Adv. Hort. Sci., 2023 37(1): 83-87 DOI: 10.36253/ahsc-13856

AHS
Advances in Horticultural Science

Image analysis to predict the maturity index of strawberries

A. Corvino [1], R. Romaniello [2], M. Palumbo [1, 2], I. Ricci [1], M. Cefola [1], S. Pelosi [1], B. Pace [1]

[1] *Institute of Sciences of Food Production, National Research Council (CNR) c/o CS-DAT, Via Michele Protano, 71121 Foggia, Italy.*

[2] *Department of Agriculture, Food and Natural Resources and Engineering, University of Foggia, Via Napoli, 25, 71122 Foggia, Italy.*

Key words: Computer vision system, cv. Sabrosa, *Fragaria × ananassa Duch*, harvest, multivariate analysis, ripening.

(*) **Corresponding author:**
antonia.corvino@ispa.cnr.it

Citation:
CORVINO A., ROMANIELLO R., PALUMBO M., RICCI I., CEFOLA M., PELOSI S., PACE B., 2023 - *Image analysis to predict the maturity index of strawberries.* - Adv. Hort. Sci., 37(1): 83-87.

Data Availability Statement:
All relevant data are within the paper and its Supporting Information files.

Competing Interests:
The authors declare no competing interests.

Received for publication 14 October 2022
Accepted for publication 26 January 2023

Abstract: Traditionally, strawberries are harvested manually when the typical colour of the cultivar does not reach at least 80% of the surface. The focus of this research activity is to develop an automatic system based on image analysis in order to objectively define the optimal harvest time. Strawberries (cv. Sabrosa), with different degrees of maturation, were analyzed in four different harvesting periods and subsequently selected and classified, based on the ripening percentage, in three maturity classes: R0-25, R50-70 and R75-100. Each class of 10 strawberries, evaluated in triplicate, was subjected to image analysis and physiological and qualitative evaluation by measuring the following parameters: respiration rate, pH, total soluble solids content, and titratable acidity. The images, captured by a digital camera, were processed using Matlab® software and all the data found were supported by multivariate analysis. The image processing has made it possible to create an algorithm measuring objectively the percentage and the saturation level of red assigning the fruit to each class. Principal component analysis (PCA) shows that discriminating parameters are the Chroma and the red Area, then used in a Partial Least Square Regression (PLSR) model to predict the TSS/TA ratio with R^2 of 0.7 and 0.6 for calibration and validation set, respectively.

1. Introduction

Strawberries are fruit, belonging to the family of the Rosaceae and genus *Fragaria*. Only at the end of 1600 strawberry is no longer considered an ornamental plant, but rather a fruit to be cultivated and marketed for its delicacy. *Fragaria chiloensis*, coming from Chile, arouses the interest of many farmers for its unusual size differing from other strawberry species. The current strawberry, called *F. ananassa*, comes from the random hybridization, that occurred in the second decade of 1700, of *F. virginiana* (coming from the eastern United States) with *F. chiloensis* (coming from the Chilean coasts of the Pacific) (Angelini, 2010). The obtained species is consumed and appreciated all over the world for sensorial and nutritional quality. Strawberries are usually consumed as fresh fruit and they represent a healthy food choice for their richness in vitamin C, micronutrients and bioactive compounds, mostly natural antioxidants such as phenols known even for anti-inflammatory action. This fruit, expressing better its potential, needs to be harvested at the right ripening because it is not climacteric. In general ripening influences the appearance, texture, flavour, and aroma due to physiological, biochemical and structural modifications. Strawberry is hand-picked evaluating visually the product when the characteristic colour is reached. As consequence, the mistake of collecting an overmature or immature fruit, by presenting a poor product on the market, is very likely to occur.

