光譜角映射法與支撐向量機應用於蜜棗表面缺陷辨識之比較

范國天、*劉乃上
南臺科技大學機械工程系
*nliou@stust.edu.tw

摘要

本研究比較與探討應用光譜角映射法與支撐向量機於蜜棗表面缺陷檢測之精確度。為了獲得蜜棗表面的高光譜影像，使用了波長範圍為 468-950 nm 的推掃式高光譜影像系統對包含銹斑、裂果、發霉、過熟腐爛等特徵的高雄 11 號(珍蜜)蜜棗進行掃描。應用相關高光譜影像資料建立了光譜角映射法與支撐向量機蜜棗表面缺陷分類模型對蜜棗表面缺陷進行辨識。由於蜜棗有弧度的表面光滑並具有蠟質，進行光譜影像掃描時表面的眩光無法避免，因此在本研究中將表面的眩光視為蜜棗表面的特徵，在建立分類模型時眩光被視為分類的類別之一。由相關模型的辨識結果建立混淆矩陣以評估模型分類有效性。由光譜角映射法與支撐向量機模型對蜜棗表面缺陷分類的混淆矩陣顯示光譜角映射法模型在分類時對於裂果與眩光的混淆度最高。支撐向量機模型對銹斑的辨識度較低，對白色發霉的部分辨識度最高。使用相關分類模型對驗證數據的分類結果顯示光譜角映射法與支撐向量機模型的分類準確性分別為 71.5%和 97.2%。由此研究可知由波長範圍為 468-950 nm 的高光譜數據所建立的支撐向量機模型能對蜜棗表面缺陷進行良好的辨識。

關鍵詞：高光譜影像，蜜棗表面缺陷檢測，支撐向量機，光譜角映射法

Investigating the Use of Spectral Angle Mapper and Support Vector Machine for Surface Defect Detection of Jujubes

Quoc Thien Pham, *Nai-Shang Liou
Department of Mechanic Engineering, Southern Taiwan University of Science and Technology

Abstract

This study investigated the classification accuracy of spectral angle mapper (SAM) and support vector machine (SVM) models for surface defect detection of jujubes with the use of hyperspectral image data. Hyperspectral imaging data of "Kaohsiung 11" jujubes with surface defects including rusty spots, decay, white fungus, black fungus, and crack were obtained by using a custom made hyperspectral imaging system. Since the surface of jujubes is smooth and waxy, the glare on the surface cannot be avoided during spectral image scanning. Therefore, in this study, the glare on the surface is regarded as the characteristics of jujubes. Confusion matrices were used to evaluate the effectiveness of models. The classification accuracy of SAM model was 71.5%. For SAM classifier, the confusion between crack and glare is the highest. The classification accuracy of SVM model was 97.2%. For SVM model, the prediction accuracy of rusty spots is the lowest and the accuracy of white fungus is the highest. The high classification accuracy of SVM model indicated that the SVM classifier, with the use of hyperspectral imaging data, is effective for classifying surface defect of jujubes.

Keywords: Hyperspectral Imaging, Jujube Surface Defect Detection, Support Vector Machine, Spectral Angle Mapper

Received: Jan. 14, 2021; first revised: Apr. 19, 2021, accepted: Apr. 2021.
Corresponding author: N.S. Liou, Department of Mechanic Engineering, Southern Taiwan University of Science and Technology, Tainan 710031, Taiwan.

I. Introduction

Hyperspectral imaging (HSI) technology is a technology which incorporates conventional imaging and spectroscopy to provide both spatial and spectral information of an object. The spectra of each spatial pixel containing the signatures of substances presented at the corresponding sample points on the hyperspectral image [1]. This technique can be used to evaluate the chemical or physical properties such as sugar [2–3], starch [4], soluble solids [1, 5] and moisture [6–7] of fruits. HSI technology is also capable of identifying and detecting spectral and spatial anomalies in agricultural products [5]. The use of HSI to detect defects/diseases of fruits has been studied: bruise on apple and pear [8–9], chilling injury in peaches [10–11], decay in citrus [12–13] and fungal infections in strawberry [14]. Jiang et al. [8] employed principal component analysis (PCA) to extract five characteristic wavelengths (472, 544, 655, 688 and 967 nm). Two wavelengths (472 and 967 nm) were used to create the classification model by applying partial least squares discriminant analysis (PLS-DA) for bruise and normal skin on Korla pears. In order to distinguish nine types of skin conditions (i.e., normal, skin injury, scarring, insect damage, puncture injury, decay, disease spots, dehiscent scarring, and anthracnose) of bi-color peaches, Li et al. [15] used band ratio ($Q_{781/848}$) method and a simple thresholding to develop classification algorithms for identifying defect surface from normal surface. Folch-Fortuny et al. [12] developed multivariate models by an N-way partial least squares regression discriminant analysis (NPLS-DA) method for detecting infected citrus fruits. Siedliska et al. [14] reported the use of near and shortwave infrared (VNIR/SWIR) hyperspectral imaging techniques for early detection of fungal infections and predicted anthocyanin content (AC) and soluble solid content (SSC) in strawberries during storage.

