Elsevier Editorial System(tm) for Innovative Food Science and Emerging Technologies
Manuscript Draft

Manuscript Number:

Title: Adaptive self-configuring computer vision system for quality evaluation of fresh-cut radicchio

Article Type: Research Article

Keywords: Computer vision system, non-destructive quality evaluation, self-configuration, automatic colors and features selection, image analysis.

Corresponding Author: Dr. Maria Cefola, Ph.D

Corresponding Author's Institution: ISPA

First Author: Bernardo Pace

Order of Authors: Bernardo Pace ; Dario P Cavallo; Maria Cefola, Ph.D; Roberto Colella; Giovanni Attolico

Abstract: An innovative Computer Vision System (CVS) that extracts color features discriminating the quality levels occurring during fresh-cut radicchio storage in air or modified atmosphere packaging was proposed. It self-configures the parameters normally set by operators and completely automates the following steps adapting to the specific product at hand: color-chart detection, foreground extraction and color segmentation for features extraction and selection. Results proved the average value of a* over the white part and the percentage of light white with respect to the whole visible surface to be the most discriminating color features to significantly separate ($P \leq 0.05$) the three desired quality levels (high, middle and poor) occurring during fresh-cut radicchio storage (whose true value was verified on the base of ammonium content and human evaluated visual quality). The proposed procedure significantly simplify the CVS design and the optimization of its performance, limiting the subjective human intervention to the minimum.

Consiglio Nazionale delle Ricerche
ISTITUTO DI SCIENZE DELLE PRODUZIONI ALIMENTARI

D. Knorr

Editor

Innovative Food Science and Emerging Technologies

Bari, 09/07/2015

Dear Prof. Knorr,

I am pleased to submit the paper entitled "**Adaptive self-configuring computer vision system for quality evaluation of fresh-cut radicchio**" by Pace Bernardo, Cavallo Dario Pietro , Cefola Maria, Colella Roberto & **Attolico Giovanni** for publication in the journal **Innovative Food Science and Emerging Technologies.**

Please find my complete address below:

Maria Cefola Ph.D.

Consiglio Nazionale delle Ricerche C.N.R.

Istituto di Scienze delle Produzioni Alimentari

Via G. Amendola 122/O, 70126 Bari - ITALY

Phone +39 080.592.9304

e-mail maria.cefola@ispa.cnr.it

Sincerely,

Maria Cefola

Industrial relevance

The non-destructive quality control represents a valuable tool to monitor fruits and vegetables along the whole chain from production to the end-user. Increased consumers' satisfaction and reduction of waste are just two examples of benefits that can come from a frequent and consistent control of food. CVS represent the most powerful and flexible way to reach these results. The current state-of-the-art make their design strongly related to the specific product at hand. Thresholds and features are characteristics that play a critical role in determining the final performance of system but are generally set by designers or operators using a–priori knowledge and/or trial-and-error processes. The proposed innovative procedure allows the CVS to self-configure most of these parameters and to easily adapt to different products regardless of the number and kind of colors associated to their surface. It results of practical applications in food processing, providing a non-destructive, automatic, cheap, fast and simple technology for the quality level evaluation, whose configuration requires a reduced, less critical and less technical human intervention.

Highlights

1. Computer Vision System for the no-destructive and automatic color analysis
2. Automatic color-chart location and separation of foreground from background
3. Automatic identification of color features able to evaluate product quality
4. Light White percentage discriminated fresh-cut radicchio quality levels
5. CVS was able to identify quality of fresh-cut radicchio stored in AIR or pMA

Adaptive self-configuring computer vision system for quality evaluation of fresh-cut radicchio

Pace Bernardo [a1], Cavallo Dario Pietro [b1], Cefola Maria [a1]*, Colella Roberto[b], Attolico Giovanni [b]

[a] Institute of Sciences of Food Production, CNR-National Research Council of Italy Via G. Amendola, 122/O – 70126 Bari, Italy

[b] Institute on Intelligent Systems for Automation, CNR-National Research Council of Italy Via G. Amendola, 122/O – 70126 Bari, Italy

* Corresponding Author: phone/fax: +39.080.5929304/9374; e-mail address: maria.cefola@ispa.cnr.it

[1] **First Authorship is equally shared**

Abstract

An innovative Computer Vision System (CVS) that extracts color features discriminating the quality levels occurring during fresh-cut radicchio storage in air or modified atmosphere packaging was proposed. It self-configures the parameters normally set by operators and completely automates the following steps adapting to the specific product at hand: color-chart detection, foreground extraction and color segmentation for features extraction and selection. Results proved the average value of a^* over the white part and the percentage of light white with respect to the whole visible surface to be the most discriminating color features to significantly separate ($P \leq 0.05$) the three desired quality levels (high, middle and poor) occurring during fresh-cut radicchio storage (whose true value was verified on the base of ammonium content and human evaluated visual quality). The proposed procedure significantly simplify the CVS design and the optimization of its performance, limiting the subjective human intervention to the minimum.

Industrial relevance

The non-destructive quality control represents a valuable tool to monitor fruits and vegetables along the whole chain from production to the end-user. Increased consumers' satisfaction and reduction of waste are just two examples of benefits that can come from a frequent and consistent control of food. CVS represent the most powerful and flexible way to reach these results. The current state-of-the-art make their design strongly related to the specific product at hand. Thresholds and features are characteristics that play a critical role in determining the final performance of system but are generally set by designers or operators using a–priori knowledge and/or trial-and-error processes. The proposed innovative procedure allows the CVS to self-configure most of these parameters and to easily adapt to different products regardless of the number and kind of colors associated to their surface. It results of practical applications in food processing, providing a non-destructive, automatic, cheap, fast and simple technology for the quality level evaluation, whose configuration requires a reduced, less critical and less technical human intervention.

Keywords: Computer vision system, non-destructive quality evaluation, self-configuration, automatic colors and features selection, image analysis.

1. Introduction

In the last years, due to the increased consumers' requests for quality, the food industry has paid great attention to measure and control the visual appearance of products (Wu & Sun, 2013a; Zhang et al., 2014). As regards fruits and vegetables, visual appearance is normally strongly related to overall quality and is therefore a very useful method to judge the quality level. In particular, in these products, accurate measurements of some chromatic properties (such as intensity, hue, chroma, color uniformity) provide important indicators of visual appearance. Generally, bright and vivid colors are associated with freshness and better quality, while dull colors are a normal consequence of an overall quality loss (Nunes, 2015). It has been proved that specific color features are correlated with quality level of fruits and vegetables. Pace et al. (2011, 2014a) reported a strong relationship of visual appearance with and $b*$ (or Chroma) in fresh-cut nectarines and with brown pigments in lettuce. Similarly, Nunes (2015) found significant correlations between color parameters and visual appearance in many fruits and vegetables. Generally, the color parameters related to visual appearance are manually acquiring by colorimeter in a proper set of sample points on the product's surface. Recently, there is an increasing attention to the use of computer vision systems (CVS). CVS are more vulnerable to instabilities of the acquisition environment than colorimeters but, if properly calibrated and used, proved to be more robust, since they can analyze the complete surface of the product and are not affected by the arbitrary and subjective choice of sampling points (Pace et al., 2011, 2013, 2014a; Wu & Sun 2013a; Zhang et al., 2014; Manninen et al., 2015). CVS can extract from an image several features including color histograms, presence and size of defective areas and color texture (Hernández-Carrión et al., 2015). However, it is important to highlight that, in several cases, only specific color features, are well related to the quality (Pace et al., 2014a).

