Hyperspectral imaging (HSI) is a powerful analytical tool but the large amount of data typically generated during HSI acquisition can limit subsequent industrial applications. In order to extract the relevant information, it is possible to adopt the classical chemometric techniques for the identification of key wavelengths for a given quality control issue. This would allow the development of multispectral devices, intended for specific product control. Acerola is a typical Brazilian fruit that at the initial stages of ripening is green, changing to yellow-reddish colour and finally to purple when it is completely ripened. Due to its high moisture content, a rapid deterioration is commonly observed on ripening. This work is focused on the changes in post-harvest quality of Acerola fruits, with the aim of developing an easy way to visualize the global fruit status during ripening. An investigation of 20 samples for five consecutive days was carried out to highlight the most important wavelengths that characterized the maturity/senescent process of the Acerola fruit, investigating the evolution of acerolas mean spectra during time. The three wavelengths selected were 1883 nm, 1407 nm and 1136 nm, related to absorption and overtones of OH and CH bonds. Grey scale images at these selected wavelengths were combined to provide a false RGB image that allowed evaluation of the ripening process in a rapid and non-destructive manner. These results will facilitate the development of a low cost multispectral imaging system characterized by a simple image based output that could improve quality monitoring of acerola.

KEYWORDS: hyperspectral. Acerola, Malpighia emarginata, wavelengths selection, ripeness

ABSTRACT

Selection of NIR wavelengths from hyperspectral imaging data for quality evaluation of Acerola fruit

Cristina Malegori1*, Aoife Gowen2, Emanuel Marquez3, Maria Fernanda Pimentel4, Sergio Tonetto de Freitas4,5, Celio Pasquinì6, Ernestina Casiraghi1

1 DeFENS Department of Food, Environmental and Nutritional Sciences, Università degli Studi di Milano – ITALY
2 Biosystems Engineering, School of Agriculture, Food Science and Veterinary Medicine, University College Dublin – IRELAND
3 Departamento de Engenharia Química, Universidade Federal de Pernambuco, Recife (PE) – BRAZIL
4 Brazilian Agricultural Research Corporation, Embrapa Semiárido, Petrolina (PE) – BRAZIL
5 Instituto de Química, Universidade Estadual de Campinas, Campinas (SP) – BRAZIL

Corresponding author: cristina.malegori@unimi.it

INTRODUCTION

Hyperspectral imaging (HSI) is a non-destructive technique to explore surfaces in more detail than single point spectroscopy. Each pixel of a hyperspectral image, usually highly correlated to its neighbours, contains the spectrum of that specific position. Thus, the hyperspectral image contains both spatial and spectral information of a sample. The data are organized in three-dimensional blocks, called hypercubes. Many devices for acquiring hyperspectral images have been manufactured and there is an increasing interest for improving the data analysis techniques applied to such complex datasets1. Despite being one of the main advantages of hyperspectral systems, the large datasets routinely encountered in HSI, can complicate the extraction of useful information since much of the information obtained is redundant2. Multivariate data analysis, or Chemometrics, is a suitable approach to reduce the dimensionality of the data while retaining the most useful spectral information. A useful methodology that can reduce the data dimension is the variable selection. In variable selection wavelengths that would be most influential on product quality evaluation are selected, removing the wavelengths that have no or low discrimination power. In this way, the data dimension is reduced while preserving the most useful information3. The variables selected depend on the behaviour of spectral responses under modification of the samples and on differences among them.
Fruit quality is defined by a series of external characteristics that make the product more or less attractive to the consumer, including suitability to be eaten as fresh or stored for reasonable period without deterioration\(^4\). Fruit quality could be considered as a multivariate concept encompassing the physical, physiological, nutritional, and pathological attributes that affect shelf life\(^3\). The ripe phenotype is the summation of biochemical and physiological changes that occur at the terminal stage of fruit development, rendering it edible and desirable to seed-dispersing animals. These changes, although variable among species, generally include modification of cell wall ultrastructure and texture, conversion of starch to sugars, increased susceptibility to post-harvest pathogens, alterations in pigment biosynthesis and accumulation, and heightened levels of flavour and aromatic volatiles\(^5\).

Nowadays, fruit are managed manually or automatically on the basis of external quality features but the high risk of human error in the classification process has been underlined as one of the most important drawbacks that machine vision can help preventing\(^6\). Vision systems for fruit sorting were traditionally based on video cameras working in the visible wavelength range, limited to obtain information on the external aspect like colour or damages presence. More information about sample composition might be obtained by computer vision, which can acquire a set of optimised monochromatic images at few selected wavelengths, and can make possible to estimate or discover features, difficult to uncover with traditional vision systems like dry matter, total soluble solids and acidity\(^7\). The use of NIR spectroscopic information and hyperspectral imaging as a non-destructive measurements of quality attributes has been implemented only recently in the post-harvest field (https://www.tomra.com/en; http://www.aweta.nl/it).