In this study, the use of hyperspectral image data to classify surface defects of Kaohsiung 11 jujubes was investigated. A line-scan hyperspectral imaging system with a wavelength range of 468-950 nm was used to obtain the reflection spectral of Kaohsiung 11 jujube with surface defects. Spectral angle mapper (SAM) and support vector machine (SVM) were used to investigate the feasibility of using hyperspectral imaging in detecting the various common defects that exist on "Kaohsiung 11" jujube surfaces. The SAM is a linear model which has been widely utilized for remote sensing image, and SVM is a supervised machine learning method with good generalization ability suitable for high-dimensional data. The performance of SAM model and SVM model was compared.

II. Material and Method

1. Jujube Samples

The Kaohsiung 11 "Honey" jujubes were used as the specimens. Jujube with normal surfaces and five types of common skin defects (i.e., rusty spots, decay, white fungus, black fungus, and crack) were obtained and stored in the inspection laboratory at 25°C before experiment. The hyperspectral images of specimens were acquired within two days after samples obtained. Typical images for normal and defect skins of jujubes are shown in Fig. 1.

Fig. 1 Skin conditions of jujube used in this study

2. Dala collection and pre-process

The spectral images were collected in a dark room using the custom-made hyperspectral imaging system [16] shown in Fig.2 Hyperspectral data collected from 80 jujube specimens were used in this study. Training spectral data of pixels were obtained from a total of 50 randomly selected jujube specimens to develop SAM and SVM pixelwise classification models. The remaining jujube samples were used to evaluate the performance of the classifiers proposed in this work. To reduce the effects of luminous source and dark currents of the camera, white and dark references were used to calibrate hyperspectral images and compute the relative reflectance. The white reference images were captured with the use of a white diffuse board and the dark reference images were acquired by turning off the luminous sources and covering the lens with a black cap [17]. The calibrated images were obtained from white reference image, black reference image and original image. In this study, glare was considered as one type of surface feature due to the mirror-like reflection or glare reflected from the smooth and waxy skin of the sphere-like jujube surface cannot be completely removed by adjusting the angle of the incident light. Thus seven types of jujube skin conditions (i.e., normal, rusty spots, decay, white fungus, black fungus, crack, and glare) were visually identified and labeled for the use of SAM and SVM methods.