Starting from these considerations, the automatic identification of the most discriminative colors by CVS could improve the non-destructive quality level evaluation of fruits and vegetables, reducing the error due to subjective identification of the most discriminating colors (and of related parameters) of the overall quality level. Automatic quality classification, based on image analysis was reported by Zhang et al., (2014). In addition, Kordecki & Palus (2014) reported an automatic color-chart detection algorithm that determines the type and locations of color-patches in images. Color-charts are placed in the scene to provide reference colors that enable the evaluation and the reduction of effects of acquisitions environment on the behavior of the CVS. The algorithm, proposed by Kordecki & Palus (2014), used k-means clustering to accomplish color quantization and segmentation of a collection of regions some of which are the color-chart's patches. Further criteria, inspired by the a-priori knowledge about the structure of the color-chart (shape, number, position, shape and size of patches) were used to select and to identify the color-chart. Recently, Zhang et al. (2015) proposed a CVS to automatically detect defective apples by combining background removal, automatic correction of lightness, counting defected regions and a relevance vector machine (RVM) classifier. The system used a camera that acquires RGB and NIR (Near InfraRed) registered images of the same scene. The binary mask of the foreground was obtained using a (experimentally identified) threshold on the NIR image. A morphological filling was therefore applied to increase the quality of the segmented region. A lightness correction was applied in the peripheral part of the apple to compensate for the change of geometry between light, camera and surface. Moreover, Avila et al. (2015) proposed an automatic method to construct olives' and grape seeds' color scales maturity: the system used an illumination invariant color model and the threshold for foreground segmentation was determined using the Otsu's method (Otsu, 1975). However, most of the cited literature is related to products characterized by mostly homogeneous color. The evaluation of color change in products characterized by nonuniform color is still a critical task (Balaban, 2008; Wu & Sun; 2013a).

Starting from these findings, in this paper an automatic procedure for the non-destructive extraction by CVS of color features was proposed and applied for the evaluation of the quality level during storage of the fresh-cut radicchio, characterized by nonuniform color.

2. Materials and Methods

2.1. Plant material and processing

Radicchio di Chioggia heads (*Cichorium intybus* L. group *rubifolium*) belonging to two hybrids (Corelli and Botticelli), were provided in two different harvest times by a farm located in Pontecagnano (south Italy) and immediately transported in cold condition to the Postharvest

Laboratory of the Institute of Sciences of Food Production. Heads of each hybrid were selected to discard damaged samples and processed as fresh-cut product. In detail, radicchio heads were prepared by removing and discarding wrapper leaves and the stem with sharp stainless steel knives. Radicchio pieces (3 × 4 cm), obtained by using a vegetable cutter (CL52 Robot Coupe, Vincennes-Cedex, France), were pooled and blended to minimize product heterogeneity. Radicchio pieces were washed in tap water at 4 °C for 4 min. After washing, pieces were dried using a manual centrifuge and packaged in polypropylene bags (25 × 30 cm, 30 μm, Carton Pack, Rutigliano, Italy) containing each one about 150 g of product either sealed in passive modified atmosphere (pMA, at equilibrium: 10% O_2 and 7% CO_2) or in unsealed bags (AIR). In total, 120 bags (two radicchio hybrids x six replicates × five quality levels x two packaging condition) were prepared and stored at 4 (± 1.0) °C. All items, at proper times during storage, were graded using a five quality level scale, based on sensory evaluation, as reported below. Images of samples belonging to each quality level were acquired and processed by Computer vision system (CVS); moreover the same samples underwent a chemical analysis for the ammonium content (NH_4^+).

2.2. Quality level classification and NH_4^+ analysis

Along the storage, fresh-cut radicchio were classified using five quality levels (QL) according to the following scale: 5 = very good (very fresh, no signs of wilting, decay or bruises), 4 = good (slight signs of shrivelling, bruises), 3 = limit of acceptability or marketability (moderate signs of shrivelling, browning, dryness, wilting, bruises), 2 = poor (severe bruises, evident signs of shrivelling, pitting, decay), and 1 = very poor (unacceptable quality due to decay, bruises, leaky juice). The QL 3 was considered the minimum threshold of acceptance for sale or consumption (Nunes et al., 2009).

For NH_4^+ analysis the method reported by Pace et al. (2014a) was used. In detail, five g of chopped sample was extracted in distilled water and, after the reaction with nitroprusside reagent and alkaline hypochlorite solution, color development was determined after incubation at 37 °C for 20 min, reading the absorbance at 635 nm (UV-1800, Shimadzu, Kyoto, Japan).

2.3. Color analysis by computer vision system (CVS)

2.3.1. CVS description and setup

The images used in the experiments were acquired using a 3CCD (Charged Coupled Device) digital camera (JAI CV-M9GE). The camera has a dedicated CCD for each color channel and provides a reliable color measure at full resolution, without the artifacts of most digital cameras (based on the Bayer filter). The uncompressed TIFF format was used to save images while avoiding any color artifacts produced by lossy compression algorithms. The camera mounted a Linos MeVis 12 mm lens system whose optical axis was perpendicular to the black background onto which the products were placed. Eight halogen lamps, divided along two rows placed at the two sides of the imaged area, were used. They were oriented using a 45° angle with respect to the optical axis of the CCD camera and to the plane on which the products were placed (Pace et al., 2014a).

2.3.2. Color-chart detection

A small color-chart (X-Rite Color Control Patches, Fig. 1) has been placed in the scene during the acquisition of each image. Its patches allow the estimation and reduction of color variations occurring in time due to the behavior of CVS components and to uncontrollable changes of environmental conditions. X-Rite releases the numerical values associated to the color of each patch in the CIE $L^*a^*b^*$ space. To use this information, the detection of the color-chart and the identification of each patch are needed when processing each image: to make these two steps automatic strongly simplify the processing of images, without requiring human intervention. The implemented method looked for two achromatic patches (the white 19th and the light grey 20th) using an iterative process that considered color, shape and size. When satisfactorily results were found, their central positions were used to look for the position of the 13^{th} patch, placed at 90 degrees with respect to the vector connecting the white and the light grey patches. Then all the other patches were identified in a straightforward way using the knowledge about the structure of the color chart. The measured $L^*a^*b^*$ values and the expected ones (supplied by X-Rite) were used to estimate and correct as much as possible any difference induced by environmental conditions and sensor characteristics. The white patch was also used to white-balance the image: a correction coefficient was evaluated for each band to reduce the distance between the measured white and the reference one. Each coefficient was therefore applied to the corresponding band over the whole image. In Fig. 2 some steps of this procedure are showed. The source image and the automatically detected 19^{th} and 20^{th} patches are shown in Fig. 2A and 2B respectively; Fig. 2C shows how all the patches are isolated to measure the corresponding color. Finally, in Fig. 2D, the source image in which the detected color-chart has been removed is reported. The detection and extraction of color-chart does not make any assumption about its position and attitude inside the scene and can be fruitfully used in more general applications and different contexts.

