Objective Colour Measurement of Tomatoes and Limes

H.M.W. Bunnik[1], D.G. Bailey[2] and A.J. Mawson[3]

[1]Farm Technology Group, Wageningen University and Research Centre, Netherlands
[2]Institute of Information Sciences and Technology, Massey University, Palmerston North, NZ
[3]Institute of Food, Nutrition and Human Health, Massey University, Palmerston North, NZ

Email: D.G.Bailey@massey.ac.nz

Abstract

Colour is an important parameter of fruit from which much information regarding the quality of the fruit can be gathered. Therefore a correct grading of fruit in research is necessary. Commonly only small areas of an object are processed, after which a rather coarse grading is applied. The aim of this study is to develop a low cost, objective technique that will measure the colour of limes and tomatoes. This system quantifies the colour of an entire object using four index numbers: the mean, standard deviation, skew and kurtosis, allowing an easy comparison between different objects or of the same object at different times.

Keywords: Colour change, tomato, lime, grading, image processing

1 Introduction

Colour is an important characteristic of fruit and vegetables. Consumers prefer bright red tomatoes over green or dark red tomatoes [1, 2]. Similarly their preference is for green limes, rather than yellow, since the green colour is linked to freshness [3]. Similar patterns can be found with many other fruits and vegetables.

Colour provides more information than just the marketability of the product. The colour of a tomato has a direct relation with the firmness [4]. Furthermore, the red colour gives a clear indication of the amount of lycopene (a carotenoid) in the tomato [5-7]. Lycopene is believed to have a preventive effect against several forms of cancer [8-10].

The importance of the colour of fruit is clear, yet there are many colour related questions still unanswered. To solve some of these questions, more knowledge is needed about the development of pigments such as chlorophyll and carotenoids as a function of time.

In colour research on fruit and vegetables it is common to classify the colour of an object into a number of bands. For instance a colour score has been given for limes by assigning a value of 0, 25, 50, 75 or 100% yellow [11]. The number of bands used can vary strongly from fruit to fruit. Tomato grading charts from auctions vary from 6 bands in the USA [12] to up to 12 classes in Israel and the Netherlands [2, 13, 14]. The majority of this grading work is done by expert panels and may result in errors [2]. Consumers in general have difficulty discriminating between adjacent colour grades in the most expanded colour charts [14].

In some cases a spectrometer is used with which a number of points upon the object are compared [1, 4, 11, 15]. The disadvantage of this method is that only a small number of points on an object are measured, which can give a misleading representation of the colour.

To obtain a more objective manner of grading the entire object, an automated grading mechanism is needed. Research has been done applying spectral image analysis, with very good results [16-18]. The disadvantage of this method is that the equipment involved is relatively expensive.

The aim of this research is therefore the development of a relatively inexpensive, automated method of objectively grading an object in its entirety. This results in a set of numbers that describe the measured object in such a way that an easy comparison is possible between objects, or of a single object over time. The method is then used to quantify the colour change (‘degreening’) of limes and tomatoes over time.

This leads to the following research questions:

- Is it possible to follow the degreening of the complete area of a lime or tomato, using a standard video camera?
- How can an object be characterised by a small set of numbers that can be easily compared?

2 Instrument Design

2.1 Image Capture

To capture the images of the object (lime or tomato), a charge coupled device (CCD) camera is used (Sony DFW-SX900, resolution 960 x 1280 pixels). The camera is linked by firewire to a PC on which the processing software is installed.

To capture more than just one side of the object using a single camera, a turntable is used. The object is placed within a cup, to allow it to stand erect. The turntable is driven by a stepper motor. To remove a stick-slip effect at low speeds, where the fruit rotates at a different speed to that of the cup, the inside of the cup is lined with rubber, which has a high coefficient of friction.

In the case of tomatoes, the inside of the cup is also lined with soft tissue, to prevent the edges of the cup from damaging the tomato. When working with limes this precaution is not needed, since the rind of a lime is less sensitive.