Acerola (Malpighia emarginata, Junco cultivar) is a fruit native to Central America and Northern South America, with some of the largest plantings occurring in Brazil, and it has recently been introduced in subtropical areas throughout the world, including Southeast Asia, India and South America. The plant provides flowers and fruits at different stages and, consequently, long fruiting periods are observed during the year. The fruit presents a short postharvest shelf-life (2 to 3 days) at room temperature\(^8\). The ripening process of acerola fruit involves a succession of complex biochemical reactions including starch hydrolysis, conversion of chloroplasts into chromoplasts with the transformation of chlorophyll and the production of carotenoids, anthocyanins, phenolics and volatile compounds that result in a colour change\(^9\). It has been increasingly recognized that in addition to vitamin C acerola contains other functional constituents such as carotenes, thiamine, riboflavin, niacin, proteins, and mineral salts, mainly iron, calcium and phosphorus which are beneficial to human health\(^10\). For these reasons Acerola is considered a super-fruit, nevertheless research on this Brazilian fruit is very limited and it has not yet been studied using non-destructive techniques for quality estimation. Furthermore, its biochemical evolution is not completely known and its perishability is very rapid, preventing the export of the fresh fruit. Table 1 shows the chemical composition of unripe and ripe acerolas.

The aim of this work is to lay the foundation for the development of a multispectral camera for the quality evaluation of acerola fruit at different ripeness stages. This overarching aim was achieved by meeting two specific objectives:

1. to evaluate the differences in hyperspectral images of acerolas at different maturity stages for the selection of useful wavelengths;
2. to develop a simple imaged based approach to allow global, rapid and non-destructive evaluation of the maturity/senescence degree of acerolas.
Table 1: Chemical composition of acerola juices at different stages of maturity (Righeto, 2005)

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Immature</th>
<th>Mature</th>
</tr>
</thead>
<tbody>
<tr>
<td>pH</td>
<td>3.62</td>
<td>3.40</td>
</tr>
<tr>
<td>Total solids (g/100 g juice)</td>
<td>5.4 ± 0.2</td>
<td>5.5 ± 0.02</td>
</tr>
<tr>
<td>Soluble solids (Brix)</td>
<td>5.1 ± 0.0</td>
<td>5.7 ± 0.0</td>
</tr>
<tr>
<td>Sugars (g/100 g of juice)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fructose</td>
<td>2.14 ± 0.04</td>
<td>3.33 ± 0.02</td>
</tr>
<tr>
<td>Glucose</td>
<td>0.99 ± 0.04</td>
<td>0.88 ± 0.02</td>
</tr>
<tr>
<td>Sucrose</td>
<td>nd</td>
<td>0.02 ± 0.01</td>
</tr>
<tr>
<td>Acids (g/100 g juice)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Malic</td>
<td>0.25 ± 0.04</td>
<td>0.38 ± 0.03</td>
</tr>
<tr>
<td>Citric</td>
<td>0.012 ± 0.001</td>
<td>0.003 ± 0.003</td>
</tr>
<tr>
<td>Tartaric</td>
<td>0.01 ± 0.00</td>
<td>0.002 ± 0.000</td>
</tr>
<tr>
<td>Ascorbic</td>
<td>1.85 ± 0.02</td>
<td>0.80 ± 0.02</td>
</tr>
<tr>
<td>Vitamin C</td>
<td>1.9 ± 0.028</td>
<td>0.97 ± 0.031</td>
</tr>
<tr>
<td>Total phenolics (mg catechin/g juice)</td>
<td>3.8 ± 0.02</td>
<td>1.35 ± 0.01</td>
</tr>
<tr>
<td>Composition (% dw basis)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sugars</td>
<td>58.5</td>
<td>75.55</td>
</tr>
<tr>
<td>Acids</td>
<td>39.3</td>
<td>21.6</td>
</tr>
<tr>
<td>Vitamin C</td>
<td>35.18</td>
<td>17.54</td>
</tr>
<tr>
<td>Total phenolics (mg catechin/g juice)</td>
<td>7.0</td>
<td>2.4</td>
</tr>
</tbody>
</table>

