

Article

# Assessment of Tomato Maturity in Different Layers by Spatially Resolved Spectroscopy

Yuping Huang <sup>1</sup>, Wan Si <sup>1</sup>, Kunjie Chen <sup>2</sup> and Ye Sun <sup>2,\*</sup>

<sup>1</sup> College of Mechanical and Electronic Engineering, Nanjing Forestry University, Nanjing 210037, China; huangyuping@njfu.edu.cn (Y.H.); siwan4221@gmail.com (W.S.)

<sup>2</sup> College of Engineering, Nanjing Agricultural University, Nanjing 210031, China; kunjiechen@njau.edu.cn

\* Correspondence: sunye@njau.edu.cn

Received: 4 November 2020; Accepted: 15 December 2020; Published: 17 December 2020



**Abstract:** Tomato maturity is important to determine the fruit shelf life and eating quality. The objective of this research was to evaluate tomato maturity in different layers by using a newly developed spatially resolved spectroscopic system over the spectral region of 550–1650 nm. Thirty spatially resolved spectra were obtained for 600 tomatoes, 100 for each of the six maturity stages (i.e., green, breaker, turning, pink, light red, and red). Support vector machine discriminant analysis (SVM) models were first developed for each of individual spatially resolved (SR) spectra to compare the classification results of two sides. The mean spectra of two sides with the same source-detector distances were employed to determine the model performance of different layers. SR combination by averaging all the SR spectra was also subject to comparison with the classification model performance. The results showed large source-detector distances would be helpful for evaluating tomato maturity, and the mean\_SR 15 obtained excellent classification results with the total classification accuracy of 98.3%. Moreover, the classification results were distinct for two sides of the probe, which demonstrated even if in the same source-detector distances, the classification results were influenced by the measurement location due to the heterogeneity for tomato. The mean of all SR spectra could only improve the classification results based on the first three mean\_SR spectra, but could not obtain the accuracy as good as the following mean\_SR spectra. This study demonstrated that spatially resolved spectroscopy has potential for assessing tomato maturity in different layers.

**Keywords:** tomato maturity; different layers; spatially resolved spectra; SVM

## 1. Introduction

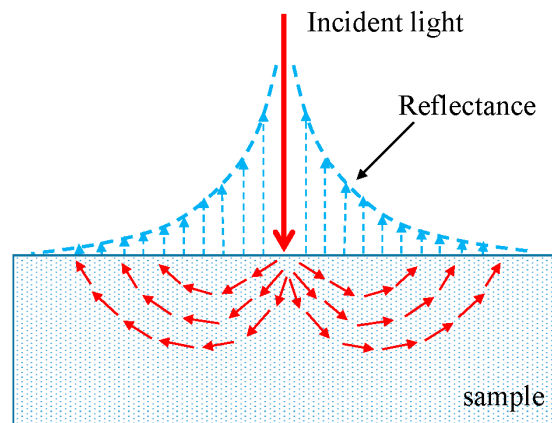
Tomato is the second most consumed vegetable product in the world due to its abundant nutrition, such as sugar, organic acids, and vitamin C [1–3]. Maturity is one of the most important indicators to determine the fruit picking time and shelf life. Color is one of the remarkable characteristics closely related to tomato maturity. As tomato matures, the tissue of tomato turns from green to red due to the changes of chlorophyll, anthocyanin and other pigments, along with the changes of the internal texture and chemical composition [4]. Therefore, color should be a significant parameter for assessing tomato maturity.

In recent years, many nondestructive detection techniques, such as magnetic resonance imaging, near-infrared spectroscopy, hyperspectral imaging, machine vision, have been used to assess the maturity or color grade of tomatoes [5–7], of which machine vision is the most widely used. Arakeri [8] and Arefi [9] effectively distinguished mature tomatoes from immature tomatoes through machine vision technology, and the classification rates were 96.5% and 96.4%, respectively. Wan et al. [10] used machine vision combined with a back-propagation neural network method to distinguish the three maturity grades (green, orange, and red) of tomato, with an accuracy rate of 99.3%.

Although machine vision technology can accurately distinguish the color grade of the fruit surface [11], it is difficult to detect the changes of internal information caused by ripeness. While the visible and near-infrared spectroscopy is capable to obtain information about the internal composition and structure of samples [12–14]. Some scholars applied this technology for the determination of tomato maturity [15,16]. Alenazi et al. [17] reported an assessment of fresh tomato fruit at five distinctive maturity stages (breaker, turning, pink, light-red, and red) using the handheld near infrared (NIR) enhanced spectrometer at a wavelength range of 285–1200 nm, and the results showed that the NIR spectroscopy is an effective tool for predicting tomato fruit quality during the harvest stage. Sirisombon et al. [18] used near infrared spectroscopy to identify the three maturity levels (mature green, pink and red) of tomatoes in the wavelength range of 1100–2300 nm, with an optimal classification accuracy rate of 96.9%. Clément et al. [19] established a mathematical model for the color parameters of tomatoes using visible and near infrared spectroscopy (400–1500 nm) with the correlation of 95%. However, conventional visible and near-infrared spectroscopy is a single-point or specific area measurement, which could not provide spatial information of the research object. Thus, for such objectives with heterogeneous structure like tomatoes, conventional visible and near-infrared spectroscopy could not accurately determine their quality.