Eleven images are captured over one complete rotation, which takes 18 seconds. Each area is recorded several times, at different angles, but in one rotation the whole object is equally covered. Applying a higher rotation speed will re-introduce the stick-slip effect. Further using fewer images will increase fluctuations in the outcome.

2.2 Lighting

Even, consistent lighting is essential to obtain images of the object that are of a high enough quality. A DC halogen lamp avoids problems of mains flicker. It is also small, enabling a compact construction.

To obtain accurate measurements of the colour, specular reflections from the surface of the object must be avoided. This may be achieved using indirect, diffuse illumination. For this purpose the halogen lamp is fitted into a tube. At 20 cm. below the bulb a circular plate is fitted in the middle of the tube (see Figure 1). This prevents the direct illumination of the object. The inside of the tube is bright white, to scatter the light around the disk. The area below and the cup itself are also white. This allows light to be scattered back onto the object to provide relatively even indirect lighting.

The lighting obtained is good, but as can be seen in Figure 2, there is still a gradation in the white background, with the lower part of the image receiving less light than the upper part. Improving the light further, by for instance by using a white sphere rather than a cylinder, would improve the outcome.

At the bottom of the tube a window is created through which the camera obtains the images.

2.3 Algorithm

To process the images the program VIPS (Visual Image Processing System) is used [19].

When considering the colour change of limes and tomatoes, they both start as green objects. Whereas the tomato progresses until it is red, the lime stops changing 'halfway', when it is yellow. This means that a similar algorithm can be used for both cases.

Although the halogen lamp performs well with respect to illumination, the colour temperature of the lamp does not match the camera. As the inside of the tube is white, this can be used as a reference for colour correction. The brightest 20% of the image contains only the background, so the average RGB value of these pixels provides an estimate of the white level. The colour correction requires that the background pixels are not saturated in any of the channels, as this would distort the average obtained. The black level is estimated empirically. A linear expansion is applied to each of the RGB channels to set the black level to 0, and the white level to 255. The result of this is shown in Figure 2.

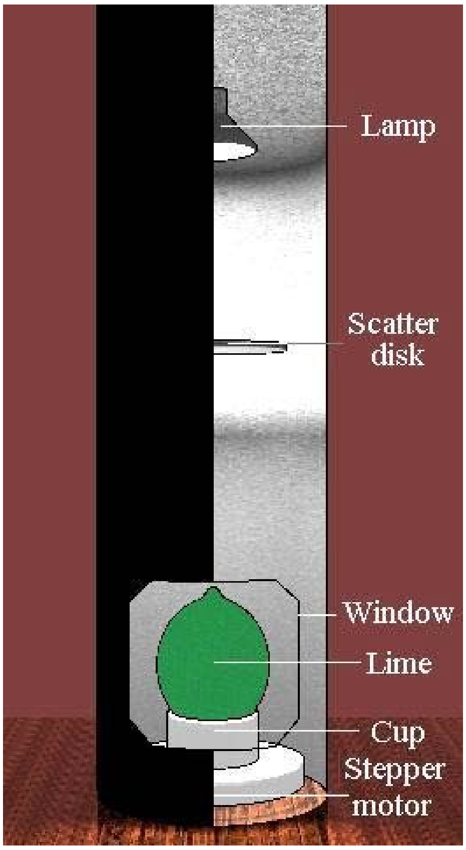

Figure 1: A sketch of the measuring device. The scatter disk ensures that only indirect lighting reaches the object.

Figure 2: Left before and right after applying the white balance and linear expansion.

The next step is to determine which pixels belong to the object being imaged. The blue channel stays approximately constant during the 'degreening' and is of a similar low level for both limes and tomatoes. Therefore this channel can be used to automatically distinguish between the background and foreground. To enable a global threshold to be used, the blue channel is normalised by dividing by the maximum of the green and red channels. For limes, the green channel is relatively uniform and consistent, although for tomatoes, the green is replaced by red as the tomato ripens.

Finally, morphological filtering (using both opening and closing) removes small isolated areas that have been mis-thresholded to give a clear silhouette of the object. A small safety margin is removed from around the edges of the mask to remove the pixels near the boundary of the object which are viewed at a very acute angle.

