors of the covariance matrix of all the pixels of the population Σ_{pilpop} . The corresponding eigenvalues are the variances σ_x^2 , σ_y^2 and σ_z^2 .



Figure 4 : RGB space - Mean of the 320 images - First two principal components.

Fig. 4 indicates the mean experimental points of each image in the RGB space and the two first principal axes X and Y; the third axis being almost perpendicular to the view is not represented.

In a further stage, a volume including all of these points was computed. Thanks to the principal component analysis, this latter was modelled as a parallelepiped centred on the mean point (\bar{r} , g, \bar{b}) having its lengths proportional to the three standard deviations. Along X, Y and Z, the lengths are respectively equal to $2^*c^*\sigma_x$, $2^*c^*\sigma_y$ and $2^*c^*\sigma_z$, where c is a coefficient that has to be adjusted.

To find the most suitable value of c, a set including 33 apples affected by various defects was considered. Most defects were scars, russet, insect holes, sun scabs and bruises. The RGB coordinates of all the pixels were transferred into the new coordinate system XYZ. After that, the value of c was chosen so that the defects pixels would be outside of the parallelepiped, as indicated on Fig. 5.

The proposed method has been tested on eighty Belgian 'Golden Delicious', containing healthy fruits and apples affected by various defects such as insects bites, scabs, russet, etc. Each coordinate of all the pixels (x, y, z) was compared to a threshold value computed as c multiplied by the corresponding standard deviation.

If:

$$\begin{split} |x| &> \sigma_x * c \\ \text{or } |y| &> \sigma_y * c \\ \text{or } |z| &> \sigma_z * c \end{split}$$

then the pixel is a defect. Otherwise it is sound tissue.



Figure 5 : RGB space with the parallelepiped including the sound tissue and excluding the bad one; left : view perpendicular to the first two principal component; right view from the I axis.

The brightest part of the fruit may be erroneously segmented as defect. This was corrected by applying a threshold on the R channel. After the segmentation, small dots like lenticels are eliminated by a morphological close filter.

Some results are presented in Fig. 6, where the left part is the original image, and the right part is the result of the segmentation. The damages are a scab (6a), a fungi attack (6b) and some russeting(6c). These blemishes are of different size, shapes and contrasts.



Figure 6 : Samples of apples with various defects, segmented with the general model; a - scab attack; b - fungi attack; c - russeting

The main defects are correctly segmented. The patch-like defects caused by scab and the spotty defects resulting from fungi attack are identified with good accuracy. Diffuse defects presenting a irregular and blurred boundary wich includes pixels having a colour close to that of healthy tissue may lead to an under segmentation. For example, in Fig 6c, the blemish areas due to russet are somewhat smaller than what the eye see. In all the cases, the stalk-calyx region is identified as a defect. This problem could be solved in further studies by applying particular techniques of image analysis such as pattern recognition.

5. CONCLUSION

Classification of 'Golden Delicious' apples reaches 94% when performing linear discriminant analysis involving the 3 R, G, B or H, S, I parameters. The accuracy is somewhat lower (91%) when making use of the first canonical variate but easier machine grading settings could be obtained with this single parameter.

Efficient defects (scabs, russet, fungi attack,...) segmentation is obtained by modelling sound tissues as a volume in the principal coordinate system and considering that all defects are located outside this volume. Using this technique, further tests can be performed to classify the apples taking into account the size, shape and number of defects.

Acknowledgement

The research is founded by Federal Ministry of Agriculture of Belgium. Project n° D1/4-6121/5703 A

REFERENCES

Dagnelie P., [1975], "Analyse statistique à plusieurs variables". Les presses agronomiques de Gembloux, Gembloux.

Heinemann P.H., Varghese Z.A., Morrow C.T., Sommer III H.J., Crassweller R.M., [1995], "Machine vision inspection of Golden Delicious apples". *Applied Engineering in Agriculture, ASAE* 11 (6), 901-906.

Journal of European Community n° L97/19, "Norme de qualité des pommes et poires de tables" réglement CEE 920/89. 11-04-89.

Singh N., Delwiche M. J., Johnson R.S., Thompson J., [1992], "Peach maturity grading with color computer vision". *1992 International Summer Meeting - ASAE meeting presentation*, Paper 92-3029.

Tao Y. Heinemann P.H., Varghese Z.A., Morrow C.T., Sommer III H.J., [1995], "Machine vision inspection of potatoes and apples". *Trans. of the* ASAE, 38(5) 1555-1561.

Yang Q., [1994], "An approach to apple surface feature detection by machine vision". *Computers and Electronics in Agriculture* 11, 249-264.

Yang Q., Marchant J. A., [1995], "Accurate blemish detection with active contour model". *Computers and Electronics in Agriculture* 14, 77-89.