Spatially-resolved spectroscopy is an emerging spectroscopic technique for measuring reflectance from the sample at different spatial distances between a point light source and detectors. Spatially-resolved (SR) spectra contain information about the sample at different layers (or depths) since the transport path of photons generally forms a “banana-shape” in biological tissues [20], shown in Figure 1. Short spatial distances between the illumination and detectors would provide top layer information of the sample, while large source-detector distances is useful to obtain deep layer information of the sample [21]. Numerous studies reported on assessment of quality attributes using spatially resolves spectroscopy with different orientations [4,22,23]. Meanwhile, the limited spatial distances and wavelength range could not provide adequate information about samples. Huang et al. [21] first developed a new spatially resolved spectroscopic system for property and quality assessment of horticultural and food products with source-detector distance of 1.5–36 mm. Then, spatially resolved spectroscopic system was applied to predict the soluble solid content (SSC), pH, and firmness of tomato over the spectral range of 550–1650 nm, these results demonstrated that the spatially resolved spectroscopy has advantages over conventional spectroscopy for enhancing quality assessment of tomatoes [24,25]. However, most of the previous studies were focused on a single quality parameter, whereas maturity is a comprehensive quality that contains a variety of quality information, and there are still many challenges that must be addressed for its application to assess tomatoes maturity. Moreover, tomato ripening is a non-uniform process, on the one hand, the color distributed variously with space during the ripeness, on the other hand, chemical constituent and texture also changed with interspace as tomato ripeness, thus single surface color indexes may not provide adequate information for tomato maturity evaluation. Since the acquisition probe of the spatially resolved spectroscopic system includes multiple fibers, it can obtain internal information at different depths about the sample within a certain detection distance. Thus, the spatially resolved spectroscopic system should be more suitable for evaluating the tomato maturity.

Therefore, in this study, the feasibility of spatially resolved spectroscopic system for assessing tomato maturity in different layers was analyzed and discussed. The specific objectives of this study were to: (1) acquire SR spectra of each tomato sample in different layers over the spectral range of 550–1650 nm; (2) reveal the relationship between quality attributes and maturity stages; (3) compare the classification results for tomato maturity in different layers; (4) analyze the classification results with various locations in the same layer.



**Figure 1.** Spatially resolved spectra at different source-detector distances.

## 2. Material and Methods

### 2.1. Samples

A total of 600 ‘Sun Bright’ tomato samples at six maturity stages were picked from a research field at Michigan State University’s Horticultural Research and Teaching Center (Holt, MI, USA). The samples were cleaned and inspected to make sure there were no external defects for tomato samples. The fruit were then classified into six maturity stages (i.e., Green, Breaker, Turning, Pink, Light-Red and Red) based on the USDA standards [26]. Each stage had 100 tomato samples. Before the experiment started, tomato samples were kept at room temperature (about 25 °C) for at least 12 h to make sure each of the whole samples had consistent temperature to avoid the temperature influence for spectra or analysis results. Spectral acquisitions were first taken for each of sample using the spatially resolved spectroscopic system over the spectral range of 550–1650 nm. Thereafter, the samples were subjected to physical and chemical experiment.

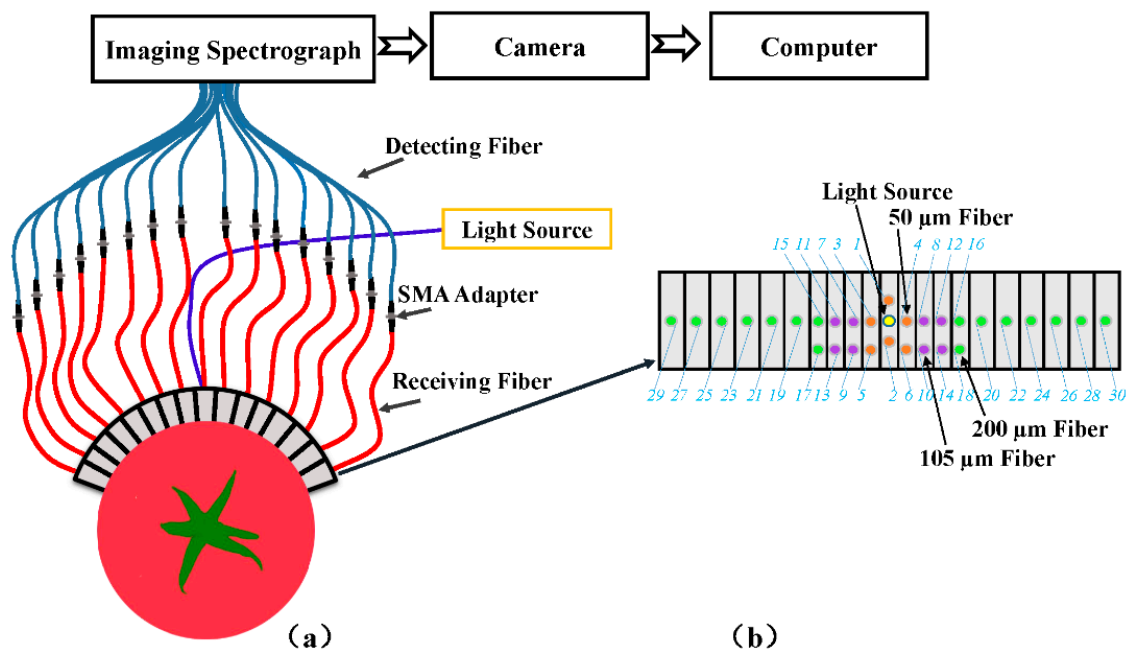
### 2.2. Spatially Resolved (SR) Spectra Measurement

Thirty spatially resolved spectra for each of tomato samples were acquired using a newly developed spatially resolved spectroscopic system (Model 1003B-10152, Headwall Photonics, Inc., Fitchburg, MA, USA), shown as Figure 2, with the advantages of large source-detector distances of 1.5–36 mm, flexibility for samples of flat or curved surface, and enhanced dynamic range by using different sized fibers. A detailed description of the acquisition probe was given by Huang [21]. During the spectral collection, the spatially resolved spectroscopic system was set with an integration time of 60 ms and the light source output power was set for 240 W. The acquisition probe was directly contact with the equatorial area of tomato fruit. The symmetric pair spectra with the same source-detector distance were performed to average in order to assess the influence of source-detector distances for tomato maturity detection, resulting in 15 mean\_SR spectra with different source-detector distances. A white cylindrical Teflon block was used as a white reference, and the same experimental procedure was taken for obtaining white reference spectra to calibrate the sample spectra. Whereas the dark reference spectra were acquired by closing the light source in a dark room. The relative spatially resolved spectra for the samples were calculated using the white, the dark and the original sample spectra [27]. The relative reflectance spectra were used in further data analysis.