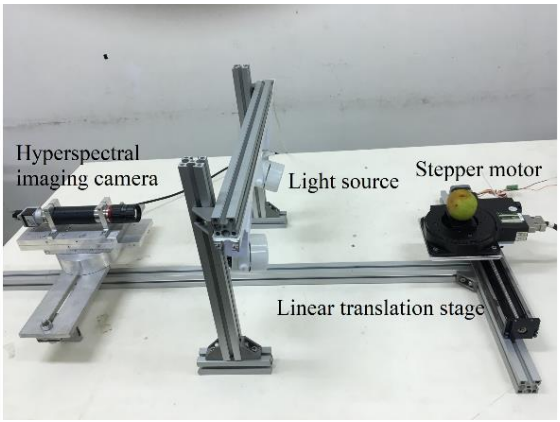


Fig. 2 The design of hyperspectral imaging system

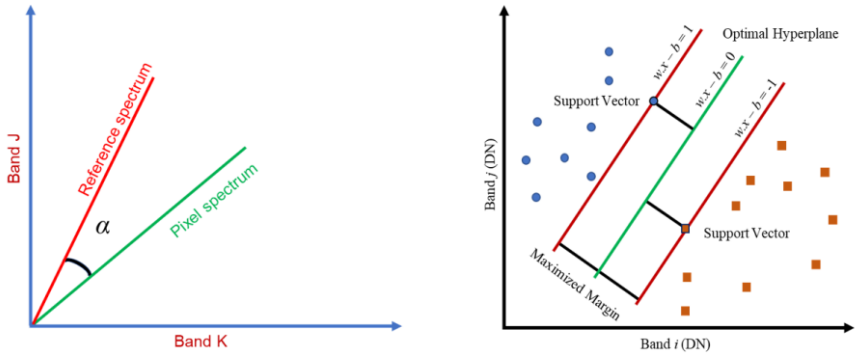
3. Classification Methods

SAM treats spectral data as vectors in n-dimensional space and measures the spectral similarity by computing the angle between image spectra to reference spectra [18]. In the application of hyperspectral data analysis, SAM determines the spectral similarity between two spectra by computing the angle between the spectra and treating them as vectors in a space with dimensionality same as the number of bands. Small angles between two spectrums indicate high similarity. For vector in two-dimensional space, the geometric interpretation of SAM can be shown in Fig 3(a) and the angle between two vectors in n-dimensional space can be computed by the following equation.

$$\alpha = \cos^{-1}\left(\frac{\sum_{i=1}^{nb} t_i r_i}{\sqrt{\sum_{i=1}^{nb} t_i^2}\sqrt{\sum_{i=1}^{nb} r_i^2}}\right). \quad (1)$$

In equation (1), nb, t and r are the number of bands, the spectrum of pixel and reference spectrum respectively and α is the spectral angle. In this study, the reference spectrum of each skin feature was computed from the training data. Support vector machine (SVM) is a supervised learning method used for feature selection, regression and classification. The SVM works well with limited training samples of high-dimensional data. In the case of linearly separable classes, the SVM attempts to find the optimal separation hyperplane between classes based on training data Fig 3(b). SVM can be generalized for a nonlinear problem with the use of kernel function. In the work, the Radial Basis Function (RBF) was chosen as the kernel function of SVM, and grid search was used to select the optimal penalty parameter C and the kernel function parameter g.

Pixels with hyperspectral data were selected from 50 randomly chosen jujube samples. Each skin feature had 10000 pixels of data. The selected data were divided into a training set (70%) and a validation set (30%). The training set was used to establish the SAM reference spectra and SVM classification model. The validation set was used for evaluating the performance of SAM and SAM models. Classification models were evaluated based on the confusion matrices of the classification performed on the validation sets.



(a)SAM method (b)SVM method

Fig. 3 Schematics of SAM and SVM methods

III. Results and Discussions

1. Spectral Characteristics

The average spectra of different types of jujube skin conditions were calculated from the training set and shown in Fig. 4. These average spectra were used as references for SAM classification. The shape of the spectrum curve of glare is similar to that of normal surface and it is possibly because that most of the glare pixels in the training set are selected from the normal glossy surface. Moreover, the spectral characteristics of white and black fungus are similar and lower than other skin conditions.

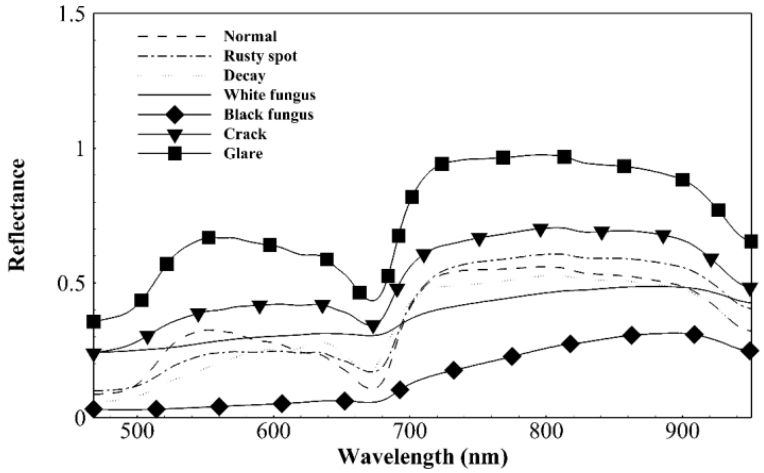


Fig. 4 Average reflectance spectra of jujube samples with different skin conditions

2. Classification Models

The classification accuracy, predicted from the validation set, of SAM and SVM are 71.5% and 97.2% respectively. It can be noted that the accuracy of the SAM model was lower than that of the SVM model. The confusion matrices were created to analyze and evaluate model performance and are shown in Table 1. The rows of the matrix represent actual classes and the columns denote predicted classes. The diagonal entries of the matrixes denote the percentage of pixels of the properly-recognized cases, while each entry outside the diagonal

means an error (the percentage of misclassification). For SVM model, the prediction accuracy of rusty spots is the lowest.