Physiologically, the green channel strongly reflects the chlorophyll content of the object, whereas the red channel indicates the yellowness (or carotenoid content) of limes, or the lycopene content of tomatoes. Dividing the red by the green gives a usable numerical ratio.

$$pixel\_colour\_index\_0 = \frac{Red}{Green} \qquad (1)$$

Taking the ratio in this way overcomes the small unevenness in illumination as both the red and green channels will be affected equally.

While this index is useful for limes, where the yellow contains a strong green component, with tomatoes, the green component becomes much less than the red, and a small change in colour results in a large change of ratio. A more uniform colour index may be obtained from equation 2.

$$pixel\_colour\_index\_1 = \frac{Red - Green}{Red + Green} \qquad (2)$$

Again this index is normalised against variations in illumination by taking a ratio. This index ranges from -1 for green pixels, to 0 for yellow, and +1 for red pixels. The colour index values can be easily converted to a percentage by adding 1 and scaling. To obtain the index of just the object pixels, the colour index image is masked to remove the background.

2.4 Derived Index Numbers

Once the pixel colour index of the object is known, this has to be captured in a set of meaningful numbers. A histogram is obtained of the pixel colour index values accumulated over all 11 images, see Figure 3. In this way the histogram represents the complete surface of the object apart from a small region near the stem which is sitting in the cup, and the small region near the top that is always viewed acutely, see Figure 2. This area is in general less than 5 % of the total area.

From these histograms four index numbers are generated:

- Mean (μ)
- Standard deviation (σ)
- Skew (S)
- Kurtosis (K)

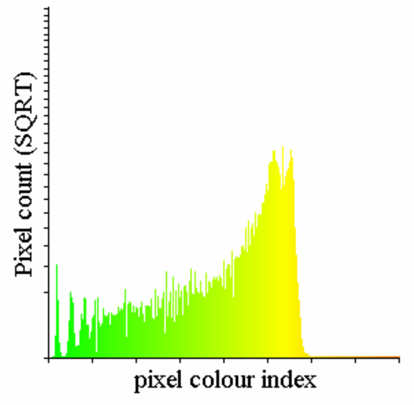

Figure 3: The histogram of a lime. Pseudo colours are applied relative to the colour of the lime.

It has been found that these four index numbers provide a good representation of the colour and colour distribution. The mean gives the average colour of the object. The standard deviation gives a measure of the range over which the colours are found. The skewness describes the symmetry in the distribution of colour. The kurtosis indicates the uniformity of the colour. These make it possible to easily compare different objects, or the same object over time.

Due to the large amount of data produced when making the measurements over a large number of objects and over an extended time, there is need for an automated processing system. This program is written in Visual Basic Excel and allows the input of large quantities of data in random order, producing the sorted data in a number of charts.

3 Initial Testing

Once the algorithm is working, the reliability of the entire setup is evaluated. For this purpose a number of tests are performed.

3.1 Consistency Testing

First the stability of the process is investigated calculating the index numbers over a long period (1 hour). In this time 100 separate measurements are made of the object, during which time the index numbers should remain constant.

The changes found in the measurements over this longer period are small and fall well within the margin compared to the coarseness of manual grading; for example see Figure 4 for a lime. During the one hour of measurement the index numbers show a stable outcome. The fluctuations stay within a range of less than 0.75 %.

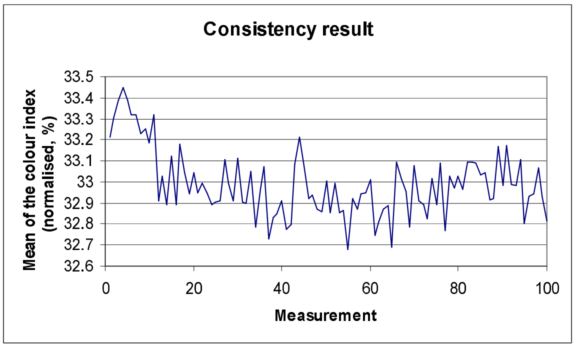

Figure 4: Fluctuations of the average colour index, during time. The y-axis is normalised between 0 and 100%. The fluctuation stays within a range of 0.75.