### 2.3. Quality Analysis of Tomato Fruit

After spectral acquisition, nondestructive and destructive firmness detection, soluble solid content (SSC) test, and pH measurement were performed on the tomato samples. Acoustic firmness of each samples was obtained by a desktop acoustic firmness sensor (AFS, AWETA, Nootdorp, The Netherlands), while the other two firmness parameters, the compression slope and the puncture

maximum force, were subjected by a Texture Analyzer (Model TA.XT2i, Stable Micro System, Inc., Surrey, UK). The SSC value for each of the tomatoes was acquired from a handheld digital refractometer (model PR-101, Atago Co., Tokyo, Japan), and pH value for each of tomato samples was carried out using a portable pH meter (model HI98161, Hanna Instruments, Inc., Woonsocket, RI, USA). The analysis of variance (ANOVA) was performed on the quality parameters (AF, CS, PA, SSC, and PH) of six maturities of tomatoes using SPSS 18.0 statistics software. The level of  $p < 0.05$  was considered significant in all analyses.



**Figure 2.** Schematic of the spatially resolved spectroscopic system (a), and the acquisition probe (b).

#### 2.4. Tomato Maturity Classification Models

Since the maturity was determined by many indicators, such as chemical composition, texture and color, the obtained spectra may not present linear relation with tomato maturity. Support vector machine discriminant analysis (SVM), as commonly used nonlinear model, uses LIBSVM package to calculate the probabilities of each sample belonging to each possible class. This method first calculates one-against-one pairwise class probabilities as a function of decision value, using the training data. The class probabilities are then found as an optimization problem using these  $k \times (k - 1)/2$  pairwise probabilities (assuming there are  $k$  classes). In this study, SVM models were first developed for each of the 30 individual SR spectra to compare the classification results for two sides of acquisition probe so that it could provide some foundation for evaluating symmetry of tomato maturity. Moreover, in order to assess the tomato maturity on different layers, 15 mean\_SR spectra with different source-detector distances and SR combination by averaging all SR spectra were also used to establish the SVM models.

The 600 tomato samples were randomly divided into training and test sets. There were 360 and 240 samples in training and test sets, respectively. SVM models were developed using MATLAB R2017a (The MathWorks, Inc., Natick, MA, USA) combined with PLS Toolbox 8.2 (Eigenvector Research, Inc., Wenatchee, WA, USA). Autoscale is an exceptionally common preprocessing method which uses mean-centering followed by division of each column (variable) by the standard deviation of that column. It is a valid approach to correct different variable scaling and units if the predominant source of variance in each variable is signal rather than noise. Therefore, autoscale preprocessing was performed in each of individual SR spectra. Venetian blinds cross-validation is suitable for the data that are already in random order [28] and was used in the training set to calibrate the models.

The performance of the models was evaluated by classification accuracy (recognition rate), sensitivity, and specificity for the test set of samples.

### 3. Results and Discussion

#### 3.1. Differences of Quality Attributes for Tomatoes at Six Maturity Stages

Table 1 shows the means and standard deviations of the quality parameters, and ANOVA for the quality parameters of tomatoes in six maturity stages using SPSS 18.0 statistics software. The SSC and pH showed an upward trend, which meant that the sugar content accumulated and the acidity content decreased as tomato matures from green to red. While the firmness parameters like acoustic firmness (AF), compression slope (CS), and puncture maximum force (PF) showed a downward trend, which indicated that as tomato fruit turns green to red, the firmness became soft, which is also agreement with previous studies [4]. All of the quality parameters showed significant differences between different maturity stages excluding acoustic firmness, which could be partly because the tomato in breaker, turning and pink stages, the color began to change, the acoustic test measured the global properties of tomato sample and the resonant frequency in these stages may have limited changes, resulting in no significant differences of these three maturity stages. Considering the results, tomato maturity was influenced by not only chemical composition, such as SSC and pH, but also structural characteristics like firmness.

**Table 1.** Means and standard deviations of the quality parameters for tomatoes with different maturities \*.

Maturity	AF	CS	PF	SSC	pH
Green	9.03 ± 1.88 <sup>a</sup>	34.57 ± 7.30 <sup>a</sup>	17.52 ± 2.13 <sup>a</sup>	4.42 ± 0.34 <sup>a</sup>	3.98 ± 0.22 <sup>a</sup>
Breaker	5.75 ± 1.33 <sup>b</sup>	18.95 ± 3.86 <sup>b</sup>	15.08 ± 2.16 <sup>b</sup>	4.61 ± 0.43 <sup>b</sup>	4.02 ± 0.10 <sup>b</sup>
Turning	5.54 ± 1.37 <sup>b</sup>	16.04 ± 3.31 <sup>c</sup>	13.89 ± 2.16 <sup>c</sup>	4.80 ± 0.47 <sup>c</sup>	4.10 ± 0.10 <sup>c</sup>
Pink	5.38 ± 1.48 <sup>b</sup>	13.91 ± 2.78 <sup>d</sup>	11.69 ± 1.73 <sup>d</sup>	5.12 ± 0.44 <sup>d</sup>	4.14 ± 0.09 <sup>d</sup>
Light Red	4.66 ± 1.36 <sup>c</sup>	10.54 ± 1.67 <sup>e</sup>	8.72 ± 1.46 <sup>e</sup>	5.37 ± 0.50 <sup>e</sup>	4.24 ± 0.11 <sup>e</sup>
Red	4.04 ± 1.36 <sup>d</sup>	8.20 ± 2.17 <sup>f</sup>	7.12 ± 1.54 <sup>f</sup>	5.73 ± 0.56 <sup>f</sup>	4.34 ± 0.12 <sup>f</sup>

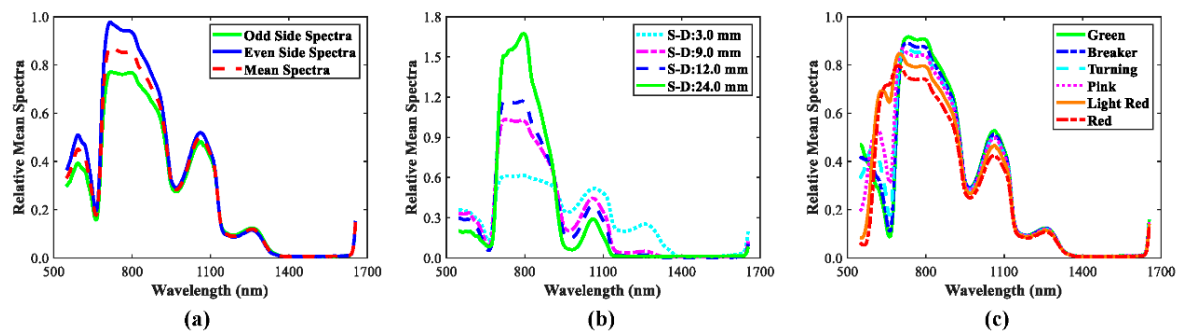
\* Numbers for the same columns with different letters (a, b, c, d, e, f) are different at the level of 0.05 based on the analysis of variance.