2.3.3. Foreground extraction

After the detection (and subsequent removal) of the color-chart, the separation of the product at hand (foreground) from the image background was needed. In our previous work, such separation was achieved using a threshold fixed experimentally (Pace et al., 2013). The new version of the CVS automatized this step. For each pixels (p_{ij}) of the image, the Euclidean distances with respect to its neighbors in an 8-neighborhood was evaluated using the saturation channel in the HSV (Hue Saturation Brithness) color space; then the histogram of these distances over the whole image was built. A typical result is shown in Fig. 3. Two Gaussian distributions were fitted to the histogram: the first (denoted fit_1 and associated to homogeneous regions) with a peak around 0 is mainly related to regions of the image such with small variations in color as the background. The second (denoted fit_2 and associated to edges) is due to strong color variations occurring at the edges that separate foreground and background. Considering fit_1 and fit_2 to be two Gaussian distributions (with means μ_1 and μ_2 and variances σ_1 and σ_2 respectively, fit_2 was obtained excluding the data in the histogram lower than $\mu_1 + 2*\sigma_1$. Finally, $\mu_2 - \sigma_2/2$ was used as the threshold selecting the strong edges between foreground and background, where the correction $\sigma_2/2$ was introduced to account for the effects of possible noise caused by pixels that were close to edges. This conservative choice was done since loosing some foreground pixels was less dangerous than including by mistake background pixels in the foreground region (Fig. 4). Fig. 4 shows this process; in particular, in Fig. 4A-B the results obtained by applying a quite high (fixed) threshold on the three color channels are showed: only the pixels whose color components were all above the threshold were set to foreground. This very conservative choice was independent from product and acquisition environment but separated only a quite small part of the product's surface, assigning to the background all the pixels having even a single color channel below the specified value. This preliminary result collected only pixels certainly belonging to the product surface: it was used as seed for the following growing schema. These original regions were extended by including all the neighboring pixels whose colors differed from the pixels already classified as foreground by a value lower than a threshold. The value of this threshold was fixed using a bottom-up, data-driven approach that automatically adapted to different products and images. The result (Fig. 4 C-D) shows that some small areas of product were erroneously classified as background. The proposed method did not apply any morphological processing to fill these gaps and holes.

2.3.4. Color segmentation

Partitioning of an image into disjoint and homogeneous regions with respect to characteristics of interest is often necessary to accomplish useful measurements with a CVS. In this work, the identification and separation of the two relevant colors (white and red) of fresh-cut radicchio was needed to measure their changes separately. A hierarchical clustering technique was used to segment colors: this technique does not require any initialization and builds a structure that can be used to dynamically select different resolutions of segmentation according to the requirements of the application at hand.

During the processing of each image, the colors expressed in the RGB space of all the pixels belonging to the foreground were converted into the CIE $L^*a^*b^*$ color space, chosen for its perceptual uniformity and device independency (Kang et al., 2008). The channel L^* (representing lightness), as known is less reliable. The other two channels were quantized (using mathematical rounding) to reduce the number of possible a^* and b^* pairs. A two dimensional histogram such as the one shown in Fig. 5 was produced, where the third dimension represents the percentage of occurrences of the corresponding (a^*, b^*) pair in the foreground region. The histograms of all the available images were therefore analysed to compute the variance of the number of occurrences at every single (a^*, b^*) pair over the whole dataset. A predefined number of (a^*, b^*) pairs were therefore selected choosing the ones with larger variance. This step was required to reduce the number of colors to a value that made the computational load manageable by the hierarchical clustering. As a side effect, less relevant colors (potential outliers) were cancelled by this approach. The selected (a^*, b^*) pairs were used to build a dendrogram (based on the Euclidean distance). This structure expresses a hierarchical subdivision of colors that can be cut at different levels to obtain different number of clusters (different color quantizations). A first subdivision identified two (a^*, b^*) pairs that separated the white pixels (W) from the red ones (R), producing an image segmentation such as the one shown in Fig. 6A-B.

Using a second subdivision, the W and R regions were further divided in two components. In detail W and R were divided in a dark (W1 or R1) and a light (W2 or R2) component. In the Fig. 6C the second segmentation on the W component is reported. Average values of L^*, a^*, b^*, on these regions and their percentage with respect to the product's surface were considered as potential features to evaluate the quality level of fresh-cut radicchio.

2.4. *Statistical analysis*

The effects of QL on the color features extracted by CVS on the fresh-cut radicchio samples were evaluated performing a one-way ANOVA with data means arranged in a completely randomized design. The mean values for QL were separated using the Student-Newman-Keuls (SNK) test. T-test was performed by using Statistica 6.0 (StatSoft, 2001).

3.0. RESULTS AND DISCUSSION

3.1. Advantages of proposed method in pre-processing phases and automatic color segmentation

The procedures proposed in this paper brought advantages to several steps of the image processing chain: color-chart detection and extraction, foreground separation and color-based product's segmentation for product classification. When processing each image, in the first step the color-chart was automatically detected, extracted and analyzed. The proposed method did not make any assumption on position and orientation of the color-chart in the scene and can be applied in any context in which it is useful to use a color reference in the image. The second step of image processing was to separate the product (foreground) from the background, including in the foreground as much product as possible and avoiding misclassification of background pixels as parts of the foreground. In fact, the latter error could affect the statistical measures made on the product and reduce the reliability of the CVS. All the methods that separates foreground and background rely on some thresholds on features that can be related to every pixel individually (grey-level, color channels, NIR data, etc) or that can be associated to properties of neighboring pixels (distance of grey-levels or colors, gradients, texture properties, etc). Setting these thresholds often involves human intervention and trial-and-error phases (Wu & Sun, 2013b). In this paper the threshold to separate the region of interest was automatically determined using a fully data-driven approach. By applying, this approach, some small areas of product were erroneously classified as background. However the experimental results showed that missing these tiny regions did not affect quality and robustness of results. Instead, the erroneous inclusion of background regions in the evaluation of statistical measures would have produced larger effects on the stability of results. The proposed procedure was applied to the fresh-cut radicchio, characterized by a nonuniform color, so the product surface needed to be segmented into its most significant parts. Indeed, in the case of products with nonuniform color, the separation of the whole surface into regions associated to relevant colors is very important (Balaban, 2008) but, at the date, no well assessed and no robust solutions are available (Wu & Sun, 2013a). In this case, the hierarchical clustering-based approach identified two macro areas: the mostly white one (W) and the red one (R). The proposed approach did not required human intervention to initialize clusters or to set the desired number of colors. The hierarchical structure was produced without such a-priori information and the best cutting level of the dendrogram was decided on the base of the performance reached by the produced color quantization in the classification and estimation phases.

3.2. Color feature selection for quality level evaluation of fresh-cut radicchio during storage in air and modified atmosphere by applying the proposed automatic procedure.