Generally, the qualitative parameters of fresh products are determined by destructive analytical techniques which involve a sample preparation phase, time-consuming and can be performed on a limited number of samples, often reducing their representativeness. In addition, the environmental impact, and the contact of the operator with the chemicals should not be overlooked, especially if they are not properly trained and experienced. For these reasons, it is important to consider eco-friendly and objective non-destructive methods that can quickly assess the proper harvest time by evaluating the quality of the product at hand. Numerous studies have investigated various non-destructive techniques and their applicability in the field for the determination of the main qualitative parameters of fruit and vegetables. Image analysis (IA) has proven to be a successful contactless tool in fruit and vegetables quality assessment. This technology captures images in the electromagnetic spectrum and extracts the most discriminating external characteristics (shape, colour and defects) and the next phase of data processing can allow, through predicting models, the estimation of chemical and physical properties of samples (Palumbo *et al.,* 2022). The objectives of this study were (1) to implement a standardized computer vision system to characterize quantitatively colour changes during the ripening of strawberries using the *L**, *a**, *b** colour space, (2) to identify features of interest that can be related with ripening stages, such as colour saturation (Chroma) and Hue angle, (3) to develop a statistical model using selected features to identify the ripening stages of strawberries from samples previously classified by expert visual inspection.

2. Materials and Methods

Plant material

Candonga strawberries (*Fragaria × ananassa Duch.*) var. Sabrosa, which have different degrees of ripeness, were provided by a cooperative company of fresh fruit (APOFRUIT Italia Soc. Coop., Scanzano Jonico, Italy) in four different consecutive harvest times from February to May (one harvest per month) called H1, H2, H3 and H4. Then, they were transported in cold conditions to the Postharvest Laboratory of CNR-ISPA of Foggia to be processed. Fruits were selected by eliminating damaged sample and were grouped into three classes, based on the visual evaluation of colour: R0-25 (from 0 to about 25% of red colour on fruit surface), R50-70 (from 50 to about 70% of red colour on fruit surface) and R75-100 (from 75 to about 100% of red colour on fruit surface) (Fig. 1). Each class, consisting of 10 strawberries, was evaluated in triplicate; each replicate was subjected to IA analysis, physiological (respiration rate) and physical-chemical (pH, total soluble solids and titratable acidity) characterization as below reported.

Computer vision system

Digital Camera AP-3200T-PGE (JAI Ltd., Yokohama, Japan), positioned inside a Photo studio box HPB-60D (HAVOX®, Vendôme, France), was used to image a batch containing 10 strawberries for each replicate. In total, for each class, three replicates were considered, for a total of 30 berries. The camera sensor was an RGB CMOS type, providing a spatial resolution of 3.2 MP at 2 fps and a colour depth of 24 bit/pixel. The lens used was a 12 mm focal length and F1.8 (KOWA Lens mod. LM12NC3 1/2) allowing a field of view (FOV) of (35 × 30 cm). The lighting was supplied by two LED handrails consisting of 20 diodes (HAVOX HPB-60, 5500K, 13,000 100 lumen CRI 93+). A Colour Checker Passport Photo 2 (X-rite Italy srl, Prato Italy)

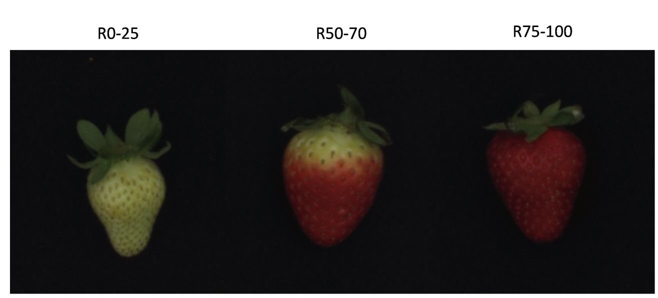


Fig. 1 - Maturity class selected for the experiment.

with 24 known colour stains was placed in the camera's FOV as a chromatic reference. The images captured by the digital camera were processed using Matlab® R2021b (MathWorks Inc., Natick, MA, USA).