#### 3.2. Spectral Features

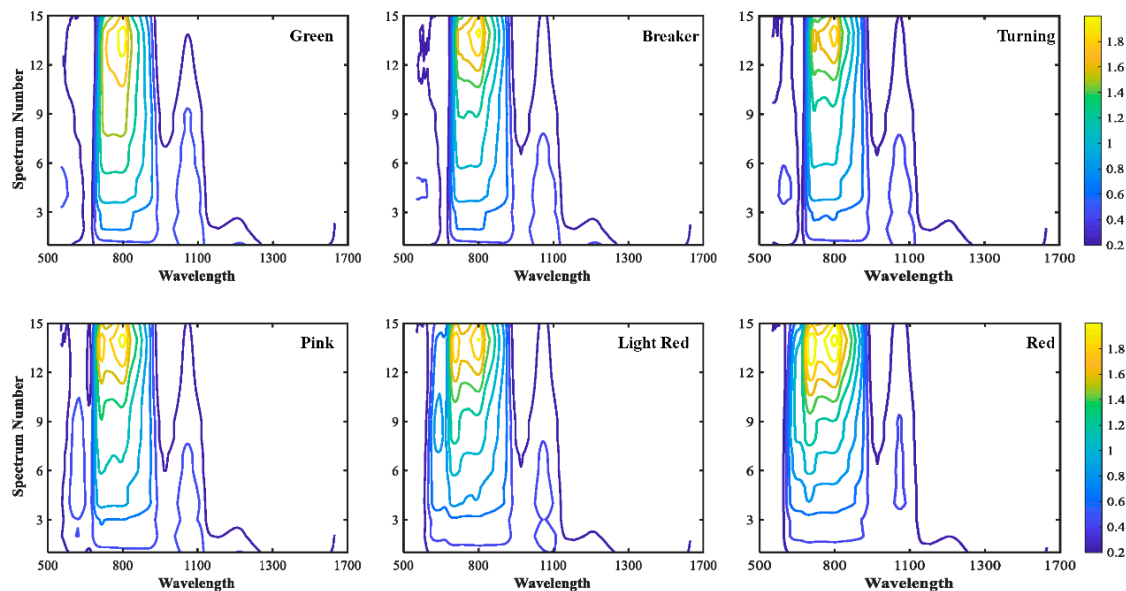
Figure 3a shows the SR spectra at source-detector distance of 6.0 mm in two sides of the probe and the mean spectra of two sides. The overall pattern for SR spectra in two sides was similar excluding in the spectral range of 700–900 nm, which suggested there should be some discrepancies for the classification results of these two SR spectra, this speculation was demonstrated in next section. Four mean\_SR spectra at different source-detector distances (3 mm, 9 mm, 12 mm, and 24 mm) were illustrated in Figure 3b. Large differences in the pattern of these spectra were observed over the whole spectral range, because these spectra contained information at different tissue layers (or depths) of the sample. It suggested each SR spectra would have different classification capabilities for tomato maturity, and therefore it is important that SVMDA models should be developed for each mean\_SR spectra so as to determine the best source-detector distance for tomato maturity detection. In the near infrared region, the spectra were influenced by water absorption bonds and C-H bonds at 970 nm, 1180 nm, and the spectral beyond 1400 nm. Figure 3c displays noticeable differences in relative reflectance around 560 nm and 680 nm for six maturity stages, which could be due to the fact that as tomato turned green to red, anthocyanin (at 560 nm) increased along with chlorophylls (at 680 nm) decreased [24,29,30]. Figure 4 shows the contour maps for 15 mean\_SR spectra for tomatoes at six maturity stages. The prominent differences were inspected from contour maps for six tomato maturity stages, particularly in the spectral range of 550–1100 nm. In addition, extremely weak intensity appeared in the spectral range around or beyond 1300 nm due to the high absorption bonds of



O-H. Overall, combined with mean\_SR spectra at different source-detector distances could provide a fingerprint for each of tomato maturity.



**Figure 3.** Mean reflectance spectra of tomatoes at turning stage for odd and even sides with source-detector distance of 6.0 mm and the mean spectra of these two spectra (a), the mean spectra with different source-detector distances for tomatoes at breaker stage (b), and the mean spectra at six maturity stages with the source-detector distance of 6.0 mm (c).

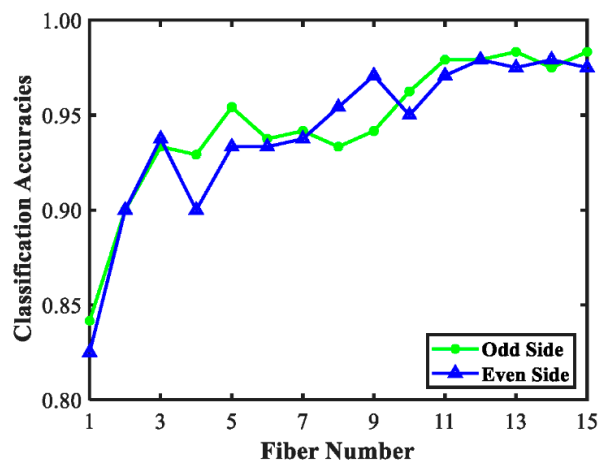


**Figure 4.** The contour maps for 15 SR spectra for tomato at six maturity stages over the spectral region of 550–1650 nm.

