Inspecting Species and Freshness of Fish Fillets Using Multimode Hyperspectral Imaging Techniques

Jianwei Qin
USDA/ARS Environmental Microbial and Food Safety Laboratory

Fartash Vasefi
SafetySpect Inc.

Rosalee S. Hellberg
Chapman University, hellberg@chapman.edu

Alireza Akhbardeh
SafetySpect Inc.

Rachel B. Isaacs
Chapman University

Recommended Citation
Inspecting Species and Freshness of Fish Fillets Using Multimode Hyperspectral Imaging Techniques

Comments
This is a pre-copy-editing, author-produced PDF of a conference proceeding accepted for publication in Proceedings of the SPIE, volume 11421, in 2020. The definitive publisher-authenticated version is available online at https://doi.org/10.1117/12.2558684

Copyright
Society of Photo-Optical Instrumentation Engineers (SPIE)

Authors
Inspecting Species and Freshness of Fish Fillets using Multimode Hyperspectral Imaging Techniques


aUSDA-ARS Environmental Microbial and Food Safety Laboratory, 303 Powder Mill Rd, Beltsville, MD 20705, USA; bSafetySpect Inc., 10100 Santa Monica Blvd., Suite 300, Los Angeles, CA 90067, USA; cSchmid College of Science and Technology, Food Science Program, Chapman University, 1 University Drive, Orange, CA 92866, USA

ABSTRACT

This study developed multimode hyperspectral imaging techniques to detect substitution and mislabeling of fish fillets. Line-scan hyperspectral images were collected from fish fillets in four modes, including reflectance in visible and near-infrared (VNIR) region, fluorescence by 365 nm UV excitation, reflectance in short-wave infrared (SWIR) region, and Raman by 785 nm laser excitation. Fish fillets of six species (i.e., red snapper, vermilion snapper, Malabar snapper, summer flounder, white bass, and tilapia) were used for species differentiation and frozen-thawed red snapper fillets were used for freshness evaluation. A total of 24 machine learning classifiers were used for fish species and freshness classifications using four types of spectral data in three different subsets (i.e., full spectra, first ten components of principal component analysis, and bands selected by a sequential feature selection method). The highest accuracies were achieved at 100% using full VNIR reflectance spectra for the species classification and 99.9% using full SWIR reflectance spectra for the freshness classification. The VNIR reflectance mode gave an overall best performance for both species and freshness inspection.

Keywords: Hyperspectral imaging, fish mislabeling, reflectance, fluorescence, Raman, machine learning

1. INTRODUCTION

With increased global trade of fish, complex supply chains, and limited monitoring, there is a rising vulnerability for fish fraud in the marketplace. A survey by the nonprofit organization Oceana found that 21% of fish sold in fish markets, grocery stores, and restaurants across the United States was mislabeled on the basis of species.1 Additional forms of fish mislabeling include labeling frozen-thawed fish as “fresh”, misrepresentation of production method, and falsification of geographical origin. Fish mislabeling is a form of economic deception, and also removes the ability for customers to make informed purchases based on conservation management practices for specific fish populations as well as potential health risks involved with certain fish. After removing morphological indicators such as heads, tails, skins, and fins, many fish fillets are similar in appearance, which makes them a vulnerable target for economically-motivated fraud. In order to avoid economic deception, there is a need for rapid detection technologies for fish mislabeling and substitution that can be used onsite by seafood importers and distributors.

Current techniques for detecting fish species are mainly based on molecular methods. DNA barcoding is commonly used to identify fish species and it has been adopted by the U.S. Food and Drug Administration.2 The DNA sequencing-based technique provides accurate identification of species through comparative analysis of sequence variation in a short fragment of the genome against an existing library of reference sequences.3 But the entire process typically requires 1–2 days of laboratory work and data analysis to identify the species of a given sample. Hence this method is not utilized onsite at processing facilities. Real-time PCR is a targeted method and cannot be used to simultaneously test for a wide range of species.4 Besides the molecular methods, traditional methods have also been used to evaluate fish and other seafoods.5 Despite high accuracies of these methods, they generally need expensive and complicated instruments and time-consuming sample preparation procedures, which prevents them from being used for rapid and high-throughput assessment of the aquatic products.