Image segmentation

Each raw image of the strawberry was separated from the background, generating a binary image. In detail, the algorithm processed the raw images by cropping the unnecessary image border and separating the three colour-components: red, green and blue (RGB). The background was thresholded using the R image, since showing the highest contrast between the object of interest (strawberry) and the background. The coarse segmentation of the strawberries was carried out by a threshold method (Gonzalez *et al.,* 2004). On the resulting binary images, a morphological filter was applied to erode the strawberry rim and a flood- filling operation was carried out to overcome the threshold defects. Using this primary mask (binary image), the total area and the red area of each strawberry were calculated to get the percentage of red coverage. In the red area, colour features have been extracted to get information on Chroma and Hue angle needed to correlate them with the analytical data.

Destructive chemical analysis

Titratable acidity (TA) and pH using a semi-automatic titrator/pH meter (PH-Burette 24 - Crison Instrument, Barcelona, Spain), were measured on about 100 g of homogenized strawberries (for each class and replicate) as reported by Cozzolino *et al*. (2021). Similarly, the total soluble solids value (TSS) was determined using a digital refractometer (DBR35-XS Instruments, Carpi, Italy) and results were expressed in °Brix. The maturity index (MI) was calculated as the ratio of TSS and TA for each class (Melgarejo *et al*., 2017).

Respiration rate

The respiration rate (RR) of strawberries was determined at 4°C using a closed system as reported by Kader (2002). Thoroughly, each replicate, about 250 g of product, was put into a 3.6 L sealed plastic container to let CO_2 accumulate up to 0.1% as the CO_2 standard concentration. At regular time intervals CO_2 concentration was monitored until the reference value is reached. A Gas sample (1 mL) was drained from the headspace through a rubber septum and injected into the gas chromatograph (p200 micro-GC-Agilent, Santa Clara, CA, USA) equipped with dual columns and a thermal conductivity detector. Carbon dioxide was analysed with a retention time of 16 s and a total run time of 120 s on a 10-m porous polymer (PPU) column (Agilent, Santa Clara, CA, USA) at a constant temperature of 70°C. The RR was expressed as mL CO_2/kg h.

Statistical analysis

The data obtained were analyzed by multifactor ANOVA for $p \leq 0.05$ to evaluate the effects of the maturity class and harvest time (fixed factors) on pH, total soluble solid, titratable acidity, colour parameters and RR (variables). Parameters affected only by maturity class were subjected to a post-hoc test (Fisher), using Statgraphics (version 18.1.12, Warrenton, VA, USA).

A principal component analysis (PCA) was performed using the software Statistica version 6.0. (Statsoft Inc., Tulsa, OK, USA) (Jolliffe, 2022) with the aim of selecting the parameters able to discriminate the maturity classes. Based on the results obtained a Partial Least Square Regression (PLSR) was applied to develop a predictive method using the Unscrambler 10.0 software (CAMO Software, Oslo, Norway). In detail, 70% of the data was used in the calibration step and the remaining 30% was used to validate the obtained model.

3. Results and Discussion

Among the analytical data, RR and the MI were affected by the interaction of the two factors (maturity class x Harvest time) as reported in figure 2. MI showed all classes were different at first harvest with values of 6.61 (± 1.22), 7.38 (± 0.09) and 9.02 (± 0.84) for R0-25, R50-70 and R75-100, respectively; at the second harvest R75-100 reported higher value than the other samples, and this difference was measured also in the last two harvests (Fig. 2A). A similar trend was observed for the RR (Fig. 2B); in detail, at the first harvest time R0-25, R50-70 and R75-100 reported values of 5.12 (± 0.06), 7.73 (± 0.45) and 12.30 (± 0.45) mL CO_2/kg h, respectively. Then, at the second harvest, the RR increased at values around 10 mL CO_2/kg h for the fruit coming from R0-25 and R50-70 maturity classes, remaining almost constant in the last two harvests. On the other hand, for the full maturity class (R75-100), an increase in RR was found during the harvest time reaching at the last the values of about 20 mL CO_2/Kg h (Fig. 2B). Regarding the classes R50-70 and R75-100, similar results were found by Cozzolino *et al*. (2021) on strawberries (cv.