### 3.3. Discrimination Models for Tomato Maturity

Figure 5 illustrates the total classification accuracies for identification of tomato maturity using 30 individual SR spectra on two sides of acquisition probe. The classification results generally presented upward trend, which indicated that the SR spectra with a relatively large source-detector distance would have better recognition results for tomato maturity. Moreover, even if the SR spectra with the same source-detector distance on two sides, the obtained results were distinct, especially for the middle SR spectra with source-detector distances of 6.0–12.5 nm. It suggested the structure and chemical composition distributed variously in the different positions of tomato tissue. Previous study reported the source-detector distance was similar to the light penetration depths for the sample [31], which meant when the light penetration depths about 6.0–12.5 nm of the tomato fruit, the single-point detection could not be accurate due to the relative large distinct results. The classification results for each individual SR spectra on two sides of the acquisition probe demonstrated that, even in the same layer for detection, the recognition for tomato maturity still had diversities due to the heterogeneity

of tomato tissue. This suggested the single-point or small-area detection could not be suitable for assessing non-uniform tissue in fruit, such as tomato fruit.



**Figure 5.** The total classification accuracies for recognition tomato maturity based on 15 individual SR spectra on two sides.

Table 2 summarizes the tomato classification results in training and test sets based on SVM/DA models using the mean\_SR spectra of two sides and the mean spectra of all spectra, to evaluate the classification results for tomato maturity in different layers and compare with the results for the combination of all layers. The recognition results in calibration set were consistently better than both cross-validation and test sets. Moreover, as the spectrum number increased, the calibration results become better and stable, specifically for the SR spectra after mean\_SR 7, each of the mean\_SR spectra could reach the best classification accuracies of 100%. The classification in cross-validation had relative consistent results except first two mean\_SR spectra. Good classification results in cross-validation indicated the established models had reliable classification performance. In test set, the overall trends for classification accuracies was ascendant along with the spectrum number increased with differences of 18.6% between the optimal and worst spectra. Moreover, when the spectrum number reached mean\_SR spectra 9, the classification results could consistently achieve above 95% for the following spectra, indicating large source-detector distances were useful for maturity evaluation. Furthermore, as shown in Table 1, tomato maturity stages were related to the chemical compositions and structural characteristics, light absorption is related to the chemical composition, while the scattering depends on the structural characteristics. The obtained SR spectra contained the compounding effect of absorption and scattering, thus the SR spectra with large source-detector distances had a long photon path and strong scattering and absorption in the sample, which would reveal more information about sample, resulting in better classification results for tomato maturity. In addition, there were large differences for the first three mean\_SR spectra, and the results were relative lower, which could be due to the limited information about samples at top layers. Mean\_SR 6 with source-detector distance of 9.0 mm had relative higher results, however, the classification accuracies decreased as the spectrum number increased until mean\_SR 9. This suggested the SR spectra with relative short spatial distances between light source and detector could not give stable classification results. The SR combination by averaging all the SR spectra, named mean\_all\_SR spectra, was also presented in Table 2, its classification result was only better than the first three mean\_SR spectra, indicating that averaging all SR spectra could not be an effective means for improvement of classification results. It also suggested that, with the mean of all data, it is possible to enlarge some irrelevant variables for model building.

**Table 2.** Total classification accuracies for tomato maturity by using SVMDA for 15 mean\_SR spectra and mean spectra of all SR spectra <sup>a</sup>.

Type of Spectra	Training Set		Test Set
	Cal (%)	CV (%)	Pred (%)
Mean_SR 1	92.5	85.8	83.3
Mean_SR 2	94.7	90.6	82.9
Mean_SR 3	98.9	93.9	90.0
Mean_SR 4	100	96.4	91.3
Mean_SR 5	98.9	95.0	92.9
Mean_SR 6	100	95.3	96.7
Mean_SR 7	98.9	95.8	93.8
Mean_SR 8	100	95.0	93.8
Mean_SR 9	100	93.9	95.4
Mean_SR 10	100	98.1	97.9
Mean_SR 11	100	97.2	97.5
Mean_SR 12	100	99.2	97.5
Mean_SR 13	100	97.2	95.4
Mean_SR 14	100	96.1	97.1
Mean_SR 15	100	96.9	98.3
Mean_all_SR Spectra	98.6	93.1	91.3

<sup>a</sup> Cal: calibration; CV: cross-validation; Pred: prediction.

Further analysis is shown in Table 3 to summarize classification results for optimal mean\_SR spectra and mean\_all\_SR spectra of tomatoes at six maturity stages. The mean\_SR 15 presented the best recognition rates for six maturity stages with the value of 100% in the training set, besides, the optimal recognition rates also showed in test set for green, pink, and light red stages. Breaker stage had relative lower classification accuracy due to two samples misclassification, while in turning and red stages, only one sample misclassified in neighboring stage. Mean\_all\_SR spectra had the same recognition results with mean\_SR 15 in green and breaker stages, but the recognition rate only reached 75% in turning stage with ten samples misclassification in neighboring stages, while mean\_SR 15 could present 30% improvement for accuracy with the value of 97.5%. Compared with the mean\_SR 15 spectra, mean\_all\_SR spectra gave lower accuracies for the following stages with differences of 5.3% for pink stage, 12.2% for light red stage and 2.9% for red stage, which demonstrated mean\_all\_SR spectra did not have prominent superiority for assessing tomato maturity.

**Table 3.** Classification results for the six maturity of tomatoes by using Support Vector Machine discriminant analysis for optimal mean\_SR spectra (mean\_SR 15) and mean spectra of all SR spectra <sup>a</sup>.

Type of Spectra	Maturity	Training Set/%	Test Set/%						Accuracy
			G	B	T	P	L	R	
Mean_SR 15	G	100	42	1	0	0	0	0	100
	B	100	0	35	1	0	0	0	94.6
	T	100	0	0	39	0	0	0	97.5
	P	100	0	1	0	40	0	0	100
	L	100	0	0	0	0	46	1	100
	R	100	0	0	0	0	0	34	97.1
Mean_all_SR Spectra	G	98.3	42	0	0	0	0	0	100
	B	98.4	0	35	5	0	0	0	94.6
	T	96.7	0	2	30	2	0	0	75.0
	P	98.3	0	0	5	38	0	1	95.0
	L	100	0	0	0	0	41	1	89.1
	R	100	0	0	0	0	5	33	94.3

<sup>a</sup> G: green; B: breaker; T: Turning; P: Pink; L: light red; R: red.