The other indexes showed similar, relatively small variations, although the kurtosis was sensitive to small variations in the low amplitude tails of the histogram.

The reason behind the small fluctuations is not known for certain. There is a small shadow between the cup and the object (visible in Figure 2) and this may introduce some uncertainty. With the rotation of the turntable, the fruit may also be moving slightly with the result that slightly different surfaces may be measured each time.

3.2 Ambient Lighting

The influence of ambient light sources, such as sunlight coming in the window, or the normal fluorescent lights in the laboratory, is determined. This is done by comparing index numbers under different light circumstances, while using the same object. Table 1 gives an overview of the effects.

Table 1: Effects of ambient lighting

Type of light	Mean	St. Dev.
Dark	55.15	0.006
Fluorescent	55.16	0.04
Sunlight	55.22	0.05

The effects of ambient lighting are negligible, as expected, due to the design of the measurement apparatus. The cylinder prevents any significant level of ambient light from reaching the object, apart from light coming in via the viewing window. The white balancing procedure corrects for the little light that does come in.

4 Measurement Performance

4.1 Test Setup

Finally, a small number of fruit (6 – 10) were measured, recording their colour change over time. To speed up the degreening process, the objects are kept at 20°C.

The tomatoes used are selected in a broad spectrum of colours. One was close to the breaker stage (<10% coloured, whereas others were already close to red (>90% coloured) [12].

The limes used were approximately 70% green, but with quite large differences in colour uniformity.

4.2 Results

The initial test with a small number of tomatoes showed to be very valuable.

Figures 5-8 show the evolution of the measured colour index with time. The mean colour index clearly shows a development similar to that reported in the literature [6, 20]. Due to the partial character of this initial test, the curve does not show the complete progression of colour change.

The standard deviation increases slowly, after which it stays constant at a rather high level. This is due to the relatively large, even tails in the histogram. This clearly shows the benefits of the kurtosis, since this gives a clear indication about the 'peakness' of the histogram.

For the tomato marked with '+' the measurement of the skew is different from the others. The tomato involved was initially very green (almost breaker). Due to the green colour, with a small red component, the skew is negative. In all other cases the colour is more uniform or with more red than green, resulting in a skew close to zero or positive respectively.

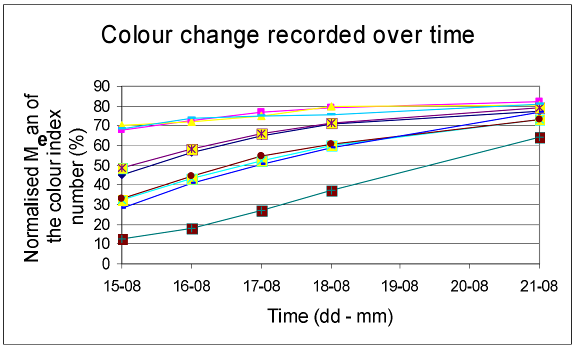

Figure 5: The development of the average colour of nine tomatoes as they change over time as recorded by the developed measuring device.

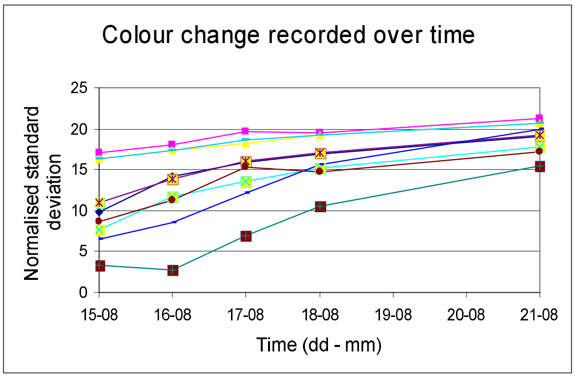

Figure 6: The development of the standard deviation of nine tomatoes as the colour changes over time.