In fresh-cut radicchio, the analysis of the NH_4^+ content allowed to improve the sensory evaluation of QL, discriminating significantly three class of quality: high (ranging from QL=5 to QL=4), middle (from QL =3 to QL =2) and poor (QL=1) (Table 1). As previously widely reported (Pace at al., 2014a; Cefola et al., 2015) NH_4^+ is a chemical objective indicator of fruits and vegetable quality, useful to standardize the sensory evaluation of visual quality performed through rating scale (Pace et al., 2014b). The color features (L^*, a^*, b^*) obtained from the entire image of fresh-cut radicchio, using the CVS significantly affected the QL. However, only two classes were identified by L^* parameters (QL= 5 and 4 from QL =3, 2, 1), while a^* allowed to discriminate only fresh samples (QL=5) from the others quality levels. Finally, a no adequate discrimination was performed by b^* parameter (Table 1). Thus, working on the entire images (without segmenting regions with different colors) produced an insufficient or improper QL classification. This is due to the nonuniform color of fresh-cut radicchio, that makes difficult the quantitative analysis of color (Wu & Sun, 2013a). The application of the fist segmentation to the entire images of fresh-cut radicchio, allowed to obtain two color component, named W or R. The color features (L^*, a^*, b^*) extracted from each component, affected significantly the QL (Table 1).

Considering the W component, the L^* parameter allowed to differentiate two class of quality (QL= 5 and 4 from QL =3, 2, 1) as previously reported considering the color parameters extracted from the non segmented images. The information provided by the lightness channel L^*, as verified in previously works (Pace et al., 2011; 2013; 2014a), is less reliable: it strongly depends on the geometry between observed surface, lights and sensor. The color parameter a^* allowed the best discrimination of the QL in three quality class, matching the results obtained by NH_4^+ (Table 1). An inadequate discrimination was performed by the b^* parameter (Table 1). The color parameters obtained from the red component showed significant changes mainly associated to the variability among samples and not related to visual quality loss. Indeed, during storage, the loss of visual quality of fresh-cut radicchio is principally due to oxidative phenomena occurring on the white component as occur in different fresh-cut vegetables (Wojciech et al., 2014).

The second segmentation, allowed to further split in two more color components both the W and R parts of fresh-cut radicchio images. On the W part two components were automatically selected: a dark white, W1, and a light white, W2. The L^*, a^* and b^* mean values of the colors associated to each part were respectively: 56.97 (± 3.66), 2.60 (± 0.57) and 10.45 (± 0.45) in W1 and 60.23 (± 3.73), 2.04 (± 0.55) and 4.75 (± 0.62) in W2. The color parameters extracted from W1 and W2 resulted significantly different for $P < 0.001$ (n=120), performing a *t*-test.

The percentages of W1 and W2 components were calculated with respect to the white part (W) and with respect to the entire surface (I) of radicchio samples: if N^{W1}, N^{W2}, and N^{W} denote the number of pixels corresponding to the W1, W2 and W regions, these percentage were N^{W1}/N^{W} and N^{W2}/N^{W}. In the same way N^{W1}/N^{I} and N^{W2}/N^{I} represent the percentages of W1 and W2 with respect to the whole surface of product (composed by N^{I} pixels). These percentages significantly affected the QL (Table

2). In detail, N^{W1}/N^{W} allowed to separate only samples with QL=5 from the other QL; while N^{W1}/N^{I} was able to separate samples with QL=5 from radicchio pieces belonging to QL of 4, 3 and 2. Unexpectedly, N^{W1}/N^{I} provided the same values for QL = 1 and QL=5 . This could be due to the similarity between very dark red and the background: in fact the number of pixels dark white that moves from QL=5 to QL=1 (due to browning) could be compensated by the pixels that become so dark to be confused with the background.

N^{W2}/N^{W} allowed to discriminate only QL=5 samples from the other QL; whereas N^{W2}/N^{I} , provided a good and significant discrimination of QL, matching the quality classes selected by NH_4^+ (Table 2). Thus, the percentage of W2 (light white) tends to decrease from QL =5 to QL = 1, discriminating the fresh-cut radicchio quality levels. Indeed, color is used as an indicator of quality for many fruits and vegetable (Wu & Sun, 2013a) and bright and white color of fresh-cut radicchio samples is considered as a freshness indicator by consumers. During storage, biochemical phenomena (such as enzymatic and not enzymatic oxidation), also induced by cutting (Kader, 2002), cause color change which affect the radicchio visual appearance: the main color changes regard the white part, which tends to become darker. Similarly, in fresh-cut iceberg, the main color changes affecting consumer assessment of visual quality were based on the white parts which turns to brown during storage (Pace et al., 2014a).

Considering separately the samples stored in AIR or pMA at the end of storage, the application of the automatic procedure proposed showed that a^* of the W part (the complete white part of the product) and the percentage N^{W2}/N^{I}, as the most discriminating color features of QL in fresh-cut radicchio (Table 1 and 2). At the end of the storage, fresh-cut radicchio stored in AIR showed an a^* mean value of 2.98 (± 0.36) significantly different ($p< 0.001$) from the mean value obtained on pMA samples (2.1± 0.38). These samples differed significantly also for the N^{W2}/N^{I} percentage which resulted 18.0 % (± 2.3) in pMA and 12.5% (± 3.2) in AIR samples. By comparing these values (a^* and N^{W2}/N^{I} percentage) with that one reported in Tables 1 and 2, it can be noted that pMA samples at the end of the storage reach a middle quality (from QL =3 to QL =2), while AIR samples reach the poor quality class (QL =1). A good match with the scores attributed by sensory evaluation was showed: at the end of storage samples stored in pMA or AIR were scored 3.0 (± 0.3) and 1.0 (± 0.18) respectively. In fresh-cut radicchio stored in pMA bags (containing at the equilibrium 10% O_2 and 7% CO_2), the browning of the white part was delayed. This accords with the general acknowledge that the use of low levels of O_2 and high concentrations of CO_2, in combination with low storage temperatures, represent the optimal conditions for storing fresh-cut vegetables, maintaining sensorial quality (Jacxsens et al. 2000; Hempel, et al., 2013).

4.0. Conclusion

The paper addresses the development of CVS for the automatic color analysis of fruit and vegetables: the goal of the system is to evaluate non-destructively the quality level of products. In particular the proposed procedures involves new methodologies that reduce the human effort and subjectivity required to design a CVS and to optimize its performance. The parameters controlling the image processing algorithms are automatically set using data driven approaches instead of being fixed by expert operators. A new technique is described that automatically locates the color-chart in the images: a critical step to evaluate and reduce the error in color measuring due to unavoidable instabilities of the CVS. A new approach is proposed to separate foreground and background: a region growing method based on adaptive thresholds derived from the images is used to select most of the product surface avoiding the inclusion of misclassified background pixels. Finally, relevant colors best suited to effectively characterize a specific product are autonomously identified using a hierarchical clustering approach that proved to be successful also on products characterized by nonuniform colors. Hierarchical clustering is completely data-driven and do not suffer from initialization problems or from the setting of the desired number of colors. It produces a structure (dendrogram) that can be exploited by the CVS to check several color quantizations (associated with different number and kind of colors). The system tests each possibility using its performance in distinguishing the quality levels during the following classification and estimation steps. Each color quantization is derived from the dendrogram by simply dynamically moving up or down on the hierarchical structure. The resulting CVS proved to be effective in separating the quality classes of interest on the fresh-cut radicchio.