Sensitivity and specificity were two important parameters for assessing the performance of SVMDA models. The definition of the sensitivity is the proportion of true positives that are correctly identified, while specificity is defined as the proportion of true negatives that are correctly identified. Further breakdowns of sensitivity and specificity at six maturity stages using mean\_SR 15 and mean\_all\_SR spectra were presented in Table 4. Mean\_SR 15 had consistently higher values for both sensitivity and specificity than that of mean\_all\_SR spectra except green stage, indicating better model performance for mean\_SR 15. In green stage, it is interesting that mean\_all\_SR spectra had excellent results for both sensitivity and specificity with the value of 1.000, which suggested mean\_all\_SR spectra had ascendancy in assessment of green tomatoes. In addition, for both of two spectra, the values of specificity in six maturity stages were more stable than sensitivities, specifically for mean\_SR 15 with only 0.5% difference between optimal and worst values, which demonstrated the models had reliable performance for tomato maturity classification.

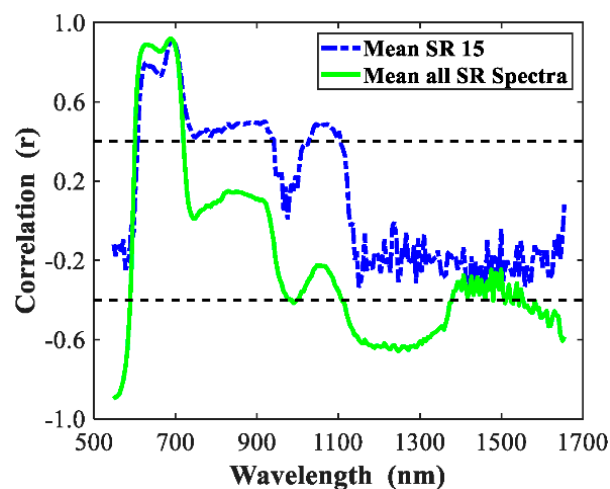
**Table 4.** Performance of the SVMDA models developed by optimal mean\_SR spectra and mean all SR spectra <sup>a</sup>.

Parameters	Spectral Type	G	B	T	P	L	R
Sensitivity	Mean_SR 15	1.000	0.946	0.975	1.000	1.000	0.971
	Mean_All_SR_Spectra	1.000	0.946	0.750	0.950	0.891	0.943
Specificity	Mean_SR 15	0.995	0.995	1.000	0.995	0.995	1.000
	Mean_All_SR_Spectra	1.000	0.975	0.980	0.970	0.995	0.976

<sup>a</sup> G: green; B: breaker; T: Turning; P: Pink; L: light red; R: red.

### 3.4. Feature Extraction

Correlation curve between the maturity stages and single wavelengths of the visible and near-infrared (Vis/NIR) spectra is presented in Figure 6. The spectral correlations of mean\_SR 15 and mean\_all\_SR\_Spectra lower than 0.4 were removed to extract these variables that had high correlation with wavelengths, which would obtain two new spectra. The classification results for these two new spectra were shown in Table 5. Compared with the full wavelength spectra, new mean\_SR 15 had relatively poor results in both the training and test sets, which suggested each of the variables for mean\_SR 15 spectra provided a contribution of model building. While for new mean\_all\_SR spectra, after removing these variables that had low correlation with wavelengths, the classification results became better in training set, and the classification results were the same in test set except turning stage with a bit change, which demonstrated that variable selection was effective for mean\_all\_SR\_Spectra.



**Figure 6.** Correlation curves for tomato between six maturity stages and relative reflectance of individual wavelengths for the spectral region of 550–1650 nm.

**Table 5.** Classification results for the six maturity of tomatoes by using Support Vector Machine discriminant analysis for new mean\_SR 15 and new mean\_all\_SR spectra.

Type of Spectra	Maturity	Training Set/%	Test Set/%						Accuracy
			G	B	T	P	L	R	
New_mean_SR 15	G	98.3	40	5	0	0	0	0	95.2
	B	88.9	2	28	0	0	0	0	75.7
	T	88.3	0	3	33	4	0	0	82.5
	P	86.7	0	1	7	35	2	1	87.5
	L	94.4	0	0	0	1	34	4	73.9
	R	98.4	0	0	0	0	10	30	85.7
New_mean_all_SR_Spectra	G	100	42	1	0	0	0	0	100
	B	98.4	0	35	5	0	0	0	94.6
	T	96.7	0	1	29	2	0	0	72.5
	P	100	0	0	5	38	1	1	95.0
	L	100	0	0	1	0	41	1	89.1
	R	100	0	0	0	0	4	33	94.3

### 3.5. Discussion

This research demonstrated that the source-detector distance had large influence on tomato maturity classification. The classification accuracies overall raised as the source-detector distance increased, and these SR spectra at source-detector distance of 16.0–36.0 mm overall gave more consistent classification results. For these SR spectra with a large source-detector distance, the transport path of photons become longer, the more interactions of photons with objects, resulting in more absorption and scattering in the sample tissue [32]. Quality parameters like SSC, pH and firmness were directly related to the maturity, which is, in turn, influenced by the pigment contents of the fruit tissue. As indicated above, light absorption is the annihilation of a photon by matter, which depends on the chemical composition of the medium, while scattering of photons is a physical process that largely depends on the structural properties of the sample. This may explain the SR spectra with large source-detector distances had relative higher classification results. Moreover, the SR spectra at the same source-detector distances with varied positions showed distinct results. Tomato fruit is heterogeneous, consisting of an outer wall of pericarp with a dense tissue structure, followed by locular cavities filled by a clear substance containing the seeds. During the ripening process, the changes of chemical constituents, structural properties, and pigments, as well as the changes of cell walls and pectin solubilisation resulted in an ununiformed composition and texture [1], which could be possible to explain the various results for different position in the same detection layer. Furthermore, the combination of SR spectra, which is the mean of all SR spectra, could not provide enhanced results, which suggested each of individual SR spectra contained interference information or irrelevant variables, after averaging all SR spectra, feature information may decrease, and introduce more interference information or irrelevant variables, leading to unsatisfying results.