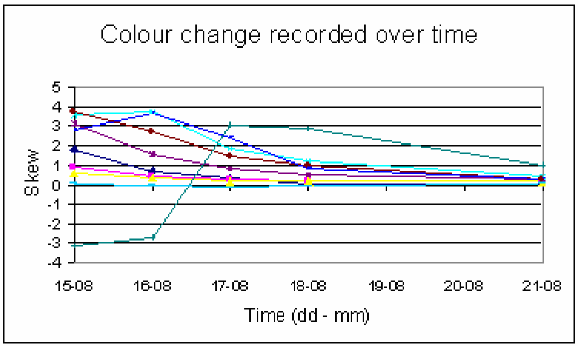

Figure 7: The development of the skew of nine tomatoes as their colour changes over time.

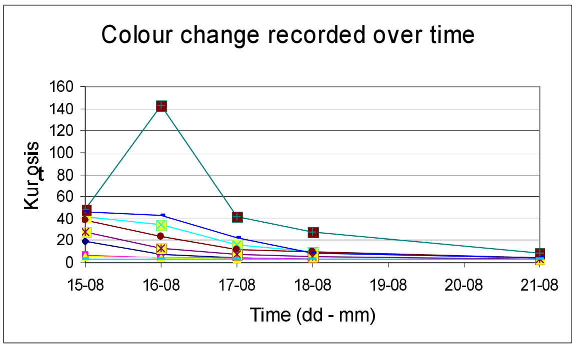

Figure 8: The development of the kurtosis of nine tomatoes as their colour change over time.

In some situations the skew and kurtosis are more difficult to interpret. Since they are 3rd and 4th order moments, they tend to be very sensitive to small fluctuations. This might obscure the outcome in certain cases. In general the outcome is good, see Figures 7 and 8. However in some cases an outlying kurtosis number is found, as for the tomato marked with a '+' in Figure 8 on 16-08 for instance.

With the limes the outcome was less good. Since the limes are stored under normal conditions, without cooling, they began dehydrating. This caused the limes to maintain their green colour, while the rind turned hard. The results out of this test where therefore less useful, other than to obtain more experience with the measuring device.

5 Conclusions

Comparing the outcome with the use of control charts, there is certainly an improvement. The number of classes into which fruit may be graded was increased significantly, while maintaining a correct classification.

The use of the four index numbers, mean, standard deviation, skew and kurtosis, gave a clear indication about the colour distribution of the object. It also allows an easy comparison using the automatic generated graphs between different objects.

The relatively slow measurement speed precludes the use of this prototype from commercial application. However, the focus of this project is more on providing a useful research tool, and the slow measurement speed is not as important.

6 Further Research

During this research it is demonstrated that the method described is useable. Therefore this method will be applied in further research in which two aspects will be investigated.

- The recording of the colour change needs to be compared to measurements made using a spectrometer.
- A larger study with more objects stored under different temperature regimes to develop kinetic models that can be used to describe colour change a range of constant or varying storage conditions.

The results of this research will be presented at a later stage.

7 Acknowledgements

I would like to acknowledge Massey University for giving me the possibility to learn much during my internship.

8 References

[1] Y. Edan, H. Pasternak, I. Shmulevich, D. Rachmani, D. Guedalia, S. Grinberg, and E. Fallik, "Color and firmness classification of fresh market tomatoes", *Journal of Food Sciences*, **62:**(4) 793-796 (1997).

[2] G. Jahns, H.M. Nielsen, and W. Paul, "Measuring image analysis attributes and modelling fuzzy consumer aspects for tomato quality grading", *Computers and Electronics in Agriculture*, **31:**(1) 17-29 (2001).

[3] E. Bosquez-Molina, J. Domínguez-Soberanes, L.J. Pérez-Flores, F. Díaz-de-León-Sánchez, and J. Vernon-Carter, "Effect of edible coatings on storage life of Mexican limes (citrus aurantifolia Swingle) Harvested in two different periods", in *XXVI IHC - Citrus, Subtropical and Tropical Fruit Crops*, **632:** 329-335 (2004).

[4] A. Batu, "Determination of acceptable firmness and colour values of tomatoes", *Journal of Food Engineering*, **61:** 471-475 (2004).