The proposed approach allows the CVS to be easily applied to different products, with nonuniform surface color (variable number and kind of colors. This significantly simplifies its practical application in food processing with a reduced and less critical human intervention for its configuration.

Acknowledgments

This research was financed by MIUR (Research Projects: ‘Smart Cities and Communities and Social Innovation Prot. 84/Ric 02/03/2012, PON4a2_Be&- Save). The authors thank Arturo Argentieri for the technical support to the configuration of the experimental set-up.

References

Avila, F., Mora, M., Oyarce, M., Zuñiga, A., & Fredes, C. (2015). A method to construct fruit maturity color scales based on support machines for regression: Application to olives and grape seeds. *Journal of Food Engineering*, 162, 9-17.

Balaban, M. O. (2008). Quantifying nonhomogeneous colors in agricultural materials part I: method development. *Journal of Food Science*, 73(9), 431-437.

Cefola, M., & Pace, B. (2015). Application of oxalic acid to preserve the overall quality of rocket and baby spinach leaves during storage. *Journal of Food Processing and Preservation*. doi:10.1111/jfpp.12502.

Hempel, A., O'Sullivan, M. G., Papkovsky, D. B., & Kerry, J. P. (2013). Nondestructive and continuous monitoring of oxygen levels in modified atmosphere packaged ready-to-eat mixed salad products using optical oxygen sensors, and its effects on sensory and microbiological counts during storage. *Journal of Food Science*, 78(7), 1057-1062.

Hernández-Carrión, M., Hernando, I., Sotelo-Díaz, I., Quintanilla-Carvajal, M. X., & Quiles, A. (2015). Use of image analysis to evaluate the effect of high hydrostatic pressure and pasteurization as preservation treatments on the microstructure of red sweet pepper. *Innovative Food Science & Emerging Technologies*, 27, 69-78.

Jacxsens, L., Devlieghere, F., De Rudder, T., & Debevere, J. (2000). Designing equilibrium modified atmosphere packages for fresh-cut vegetables subjected to changes in temperature. *LWT – Food Science and Technology*, 33, 178–87.

Kader, A. A. (2002). Biology and technology: An overview. Postharvest Technol. Hortic. Crops, 3311, 39-48.

Kang, S. P., East, A. R., & Trujillo, F. J. (2008). Colour vision system evaluation of bicolour fruit: A case study with ‘B74’mango. *Postharvest Biology and Technology*, 49(1), 77-85.

Kordecki, A., & Palus, H. (2014). Automatic detection of colour charts in images. *Przegląd Elektrotechniczny*, 90(9), 197-202.

Manninen, H., Paakki, M., Hopia, A., & Franzén, R. (2015). Measuring the green color of vegetables from digital images using image analysis. *LWT-Food Science and Technology*. doi:10.1016/j.lwt.2015.04.005

Nunes, M. C. N., Emond, J. P., Rauth, M., Dea, S., & Chau, K. V. (2009). Environmental conditions encountered during typical retail display affect fruit and vegetable quality and amount of waste. *Postharvest Biology and Technology*, 51, 232–241.

Nunes, M. C. N. (2015). Correlations between subjective quality and physicochemical attributes of fresh fruits and vegetables. *Postharvest Biology and Technology*, 107, 43-54.

Otsu, N. (1975). A threshold selection method from gray-level histograms. *Automatica*, 11, 23-27.

Pace, B., Cefola, M., Renna, F., & Attolico, G. (2011). Relationship between visual appearance and browning as evaluated by image analysis and chemical traits in fresh-cut nectarines. *Postharvest Biology and Technology*, 61(2), 178-183.

Pace, B., Cefola, M., Renna, F., Renna, M., Serio, F., & Attolico, G. (2013). Multiple regression models and Computer Vision Systems to predict antioxidant activity and total phenols in pigmented carrots. *Journal of Food Engineering*, 117(1), 74-81.

Pace, B., Cefola, M., Da Pelo, P., Renna, F., Attolico, G. (2014a). Non-destructive evaluation of quality and ammonia content in whole and fresh-cut lettuce by computer vision system. *Food Research International*, 64, 647-655.

Pace B., Cardinali A., D'Antuono I., Serio F., Cefola M. (2014b). Relationship between quality parameters and the overall appearance in lettuce during storage. *International Journal of Food Processing and Technology*, 1, 18-26.

StatSoft, I. N. C. (2001). STATISTICA (data analysis software system), version 6. Tulsa, USA, 150.

Wojciech, J., Florkowski, R. L., Shewfelt, B. B., & Stanley, E. P. (2014). *Postharvest handling. A systems approach* (2nd ed., pp. 153–204). Amsterdam: Elsevier.

Wu, D., & Sun, D. W. (2013a). Colour measurements by computer vision for food quality control–A review. *Trends in Food Science & Technology*, 29(1), 5-20.

Wu, D., & Sun, D. W. (2013b). Advanced applications of hyperspectral imageing technology for food quality and safety analysis and assessment: a review – Part I: Fundamentals. *Innovative Food Science and Emerging Technologies*, 19, 1-14.

Zhang, B., Huang, W., Li, J., Zhao, C., Fan, S., Wu, J., & Liu, C. (2014). Principles, developments and applications of computer vision for external quality inspection of fruits and vegetables: A review. *Food Research International*, 62, 326-343.

Zhang, B., Huang, W., Gong, L., Li, J., Zhao, C., Liu, C., & Huang, D. (2015). Computer vision detection of defective apples using automatic lightness correction and weighted RVM classifier. *Journal of Food Engineering*, 146, 143-151.

Figure Captions

Figure 1. The X-Rite color-chart used in the experiments. It provides a set of grey levels on the bottom row. Several colors, properly distributed in the color space, allow the evaluation of color fidelity of the acquisition system: the expected $L^*a^*b^*$ values are provided by the manufacturer.

Figure 2. A typical acquired image (A) before the processing to automatically detect the color-chart. The two lightest greys are detected (B) and provide the required information about position and attitude of the color-chart. Then all the other patches are identified and isolated (C). Now the color-chart is removed (D) to simplify the following processing.

Figure 3. The purple lines represent a typical histogram of color distances measured over the whole image between a pixel and its 8-neighborings. The significant peak (around zero) relates to distances measured over mostly homogeneous regions: they are approximated by a normal distribution ($fit_{1,}$ red line). A second peak ($fit_{2,}$ brown line) is related to the distances occurring at the edges between foreground and background: this distribution allows the set of the threshold used by the growing process to the value μ_2 - $\sigma_2/2$. This choice does not include some pixels belonging to the product's surface but prevents any background pixel from being erroneously classified as foreground.

Figure 4. The picture (A) shows typical starting regions used to separate foreground from background: a very conservative threshold is used that does not need to be adapted to different products or acquisition conditions. In fact, (B) shows that this conservative approach avoids any misclassification of background pixels but selects a very limited part of the product's surface. In a second phase of the processing, these starting regions grow to include neighboring pixels whose distance is lower than the threshold automatically set using the histogram of distances. This produces

the final mask (C) that, as shown in (D), selects most of the product's surface: this pixels are used for further processing.