Table 1 Confusion matrices of SAM (upper) and SVM (lower) models

Actual	Classified as						
	Normal (%)	Rusty spots (%)	Decay (%)	White fungus (%)	Black fungus (%)	Crack (%)	Glare (%)
Normal	77.23	7.37	0.01	0	0	0	15.37
	98.71	0.4	0.14	0	0	0	0.75
Rusty spots	15.14	78.07	4.28	0	0.97	1.26	0.27
	4.72	89.14	4.28	0.21	0.42	0.53	0.69
Decay	1.15	16.8	63.41	0	9.85	5.96	2.82
	0.11	0.83	98.48	0.02	0.45	0.07	0.03
White fungus	0	3.33	0.13	79.8	6.49	10.2	0.06
	0	0.18	0.1	99.11	0.42	0.2	0
Black fungus	0	3.88	0.53	6.71	88.48	0.39	0
	0.02	0.28	4.29	2.32	93.09	0	0
Crack	0.59	11.15	5.97	6.27	0.2	44.21	31.61
	0.07	1.12	0.86	3.33	0	93.77	0.86
Glare	9.06	2.95	0.1	2.74	0.02	13.39	71.74
	0.96	0.58	0.27	0.12	0	1.34	96.72

SAM SVM

For SAM classifier, the confusion between crack and glare is the highest. The reason could be the color of the newly cracked jujube is close to glare, leading to labelling error from RGB images. Besides, spectral curves are quite similar but differ in reflectance value. Pixels of crack located in areas of high intensity, which can be misrecognized as glare. The angle between glare reference and crack reference is 3.7 degree. The average angles and standard deviation of crack sample respect to crack reference are 5.13 and 2.36 degree respectively. And the average angles and standard deviation of glare sample respect to glare reference are 4.78 and 2.34 degree respectively. It can be seen from the bell shape curves generated from the average, standard deviation and the angle between glare and crack references that glare and crack could be easily misclassified by SAM algorithm (Fig. 5.). Furthermore, the confusion between decay and rusty spots is hight. The confusion could be due to some decay pixels and rusty spot pixels have very similar spectral signatures.

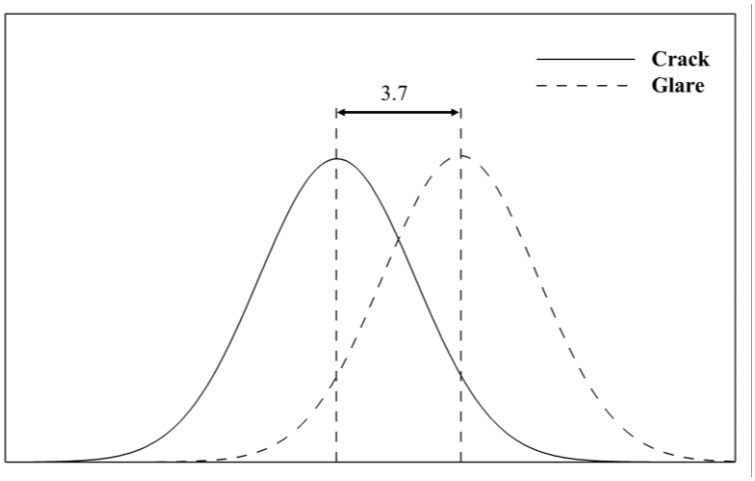


Fig. 5 Bell-shaped curves generated from the average, standard deviation and the angle between glare and crack references

Fig. 6. shows the classification results of five jujube samples using the proposed SAM and SVM classification models. For each sample, the RGB image was shown in the first row, and the classification results of SAM and SVM models were shown in the second and third rows respectively. For the SAM model, the misclassification of normal as rusty spot and decay as rusty spot can be seen in the predicted result of all jujube samples. The possible reason could be the intensity obtained from the pixels of normal surface near the edge is low due to the curved shape of jujube. Thus normal surface near the edge was misclassified as a low-intensity skin feature such as rust surface. With the use of SVM model, most of the fruit pixels, excepting the pixels near the left boundary of the first sample, can be properly classified.