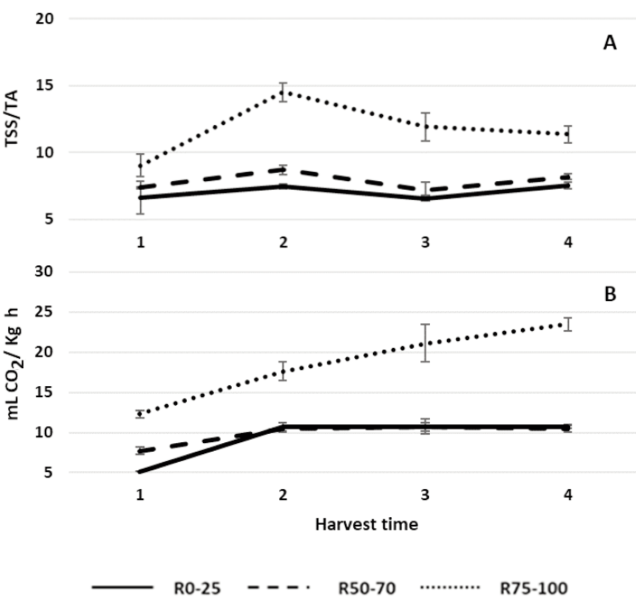
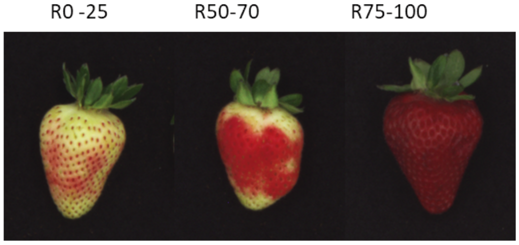


Fig. 2 - Effect of the maturity class and harvest time on maturity index (TSS/TA) (A) and respiration rate (mL CO_2/Kg h) at different harvest time (B).

Sabrosa) collected at two different ripening stages, namely half-red (in ripening phase, fully expanded and 50% red) and red (in ripening phase, fully expanded and 100% red) in three consecutive harvests.

The image processing allowed us to measure the percentage of red (Area red) and the colour parameters Chroma and Hue angle, which enabled the three classes' differentiation as indicated in figure 3.



	R0 -25	R50-70	R75-100
Area red %	21.73 ± 7.36 a	54.90 ± 13.02 b	100.00 ± 0.00 c
Hue angle (h°)	64.11 ± 14.80 c	40.32 ± 5.37 b	30.83 ± 3.57 a
Chroma	23.33 ± 1.16 a	27.55 ± 1.75 c	25.37 ± 4.12 b

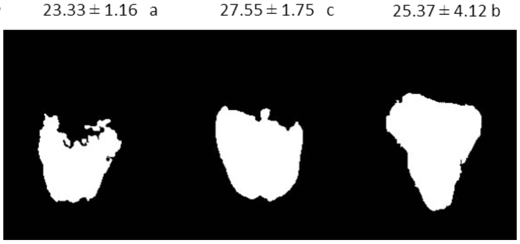


Fig. 3 - Strawberries image, mask, and correlated colour data (Area red , Hue angle and Chroma). Different letters indicate statistically significant differences according to the least significant difference (LSD) Fisher's test.

Since, harvest time affected only two quality parameters, data coming from the different harvests were collected and used for the multivariate analysis (PCA and PLSR). Regarding PCA, also confirming data of ANOVA analysis, Area red and Chroma were able to discriminate the maturity class on the first component, which accounted for 94% of the variability (Fig. 4A). Thus, these two parameters were used as predictors of maturity index TSS/TA, in a resulted PLS model, which showed R^2 of 0.7 in calibration and 0.6 in validation (Fig. 4B).

On the basis of these findings, an algorithm was developed using the Matlab® software, for the objective measurement of the percentage of red (Red Area) and the saturation level (Chroma) of a strawberry starting from the acquired images to automatically and non-destructively attribute the fruit to each class, applying the PLSR model.