[5] R. Arias, T. Lee, L. Logendra, and H. Janes, "Correlation of lycopene measured by HPLC with the L, a, b color readings of a hydroponic tomato and the relationship of maturity with color and lycopene content", *Journal of Agricultural and Food Chemistry*, **48:** 1697-1702 (2000).

[6] S. Brandt, Z. Pék, É. Barna, A. Lugasi, and L. Helyes, "Lycopene content and colour of

ripening tomatoes as affected by environmental conditions", *Journal of the Science of Food and Agriculture*, **86:** 568-572 (2006).

[7] A. Wold, H.J. Rosenfeld, K. Holte, H. Baugerød, R. Blomhoff, and K. Haffner, "Colour of post-harvest ripened and vine ripened tomatoes (Lycopersicon esculentum Mill.) as related to total antioxidant capacity and chemical composition", *International Journal of Food Science and Technology*, **39:** 295-302 (2004).

[8] S.K. Clinton, T.M.W. Boileau, and J.W. Erdman, "Effect lycopene or tomato powder upon prostate cancer (correspondence)", *Journal of the National Cancer Institute*, **96:**(7) 554-555 (2004).

[9] P.H. Gann and F. Khachik, "Tomatoes or Lycopene versus prostate cancer; Is evolution Anti reductionist (Editorial)", *Journal of the National Cancer Institute*, **95:**(21) 1563-1565 (2003).

[10] N.I. Krinksy and E.J. Johnson, "Carotenoid actions and their relation to health and disease", *Molecular Aspects of Medicine*, **26:** 459-516 (2005).

[11] T. Pranamornkith, J.A. Heyes, and A.J. Mawson, "Effects of CA and alternative postharvest treatments of lime (Citrus latifolia Tanaka) fruit", in *International Controlled Atmosphere Research Conference*, Michigan State University, USA, (July 5-10, 2005).

[12] United States Department of Agriculture, *United States standards for grades of fresh tomatoes.* 1997.

[13] M.M. Lana, *Modelling quality of fresh-cut tomato based on stage of maturity and storage conditions*. Horticultural Production Chains Group, Plant Science. Ph.D. Thesis. Wageningen: Wageningen University. (2005).

[14] A. Lopez Camelo and P.A. Gomez, "Comparison of color indices for tomato ripening", *Horticultura Brasileira*, **22:**(3) 534-537 (2004).

[15] R. Gómez, J. Costa, M. Amo, A. Alvarruiz, M. Picazo, and J. Pardo, "Physicochemical and colorimetric evaluation of local varieties of tomato grown in SE Spain", *Journal of the Science of Food and Agriculture*, **81:** 1101-1105 (2001).

[16] G. Polder, G.W.A.M. van der Heijden, and I.T. Young, "Hyperspectral image analysis for measuring ripeness of tomatoes", in *2000 ASAE International Meeting*, Milwaukee, Wisconsin, (July 9-12, 2000).

[17] G. Polder, G.W.A.M. van der Heijden, H. van der Voet, and I.T. Young, "Measuring surface distribution of carotenes and chlorophyll in ripening tomatoes using image spectrometry", *Postharvest Biology and Technology*, **34:** 117-129 (2003).

[18] G. Polder, G.W.A.M. van der Heijden, and I.T. Young, "Tomato sorting using independent component analysis on spectral images", *Real-Time Imaging*, **9:** 253-259 (2003).

[19] D.G. Bailey and R.M. Hodgson, "VIPS - a Digital Image Processing Algorithm Development Environment", *Image and Vision Computing*, **6:** 176-184 (1988).

[20] M.L.A.T.M. Hertog, J. Lammertyn, M. Desmet, N. Scheerlinck, and B.M. Nicolai, "The impact of biological variation on postharvest behaviour of tomato fruit", *Postharvest Biology and Technology*, **34:** 271-284 (2004).

[21] F. Artes, E. Sanchez, and L.M.M. Tijskens, "Quality and shelf life of tomatoes improved by intermittent warming", *Lebensmittel-Wissenschaft und-Technologie*, **31:** 427-431 (1998).