In this work, SAM and SVM models were established by using the full range spectral data. The high information redundancy in the full range hyperspectral data can limit the performance of classification models such as SAM and SVM models. This greatly affects the classification time in the case of developing an online classification system. To improve classification time, dimension reduction techniques should be used to reduce data redundancy of the models while retaining important features of the hyperspectral data.

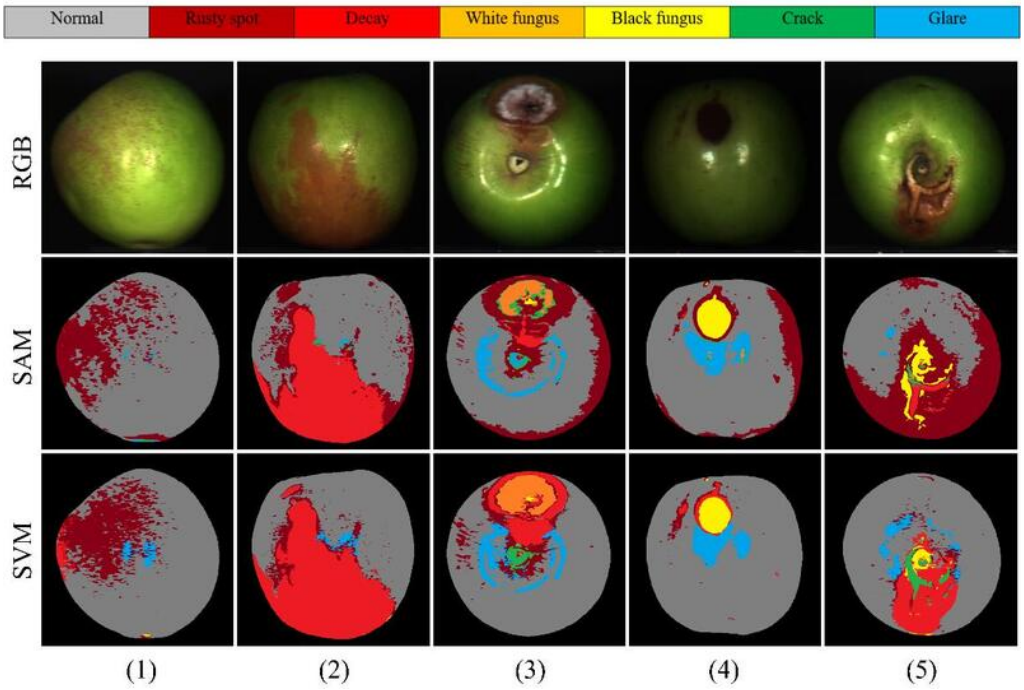


Fig 6. Color images and skin types predicted by SAM and SVM models

IV. Conclusion

In this study, SAM and SVM models were used to classify skin defects of Kaohsiung 11 "Honey" jujube, and the performance of these two models was investigated. The results showed that the accuracy of SVM model is higher than that of SAM model. The classification accuracy of SVM was 97.2%, which indicated that the SVM classifier, with the use of hyperspectral imaging data, is effective for classifying the surface defect of jujube. Hyperspectral data of the whole spectral range were to train the classification models in this work. Classification speed is critical when using hyperspectral imaging technique for online application. To improve the classification speed, data reduction method such as principal component analysis can be used to select important wavelengths for skin defect classification.

Reference

[1] B.H. Zhang, J.B. Li, S.X. Fan, W.Q. Huang, C.J. Zhao, C.L. Liu and D.F. Huang. (2015). Hyperspectral imaging combined with multivariate analysis and band math for detection of common defects on peaches (Prunus persica). *Comput. Electron. Agric.*, *114*, 14–24.

[2] J.B. Li and L.P. Chen (2017). Comparative analysis of models for robust and accurate evaluation of soluble solids content in 'Pinggu' peaches by hyperspectral imaging. *Comput. Electron. Agric.*, *142*, 524–535.