* jianwei.qin@usda.gov; phone 1-301-504-8450 ext. 244
Optical sensing techniques have been developed for quality evaluation of whole fish and fish fillet, which provide a simple, fast, low-cost, and nondestructive alternative to the conventional methods. Various spectroscopy techniques have been investigated, such as visible (VIS), near-infrared (NIR), mid-infrared (MIR), fluorescence, Raman, impedance, and nuclear magnetic resonance (NMR). Hyperspectral imaging (HSI) techniques have become a powerful tool to inspect food and agricultural products, and they have been used for quality analysis of fish and other seafoods. To our knowledge, reflectance measurement is the only HIS mode used for fish applications in the published studies, and it has been mainly carried out in visible and near-infrared (400–1000 nm) and near-infrared (900–1700 nm) wavelength ranges. Other HSI modes have not been explored. This study aimed to investigate multimode hyperspectral imaging techniques, including reflectance, fluorescence, and Raman, to detect substitution and mislabeling of fish fillets.

2. MATERIALS AND METHODS

2.1 Multimode Hyperspectral Imaging Systems

Three line-scan hyperspectral systems were used to collect four types of image data from fish fillet samples: (1) reflectance images in visible and near-infrared (VNIR) region, (2) fluorescence images by 365 nm UV excitation, (3) reflectance images in short-wave infrared (SWIR) region, and (4) Raman images by 785 nm laser excitation. A VNIR hyperspectral system was used to acquire both reflectance and fluorescence images. The reflectance and fluorescence images were acquired in spectral regions of 419–1007 nm (125 bands) and 438–718 nm (60 bands), respectively. Another similar hyperspectral system was used to acquire reflectance images in the SWIR spectral region. The SWIR reflectance images were acquired in a wavelength range of 842–2532 nm (287 bands). Raman images were acquired by a line-scan hyperspectral Raman system. The system covered a wavenumber range of 103–2831 cm⁻¹ (846 bands) with a spectral resolution of 14 cm⁻¹.

2.2 Experimental Samples and Procedures

Four fish fillets labeled as “snapper”, “flounder”, “white bass”, and “tilapia” were purchased from a local seafood market. In addition, a total of 10 fish fillets labeled as “red snapper” were purchased from three online retailers. All 14 fillets were used for the fish species differentiation study. The fish freshness evaluation study was limited to use of the red snapper fillets. The fillet samples were transported with ice packs to our laboratory and they were imaged immediately. After imaging, the red snapper fillets were frozen in a −20 °C freezer for 24 h and then thawed in a 4 °C refrigerator for 24 h. The frozen-thawed samples were reimaged using the same three systems. The same freezing and thawing process was repeated for a second cycle, and the samples were imaged for the third time to finish the image acquisition. As a result, three sets of the hyperspectral images were collected from each red snapper fillet, including an “as received” (AR) image and two images corresponding to the two freeze-thaw cycles (FT1 and FT2). Each fish fillet was placed in a sample holder with a sampling volume of 150×100×25 mm³. The lens-to-sample distance in each system was adjusted so that the length of the instantaneous field of view (IFOV) of the camera unit was slightly longer than the length of the sample holder (150 mm). Under these settings, the spatial resolutions along the IFOV direction of all three imaging systems were determined as 0.4 mm/pixel. Each fillet sample was scanned along the width direction of the sample holder using an incremental size of 0.4 mm to match the spatial resolution of the length direction.

2.3 Fish Species Authentication with DNA Barcoding

All fillet samples were DNA tested for species authentication. Before imaging, a small piece of sample was removed from the interior of each fillet using a disposable scalpel and sterile forceps and then placed in a 50 mL sterile Falcon tube. The samples were immediately frozen at −80 °C for 24 h and then shipped overnight with ice to Chapman University for DNA-based identification. DNA was extracted from ~10 mg of each sample using the DNeasy Blood and Tissue Kit (Qiagen, Germantown, MD, USA). Consensus sequences were identified based on the top species match in the Barcode of Life Database (BOLD) Animal Identification Request Engine (http://www.boldsystems.org/), Full Length Published Records.