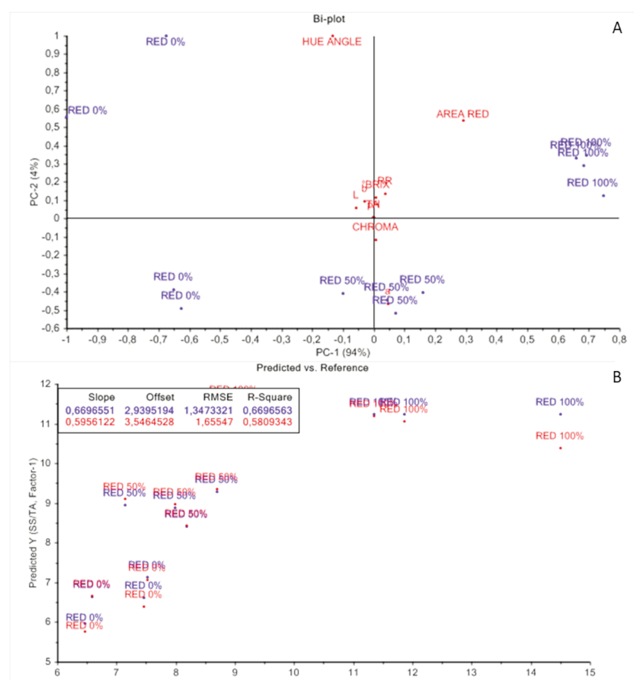


Fig. 4 - Principal Component Analysis (PCA) (A), and Partial Least Square Regression (PLSR) (B) model.

4. Conclusions

Results demonstrated that is possible to predict TSS/TA index starting by colour parameters extracted by IA on strawberry (cv. Sabrosa), collected in four consecutive harvests. The performance of the predictive model obtained might be improved by increasing the number of samples and extending the analysis also to other cultivars trying to build an algorithm available for handheld devices used in the field or in general for applications available to consumers to consciously buy the product.

Acknowledgements

The authors thank Giuseppe Sicuro and Gianluca Faliero of Apofruit Italia Soc. Coop. Agricola for the agronomy support during fruit harvest, Massimo Franchi of CNR-ISPA for the technical support, and Joseph Massaro for reviewing the English form of the paper.

This research was funded by the Project PON «R&I» 2014-2020-Azione II - "E-crops- Technologies for Digital and Sustainable Agriculture" funded by the Italian Ministry of University and Research (MUR) under the PON Agrifood Program (Contract ARS01-01136).

References

ANGELINI R., 2010 - *La fragola.* - Script, Bologna, Italy, pp. 548.

COZZOLINO R., PACE B., PALUMBO M., LAURINO C., PICARIELLO G., SIANO F., DE GIULIO B., PELOSI S., CEFOLA M., 2021 - *Profiles of volatile and phenolic compounds as markers of ripening stage in candonga strawberries*. - Foods, 10(12): 3102.

GONZALEZ R.C., WOODS R.E., EDDINS S.L., 2004 - *Digital image processing using MATLAB.* - Pearson Prentice Hall, Upper Saddle River, NJ, USA.

JOLLIFFE I.T., 2002 - *Principal component analysis*. - Springer Series in Statistics, 2nd ed. Springer, New York, NY, USA.

KADER A.A., 2002 - *Methods of gas mixing, sampling and* analysis, pp. 145-148. - In: KADER A.A. (ed.) *Postharvest technology of horticultural crops.* University of California Agriculture and Natural Resources, Oakland, CA, USA, pp. 535.

MELGAREJO P., LEGUA P., PÉREZ-SARMIENTO F., MARTÍNEZ-FONT R., MARTÍNEZ-NICOLÁS J. J., HERNÁNDEZ F., 2017 - *Effect of a new remediated substrate on fruit quality and bioactive compounds in two strawberry cultivars.* - J. Food Nutr. Res., 5(8): 579-586.

PALUMBO M., COZZOLINO R., LAURINO C., MALORNI L., PICARIELLO G., SIANO F., STOCCHERO M., CEFOLA M., CORVINO A., ROMANIELLO R., PACE B., 2022 - *Rapid and non-destructive techniques for the discrimination of ripening stages in candonga strawberries.* - Foods, 11: 1534.