[3] A. Kjaer, G. Nielsen, S. Staerke, M.R. Clausen, M. Edelenbos and B. Jorgensen (2017). Detection of glycoalkaloids and chlorophyll in potatoes (solanum tuberosum l.) by hyperspectral imaging. *Am. Potato J., 94*, 573–582.

[4] S.X. Fan, C.Y. Li, W.Q. Huang and L.P. Chen (2017). Detection of blueberry internal bruising over time using NIR hyperspectral reflectance imaging with optimum wavelengths. *Postharvest Biol. Technol., 134*, 55–66.

[5] B.H. Zhang, S.X. Fan, J.B. Li, W.Q. Huang, C.J. Zhao, M. Qian and L. Zheng (2015). Detection of Early Rottenness on Apples by Using Hyperspectral Imaging Combined with Spectral Analysis and Image Processing. *Food Anal. Methods*, *8*, 2075–2086.

[6] B.S. Zhan, J.H. Ni and J. Li (2014). Hyperspectral Technology Combined with CARS Algorithm to Quantitatively Determine the SSC in Korla Fragrant Pear. *Spectrosc Spect. Anal.*, *34*, 2752–2757.

[7] D. Yang, D.D. He, A.X. Lu, D. Ren and J.H. Wang (2017). Combination of spectral and textural information of hyperspectral imaging for the prediction of the moisture content and storage time of cooked beef. *Infrared Phys. Technol.*, *83*, 206–216.

[8] H. Jiang, C. Zhang, Y. He, X.X. Chen, F. Liu and Y.D. Liu (2016). Wavelength selection for detection of slight bruises on pears based on hyperspectral imaging. *Appl. Sci.*, *6*, 450.

[9] V. Leemans, H. Magein and M.F. Destain (2002). On-line fruit grading according to their external quality using machine vision. *Biosyst Eng.*, *83*, 397–404.

[10] L.Q. Pan, Q. Zhang, W. Zhang, Y. Sun, P.C. Hu and K. Tu (2016). Detection of cold injury in peaches by hyperspectral reflectance imaging and artificial neural network. *Food Chem.*, *192*, 134–141.

[11] Y. Sun, X.Z. Gu, K. Sun, H.J. Hu, M. Xu, Z.J. Wang, K. Tu and L.Q. Pan (2017). Hyperspectral reflectance imaging combined with chemometrics and successive projections algorithm for chilling injury classification in peaches. *LWT-Food Sci. Technol.*, *75*, 557–564.

[12] A. Folch-Fortuny, J.M. Prats-Montalban, S. Cubero, J. Blasco and A. Ferrer (2016). VIS/NIR hyperspectral imaging and N-way PLS-DA models for detection of decay lesions in citrus fruits. *Chemom. Intell. Lab. Syst.*, *156*, 241–248.

[13] S.X. Fan, C.Y. Li, W.Q. Huang and L.P. Chen (2018). Data fusion of two hyperspectral imaging systems with complementary spectral sensing ranges for blueberry bruising detection. *Sensors*, *18*, 4463.

[14] A. Siedliska, P. Baranowski, M. Zubik, W. Mazurek and B. Sosnowska (2018). Detection of fungal infections in strawberry fruit by VNIR/SWIR hyperspectral imaging. *Postharvest Biol. Technol.*, *139*, 115–126.

[15] J.B. Li, L.P. Chen, W.Q. Huang, Q.Y. Wang, B.H. Zhang, X. Tian, S.X. Fan and B. Li (2016). Multispectral detection of skin defects of bi-colored peaches based on vis-NIR hyperspectral imaging. *Postharvest Biol. Technol.*, *112*, 121–133.

[16] J.U. Porep, D.R. Kammerer and R. Carle (2015). On-line application of near infrared (NIR) spectroscopy in

food production. *Trends Food Sci. Technol., 46*, 211–230.

[17] O. Kleynen, V. Leemans and M.F. Destain (2005). Development of a multi-spectral vision system for the detection of defects on apples. *J. Food Eng.*, *69*, 41–49.

[18] J.H. Cheng, D.W. Sun, H.B. Pu, X.H. Chen, Y.L. Liu, H. Zhang and J.L. Li (2015). Integration of classifiers analysis and hyperspectral imaging for rapid discrimination of fresh from cold-stored and frozen-thawed fish fillets. *J. Food Eng.*, *161*, 33–39.