The classification results in this study were also comparable or better with previous studies using near infrared spectroscopy. Sirisomboon et al. [18] obtained classification results of 96.85% and 100% for mature green and red tomatoes, respectively, Tiwari et al. [33] correctly predicted 71% of immature and 85% of mature green tomatoes and Clément [34] predicted the tomato maturity stage with a coefficient of determination of 0.93. Accurate comparison of the results obtained by spatially resolved spectroscopic system in this research with previous studies can be difficult, because many factors (e.g., growth condition, variety, methods used for sampling and data processing, etc.) can influence the classification results. From the results in Tables 2 and 3, the classification results of the six maturity stages of tomato by the spatially resolved spectroscopic system basically exceeds 90%, which demonstrated this technology is effective for classifying tomato maturity stages. Moreover, a follow-up study should be considered to further improve the classification results, such as further optimization of the data (i.e., feature extraction, variable selection, advanced data preprocessing, and modeling techniques).

#### 4. Conclusions

The paper reported on assessment of tomato maturity in different layers using spatially resolved spectroscopy over the wavelength range of 550–1650 nm. Thirty individual SR spectra were used to establish the SVMDA models and the classification results for two sides of an acquisition probe were compared. Mean\_SR spectra were adopted to evaluate the influence of source-detector distances for maturity classification. Moreover, the mean of 30 individual SR spectra was also applied to analyze and compare the performance of the models. The classification results indicated that tomato maturity classifications were influenced by the source-detector distances as well as the detection location of the tomato. The mean of all SR spectra could not achieve satisfactory results, which demonstrated that more information about samples sometimes could not obtain improved results. The mean\_SR 15 acquired the optimal classification results for tomato maturity with recognition rates of 100%, 94.6%, 97.5%, 100%, 100%, and 97.1% for green, breaker, turning, pink, light red, and red stages. In this study, spatially resolved spectroscopy was employed to assess the feasibility for tomato maturity measurement in different layers.

**Author Contributions:** Manuscript writing and data analysis, Y.H.; Experimental methodology, W.S.; Reviewing and editing, K.C.; Reviewing and Manuscript writing, Y.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** The work was partly supported by the Natural Science Fund for Colleges and Universities in Jiangsu Province (19KJB210003), the Fundamental Research Funds for the Central Universities (KJQN202045), Natural and Science Foundation of China (31901769), the Science and Technology Innovation Project for Overseas Students in Nanjing (013040106).

**Acknowledgments:** The authors would like to thank Zhao Zhang in the USDA/ARS postharvest engineering laboratory at Michigan State University for his help in carrying out the experiment reported in this paper.

**Conflicts of Interest:** The authors declare no conflict of interest.

#### References

1. Zhu, Q.; He, C.; Lu, R.; Mendoza, F.; Cen, H. Ripeness evaluation of ‘Sun Bright’ tomato using optical absorption and scattering properties. *Postharvest Biol. Technol.* **2015**, *103*, 27–34. [[CrossRef](#)]
2. Xie, L.; Ying, Y.; Ying, T.; Yu, H.; Fu, X. Discrimination of transgenic tomatoes based on visible/near-infrared spectra. *Anal. Chim. Acta* **2007**, *584*, 379–384. [[CrossRef](#)] [[PubMed](#)]
3. Dorais, M.; Ehret, D.L.; Papadopoulos, A.P. Tomato (*Solanum lycopersicum*) health components: From the seed to the consumer. *Phytochem. Rev.* **2008**, *7*, 231. [[CrossRef](#)]
4. Qin, J.; Lu, R. Measurement of the optical properties of fruits and vegetables using spatially resolved hyperspectral diffuse reflectance imaging technique. *Postharvest Biol. Technol.* **2008**, *49*, 355–365. [[CrossRef](#)]
5. Zhang, L.; McCarthy, M.J. Measurement and evaluation of tomato maturity using magnetic resonance imaging. *Postharvest Biol. Technol.* **2012**, *67*, 37–43. [[CrossRef](#)]
6. Takahashi, N.; Maki, H.; Nishina, H.; Takayama, K. Evaluation of Tomato Fruit Color Change with Different Maturity Stages and Storage Temperatures Using Image Analysis. *IFAC Proc. Vol.* **2013**, *46*, 147–149. [[CrossRef](#)]
7. Huang, Y.-P.; Liu, Y.; Yang, Y.-T.; Zhang, Z.-W.; Chen, K.-J. Assessment of Tomato Color by Spatially Resolved and Conventional Vis/NIR Spectroscopies. *Spectrosc. Spectr. Anal.* **2019**, *39*, 3585–3591.
8. Arakeri, M.P. Computer Vision Based Fruit Grading System for Quality Evaluation of Tomato in Agriculture industry. *Procedia Comput. Sci.* **2016**, *79*, 426–433. [[CrossRef](#)]
9. Arefi, A.; Motlagh, A.M.; Mollazade, K.; Teimourlou, R.F. Recognition and localization of ripen tomato based on machine vision. *Aust. J. Crop Sci.* **2011**, *5*, 1144.
10. Wan, P.; Toudeshki, A.; Tan, H.; Ehsani, R. A methodology for fresh tomato maturity detection using computer vision. *Comput. Electron. Agric.* **2018**, *146*, 43–50. [[CrossRef](#)]
11. Xie, C.; Liu, Y.; Zeng, W.; Lu, X. An improved method for single image super-resolution based on deep learning. *Signal Image Video Process.* **2018**, *13*, 557–565. [[CrossRef](#)]