Figure 5. The two-dimensional histogram of color (related to the a^* and b^* channels of the $L^*a^*b^*$ color space). The L^* is not considered because proved to be less stable and significant to characterize colors. These representations, derived from each image, enable the identification of the most relevant colors. The hierarchical clustering applied on these colors selects their quantization best suited to separate the quality levels.

Figure 6. The surface of the fresh-cut radicchio shown in image (A) is separated into its white and red regions (B) using the first level of the dendrogram structure produced by the hierarchical clustering. The approach is completely data driven and does not require any setting based on heuristics or on a-priori knowledge. The white region of radicchio is further segmented using the next level of the dendrogram (C). The approach separates the light white and the dark white regions without any further processing: all the possible color quantization are coded into the hierarchical structure and can be dynamically explored to identify the most significant and best performing colors.

Figure 1

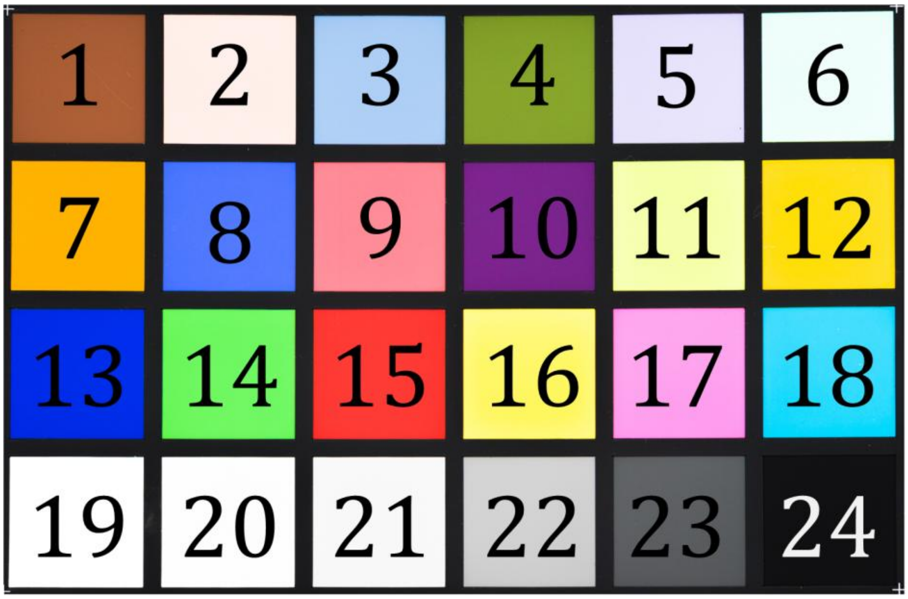

Figure 1.

Figure 2

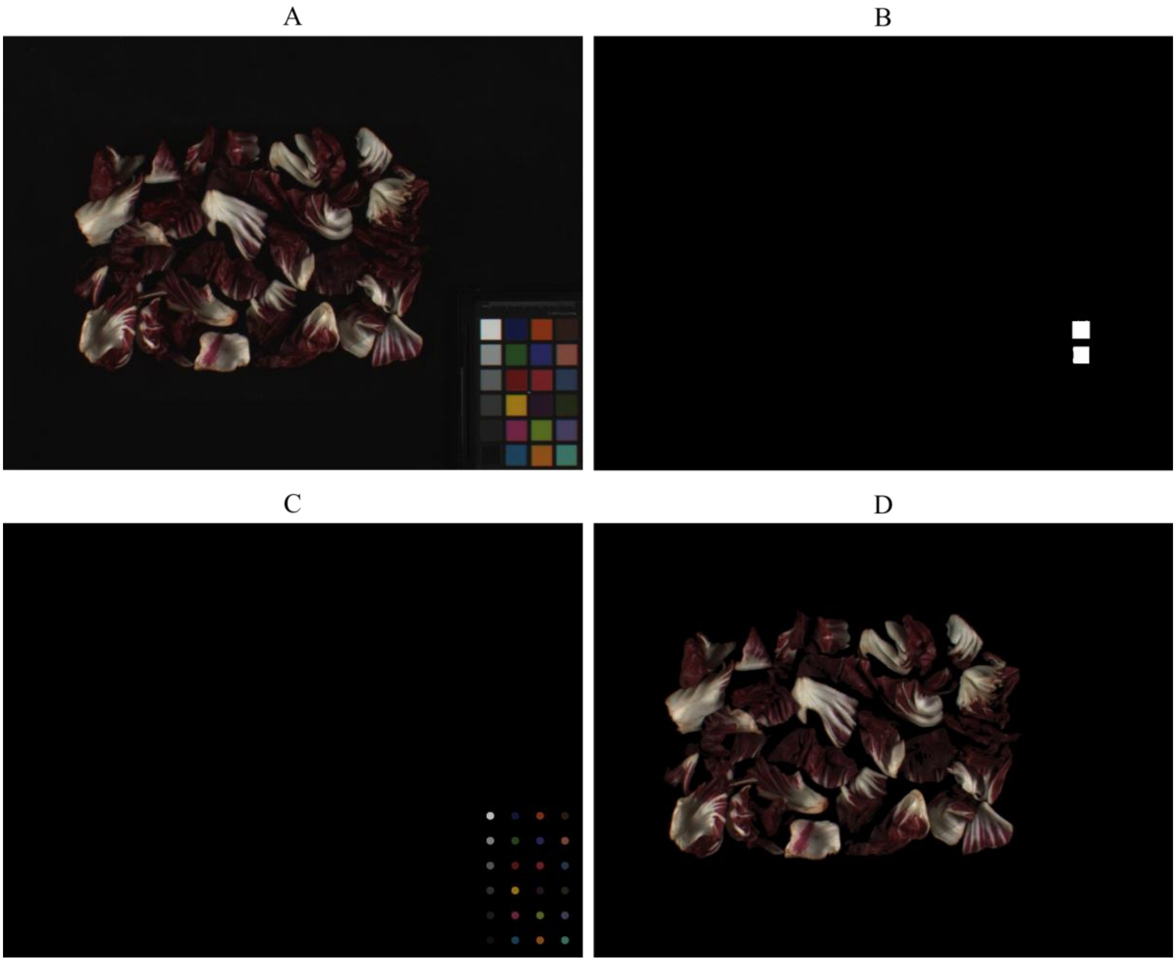

Figure 2.