MATERIALS AND METHODS

Hyperspectral images of 20 acerolas were acquired for 5 consecutive days, giving a total of 100 images. The fruits were selected at different initial ripeness degrees, based on the color of the peel, and were stored at room temperature (25 ± 2°C). According to literature, green was chosen as the initial stage, yellow as the middle and red as the last stage of maturity. A SisuChema Hyperspectral Imaging System (900 - 2500 nm) was used, equipped with a 50 mm lens with a minimum spatial resolution of 150 μm. One of the main aspects to be taken into account is the spherical shape of acerolas. In this case, two problems arise: the presence of bright spots caused by the reflection of light and the progressive darkness of the borders, both caused by the effect of Lambert’s cosine law. To elaborate the huge amount of data and try to overcome the light artefacts, a chemometric approach was applied. A binary mask was first created to produce an image containing only the fruit, avoiding any interference from the background. For this task, the image at the wavelength in which the fruit appeared opaque, compared with the background, and can be segmented easily by simple thresholding, was selected (1111 nm). The spectra of the fruit (without background) were pre-processed using automatic baseline correction (SNV) and multiplicative scatter correction (MSC) and then the data were smoothed (Savitzky-Golay smoothing, 11 points) and mean centred. Other pre-processing treatments, such as first and second derivatives were tested, but the results were not improved; no outliers, dead pixels or spikes were detected in the hyperspectral images. An investigation of the change in acerolas mean spectra during time was carried out to highlight the most important three wavelengths that characterized the ripeness process of the acerola fruit. This was done by subtracting the mean spectrum at day 1 from all subsequent mean spectra, for a given fruit. Consequently, for each fruit, the images at the three wavelengths with the widest variation along time, were plotted in grey scale and combined to make a false RGB image allowing an intuitive understanding of the qualitative status of each fruit.

RESULTS

Figure 2 shows the mean spectra of all the fruits analysed during the five days shelf life. Simple visual approaches to evaluate the changes in spectra were not satisfactory since did not highlight remarkable differences. For this reason, the mean spectrum of each fruit at day 1 was subtracted from all subsequent images of the same fruit (day 2, 3, 4, 5), to identify the main changes, as shown in Figure 2. As it can be seen in the right part of figure 2, the spectrum of the initially less ripe acerola does not change much during time, while from the image of the initially more ripe acerola (that in five days became senescent) it is possible to highlight the most important three wavelengths that characterized the acerola spectra. The three wavelengths, that appear to exhibit consistent and large changes with time, were 1883 nm, due to the second overtone of the C=O double bond, 1407 nm,
the absorption band of the CH first overtone combination and first OH overtone, and 1136 nm characterized by
the second overtone of CH bond. It difficult to interpret why these bands are changing with ripening due to the
biochemical complexity of the ripening process that as previously said, has not been widely studied. To describe
acerola using these most informative wavelengths, a RGB false image was created using the selected wave-
lengths as the red, green and blue channel.

Figure 2: mean spectra of all the fruits analysed during the five days shelf life with a zoom on acerola number 1
and acerola number 20

As an example, the false colour images of the less ripe acerola number 1, and the more ripe, acerola number
20, (that in five days passed from ripe to overripe), are presented in Figure 3 allowing monitoring of the ripening/
senescence process. It is evident that the changes of the peel colour in acerola number 1 during ripening are
not well represented in the false images while the senescence/degradation, which occurs on acerola number 20,
is well-highlighted (figure 3).

Figure 3: Image of acerola 2 and acerola 20 at those selected wavelengths were combined to produce an RGB
false image that was fallowed during 5 days

DISCUSSION AND CONCLUSION

The methodological approach for the wavelengths selection allowed the initial aim to be reached: the evaluation
of Acerola post-harvest quality during ripening/senescence process, in a view of developing a simple multi-
spectral device for qualitative evaluation. During the work, unexpected variability of the acerola NIR profile was
highlighted during time. In fact, a colour change from green to red, which is considered as a ripeness evolution
index, does not correspond to a modification of the fruit composition, commonly reflected in NIR spectra. Spectral
modification, that allowed variable selection, appears more evident when acerola colour turns from red to
purple-brown during the degradation process of the fruit. As a matter of fact, the distinction between ripening
and senescence has never been finely drawn\textsuperscript{13} so the phenomena that we highlighted with the hyperspectral
approach could be defined as overripe or initial stage of senescence. Due to the poor knowledge about acerola,
a step back is necessary in order to well delineate the biochemical evolution during ripening. First of all, it will
be important to focus on ethylene production and cellular respiration defining if acerola could be classified as a climacteric fruit. So far, there is no evidence in literature, but if acerola could be considered a climacteric fruit, like banana, apple and mango, it could be harvested unripe and easily exported to other countries. This will be of paramount importance and will widely increase the potential market of this super-fruit.

Acknowledgments
The corresponding author thanks the ICNIRS for the John Shenk Travel Grants and the SISNIR for funding the participation to the NIR 2015 Conference.

References