12. Shen, L.; Wang, H.; Liu, Y.; Liu, Y.; Zhang, X.; Fei, Y. Prediction of Soluble Solids Content in Green Plum by Using a Sparse Autoencoder. *Appl. Sci.* **2020**, *10*, 3769. [[CrossRef](#)]
13. Nicolai, B.M.; Beullens, K.; Bobelyn, E.; Peirs, A.; Saeys, W.; Theron, K.I.; Lammertyn, J. Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biol. Technol.* **2007**, *46*, 99–118. [[CrossRef](#)]
14. Jiang, H.; Jiang, X.; Ru, Y.; Chen, Q.; Xu, L.; Zhou, H. Sweetness detection and grading of peaches and nectarines by combining short- and long-wave fourier-transform near-infrared spectroscopy. *Anal. Lett.* **2020**, 1–20. [[CrossRef](#)]
15. Ecartot, M.; Tessarotto, L.; Chervin, C. Rapid phenotyping of the tomato fruit model, Micro-Tom, with a portable VIS-NIR spectrometer. *Plant Physiol. Biochem.* **2013**, *70*, 159–163. [[CrossRef](#)] [[PubMed](#)]
16. Tilahun, S.; Park, D.S.; Mu, H.S.; Hwang, I.G.; Kim, S.H.; Han, R.C.; Jeong, C.S. Prediction of lycopene and  $\beta$ -carotene in tomatoes by portable chroma-meter and VIS/NIR spectra. *Postharvest Biol. Technol.* **2018**, *136*, 50–56. [[CrossRef](#)]
17. Alenazi, M.M.; Shafiq, M.; Alsadon, A.A.; Alhelal, I.M.; Saad, M.A.O. Non-destructive Assessment of Flesh Firmness and Dietary Antioxidants of Greenhouse-grown Tomato (*Solanum lycopersicum* L.) at Different Fruit Maturity Stages. *Saudi J. Biol. Sci.* **2020**, *27*, 2839–2846. [[CrossRef](#)]
18. Sirisomboon, P.; Tanaka, M.; Kojima, T.; Williams, P. Nondestructive estimation of maturity and textural properties on tomato ‘Momotaro’ by near infrared spectroscopy. *J. Food Eng.* **2012**, *112*, 218–226. [[CrossRef](#)]
19. Clément, A.; Dorais, M.; Vernon, M. Nondestructive Measurement of Fresh Tomato Lycopene Content and Other Physicochemical Characteristics Using Visible–NIR Spectroscopy. *J. Agric. Food Chem.* **2008**, *56*, 9813–9818. [[CrossRef](#)]
20. Ohnishi, M.; Kusakawa, N.; Masaki, S.; Honda, K.; Shimada, Y.; Fujimoto, I.; Hirao, K. Investigation on deep layer measurements in the cerebral cortex within the adult head by near infrared spectroscopy using an absorbance difference technique. *J. Near Infrared Spectrosc.* **2003**, *11*, 27–38. [[CrossRef](#)]
21. Huang, Y.; Lu, R.; Chen, K. Development of a multichannel hyperspectral imaging probe for property and quality assessment of horticultural products. *Postharvest Biol. Technol.* **2017**, *133*, 88–97. [[CrossRef](#)]
22. Do Trong, N.N.; Erkinbaev, C.; Tsuta, M.; De Baerdemaeker, J.; Nicolai, B.; Saeys, W. Spatially resolved diffuse reflectance in the visible and near-infrared wavelength range for non-destructive quality assessment of ‘Braeburn’ apples. *Postharvest Biol. Technol.* **2014**, *91*, 39–48. [[CrossRef](#)]
23. Xia, J.; Weaver, A.; Gerrard, D.E.; Yao, G. Distribution of optical scattering properties in four beef muscles. *Sens. Instrum. Food Qual. Saf.* **2008**, *2*, 75–81. [[CrossRef](#)]
24. Huang, Y.; Lu, R.; Chen, K. Assessment of Tomato Soluble Solids Content and pH by Spatially-Resolved and Conventional Vis/NIR Spectroscopy. *J. Food Eng.* **2018**, *236*, 19–28. [[CrossRef](#)]
25. Huang, Y.; Lu, R.; Hu, D.; Chen, K. Quality assessment of tomato fruit by optical absorption and scattering properties. *Postharvest Biol. Technol.* **2018**, *143*, 78–85. [[CrossRef](#)]
26. USDA. *U.S. Standards for Grades of Fresh Tomatoes*. United States Department of Agriculture; Agricultural Marketing Service: Washington, DC, USA, 1991.
27. Huang, Y.; Yang, Y.; Sun, Y.; Zhou, H.; Chen, K. Identification of Apple Varieties Using a Multichannel Hyperspectral Imaging System. *Sensors* **2020**, *20*, 5120. [[CrossRef](#)]
28. Huang, Y.; Lu, R.; Chen, K. Detection of internal defect of apples by a multichannel Vis/NIR spectroscopic system. *Postharvest Biol. Technol.* **2020**, *161*, 111065. [[CrossRef](#)]
29. Huang, Y.; Lu, R.; Chen, K. Prediction of Firmness Parameters of Tomatoes by Portable Visible and Near-Infrared Spectroscopy. *J. Food Eng.* **2018**, *222*, 185–198. [[CrossRef](#)]
30. Huang, Y.; Lu, R.; Xu, Y.; Chen, K. Prediction of tomato firmness using spatially-resolved spectroscopy. *Postharvest Biol. Technol.* **2018**, *140*, 18–26. [[CrossRef](#)]
31. Wang, A.; Wei, X. A sequential method for estimating the optical properties of two-layer agro-products from spatially-resolved diffuse reflectance: Simulation. *Artif. Intell. Agric.* **2019**, *3*, 69–78. [[CrossRef](#)]
32. Lu, R. *Overview of Light Interaction with Food and Biological Materials*; CRC Press: Abingdon, UK, 2016; pp. 19–42.
33. Tiwari, G.; Slaughter, D.C.; Cantwell, M. Nondestructive maturity determination in green tomatoes using a handheld visible and near infrared instrument. *Postharvest Biol. Technol.* **2013**, *86*, 221–229. [[CrossRef](#)]

34. Clément, A.; Dorais, M.; Vernon, M. Multivariate approach to the measurement of tomato maturity and gustatory attributes and their rapid assessment by Vis–NIR Spectroscopy. *J. Agric. Food Chem.* **2008**, *56*, 1538–1544. [[CrossRef](#)] [[PubMed](#)]

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).