Figure 3

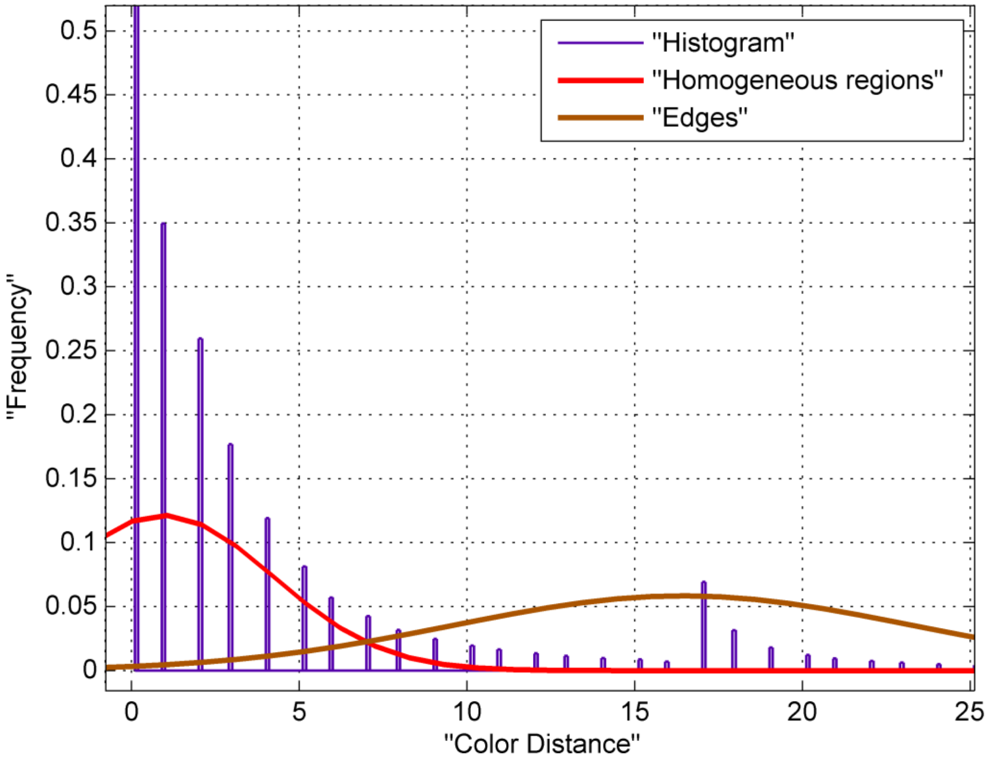

Figure 3.

Figure 4

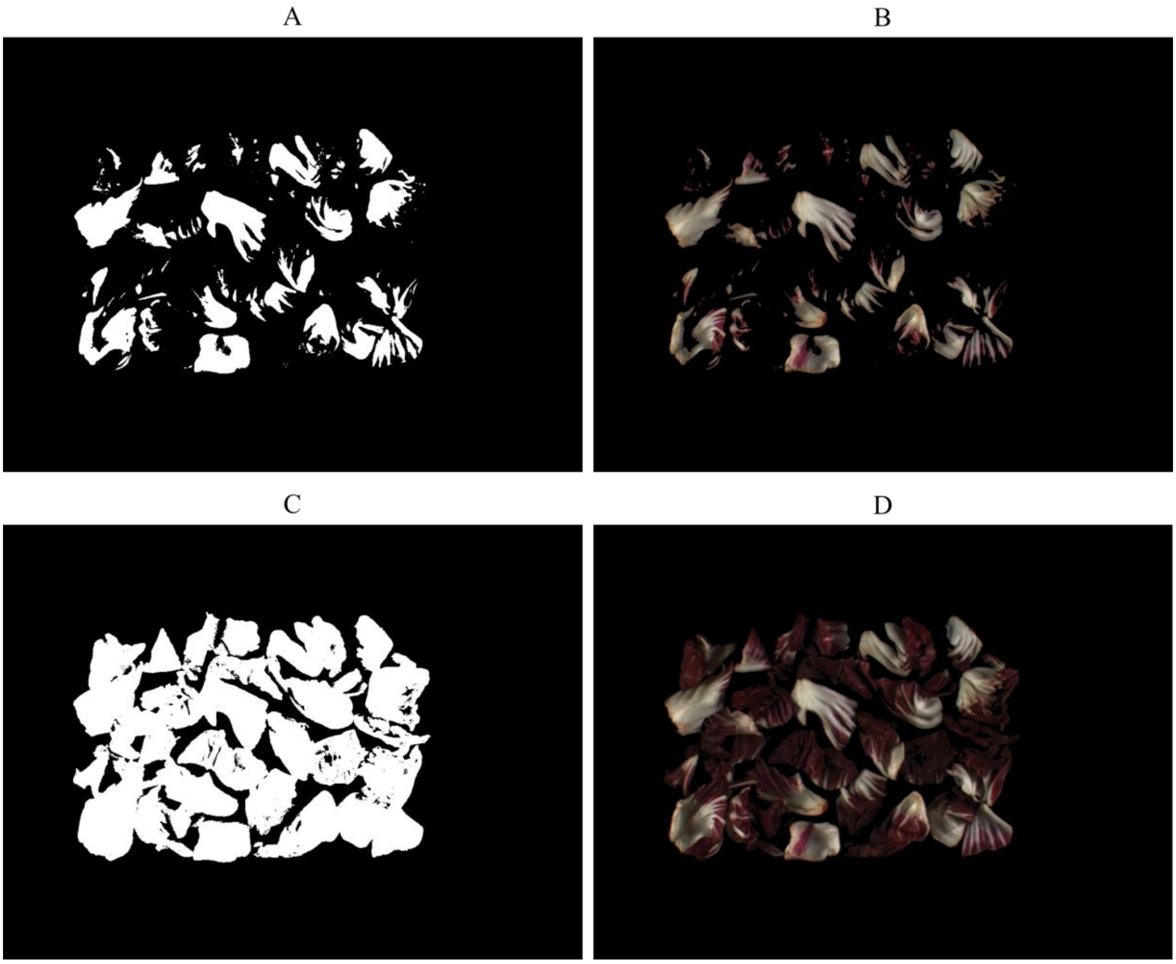

Figure 4.

Figure 5

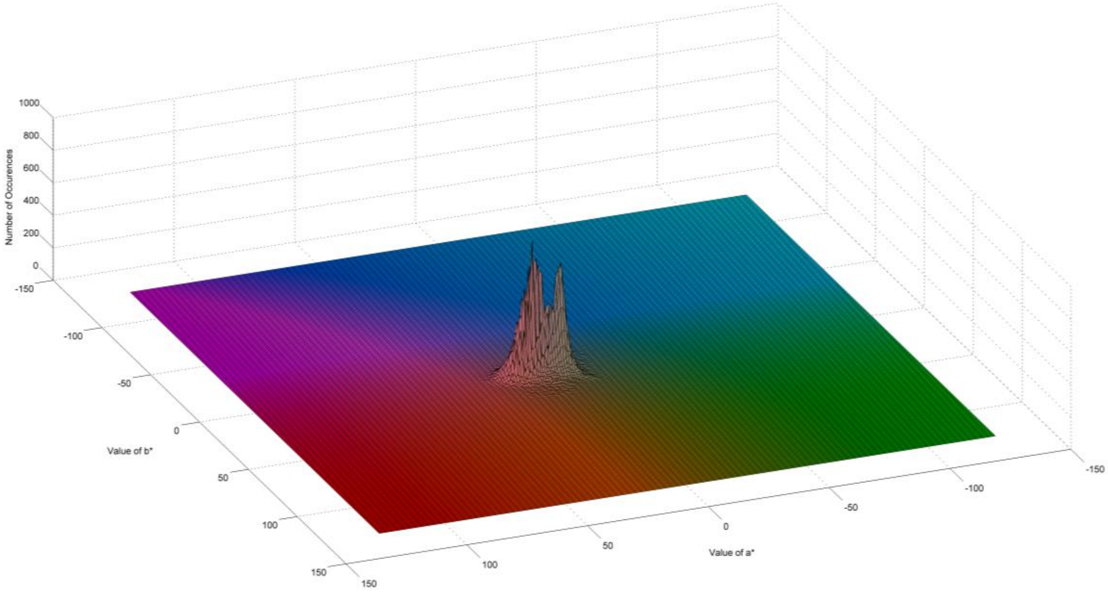

Figure 5.

Figure 6

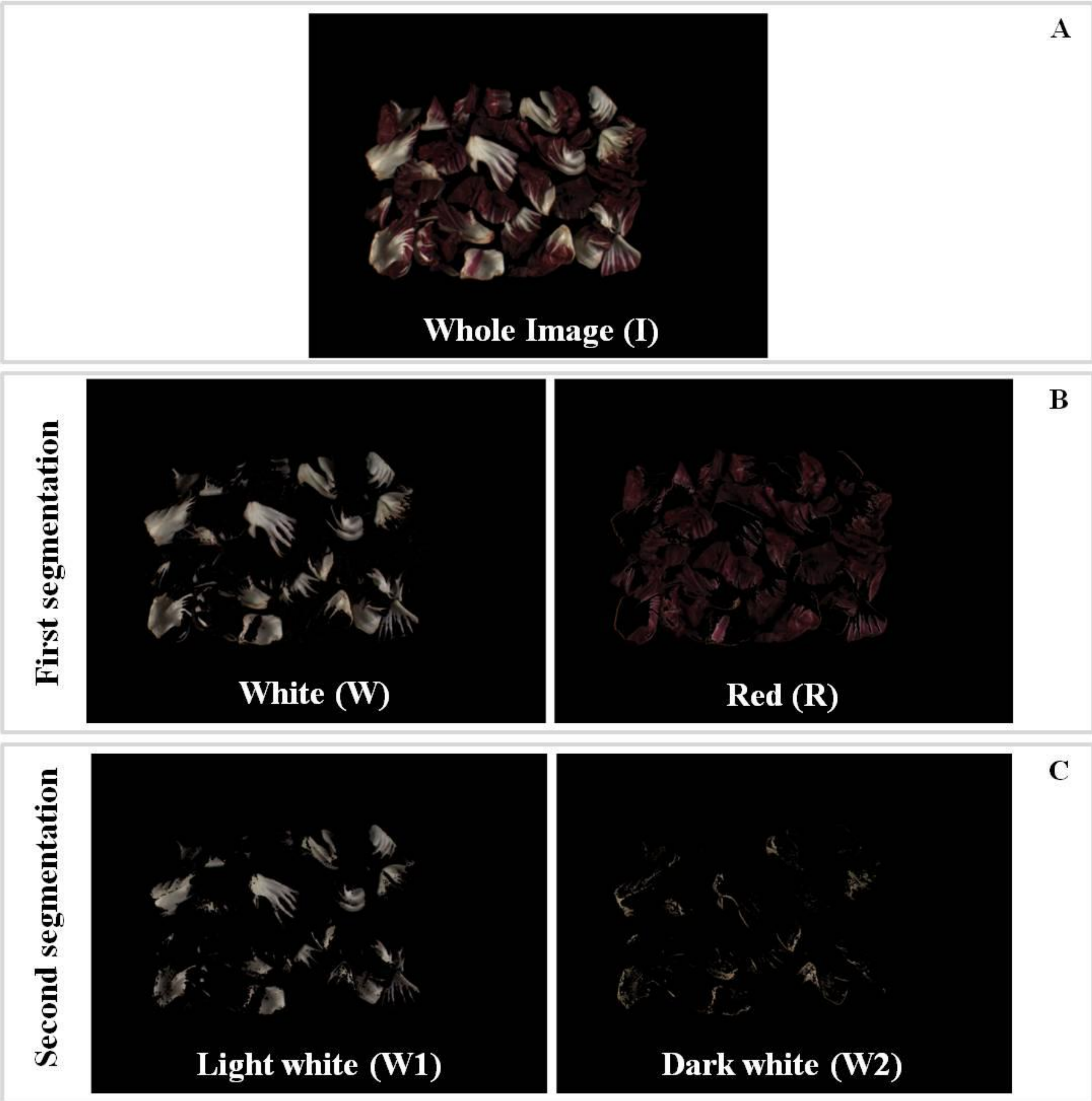

Figure 6.

Table 1

Table 1. Main effects of quality levels (very good, good, acceptable, poor and very poor) on the ammonium (NH_4^+ µmole/g) and colour features (*L**, *a** and *b**) automatically extracted from the entire and segmented images by computer vision system in fresh-cut radicchio.

Fresh-Cut Radicchio Images		Very good 5		Good 4		Acceptable 3		Poor 2		Very poor 1		*P* value
	NH_4^+	0.07	c	0.10	c	0.22	b	0.30	b	0.66	a	***
Entire Image	*L**	41.63	a	40.59	a	37.10	b	37.50	b	35.75	b	***
	*a**	11.12	a	10.56	b	10.62	b	10.67	b	12.15	b	***
	*b**	4.41	c	4.97	b	5.58	a	5.54	a	5.10	b	***
I segmentation												
White (W)	*L**	64.07	a	61.96	a	53.56	b	59.15	b	58.65	b	***
	*a**	1.56	c	1.66	c	2.21	b	2.13	b	2.69	a	***
	*b**	4.90	c	5.59	b	6.47	a	6.37	a	6.08	ab	***
Red (R)	*L**	32.20	a	32.07	ab	30.89	c	31.03	bc	31.08	c	**
	*a**	16.98	a	16.16	b	14.85	c	15.17	c	15.91	b	***
	*b**	4.64	c	5.07	bc	5.54	a	5.48	a	5.02	b	***

Mean values of 6 replicates. For each parameter the mean values followed by different letters, are significantly different (*P*-value < 0.05) according to Student-Newman-Keuls (SNK) test.
Significance: **and *** = significant at *P*-value ≤ 0.01 and 0.001, respectively.

Table 2

Table 2. Main effects of quality levels (very good, good, acceptable, poor and very poor) on the percentages (%) of the W1 (dark component of the white part) and W2 (clear component of the white part) components calculated as number of pixel of each component on the white part W (N^{W1}/N^{W} or N^{W2}/N^{W}) or on the entire surface image (I) (N^{WI}/N^{I} or N^{W2}/N^{I}) of fresh-cut radicchio images.

White components percentages (%)	Very good		Good		Acceptable		Poor		Very poor		*P* value
	5		4		3		2		1		
N^{W1}/N^{W}	14.02	b	21.8	a	27.63	a	27.2	a	25.36	a	***
N^{W1}/N^{I}	4.44	b	7.24	a	6.99	a	7.34	a	4.50	b	***
N^{W2}/N^{W}	85.98	a	78.2	b	72.37	b	72.8	b	74.63	b	***
N^{W2}/N^{I}	28.88	a	26.3	a	19.09	b	19.6	b	13.33	c	***

Mean values of 6 replicates. For each parameter the mean values followed by different letters, are significantly different (*P*-value < 0.05) according to Student-Newman-Keuls (SNK) test.
Significance: *** = significant at *P*-value ≤ 0.001.