2.4 Spectral Processing and Machine learning Classifications

Flat-field corrections were conducted on VNIR and SWIR reflectance images to convert original intensities to relative reflectance. Similar corrections were also used for original fluorescence images to obtain the relative fluorescence intensities. Fluorescence background in original Raman images was removed by a baseline correction method using adaptive iteratively reweighted penalized least squares. After preprocessing, a spatial mask was created for each image type using a selected single-band image to remove the background. Then the mean (M) and standard deviation (STD) of the fish pixel intensities within the mask were calculated. All the fish pixels were grouped into 10×10 pixel regions to
mimic point spectroscopy measurements. When 10% of the pixels were beyond the range of M±2STD, the whole 10×10 pixel region were excluded for further analysis. The 100 spectra in each remaining pixel region were averaged, and all the mean spectra were used for machine learning classifications. The four types of the spectral data were labeled using DNA test results for fish species and freshness status for red snapper fillets. The labeled data were input to Classification Learner app in MATLAB (MathWorks, Natick, MA, USA) for classifications. In addition to the full spectra, principal component analysis (PCA) was performed for the purpose of dimension reduction. Also, a sequential feature selection algorithm was used to select a subset of significant bands for each of the four types of the images. The full spectra, the first ten PCA components, and the spectral data at selected bands were all used for the machine learning classifications. Two-fold cross-validation was used to evaluate the predictive accuracies of all the classification models. Each spectral dataset was randomly partitioned into two equal-size disjoint folds. A model was trained using out-of-fold data and the performance was assessed using in-fold data. The overall accuracy of each model was obtained by calculating the average error over the two folds. A total of 24 classifiers in six general categories (i.e., decision tree, discriminant analysis, Naive Bayes, support vector machine (SVM), k-nearest neighbor (KNN), and ensemble) were tested to compare the classification accuracies for each type of the spectral data.

3. RESULTS AND DISCUSSION

3.1 DNA Test Results
The four fish fillets labeled by the local seafood market as “snapper”, “flounder”, “white bass”, and “tilapia” were identified by DNA barcoding as vermilion snapper, summer flounder, white bass, and tilapia, respectively. The DNA tests also confirmed that six “red snapper” fillets purchased from two online retailers (three from each) were correctly labeled. However, four fish fillets labeled by one online retailer as “red snapper” were determined to be mislabeled. The four samples were identified as Malabar snapper. The samples identified as red snapper, vermilion snapper, Malabar snapper, summer flounder, white bass, and tilapia were used for the species classification study, and the six authenticated red snapper fillets were used for the freshness classification study.

3.2 Hyperspectral Images and Spectra of Fish Fillets
Fig. 1 shows four types of hyperspectral images from a red snapper fillet. The four images were extracted at selected spectral peak positions. The fillet surface appears more consistent in the VNIR and SWIR reflectance images than in the fluorescence and Raman images, revealing that the fluorescence and Raman signals may be more sensitive to the fish tissue variations than the VNIR and SWIR reflectance signals.

![Red snapper (as received)](#)

![VNIR reflectance (699 nm)](/)

![Fluorescence (509 nm)](/)

![SWIR reflectance (1084 nm)](/)

![Raman (1651 cm⁻¹)](/)

Fig. 1. Four single-band images extracted from hyperspectral data collected from a red snapper fillet.
Mean spectra of six fish species are plotted in Fig. 2. The VNIR reflectance spectra (Fig. 2a) show different patterns due to compositional variations of the fish fillets. The broad reflectance valley at 560 nm and two small valleys at 546 and 578 nm likely correspond with the absorption peaks of the heme pigments in the fish tissue, such as hemoglobin in the blood filled vessels and myoglobin in the muscle. Main spectral features of the SWIR reflectance (Fig. 2c) appear in the wavelength range of 900–1500 nm. Two major SWIR reflectance valleys were observed at 984 and 1208 nm, which are associated with the first O-H stretching overtone of water and the second C-H stretching overtone of fat, respectively. The fluorescence spectra (Fig. 2b) show distinctive differences, which could arise from different protein-protein interactions and collagen structures among the various fish species. Major Raman peaks of the fillet samples can be assignable to the lipid component in the fish, and their vibrational modes and chemical bonds are marked in Fig. 2d.

3.3 Fish Species Classifications

Fig. 4 summarizes fish species classification results by 24 machine learning classifiers using four types of spectral data in three different subsets. For the full spectra (Fig. 4a), the VNIR reflectance data achieved two perfect classifications using linear discriminant and subspace discriminant classifiers. Linear, quadratic, and cubic SVMs gave high accuracies (97.6–99.5%) for the VNIR reflectance, fluorescence, and Raman data. Naïve Bayes classifiers yielded the worst results (<80%) for all four types of the spectra. The accuracies using the PCA data (Fig. 4b) and the selected bands (Fig. 4c) exhibited some similar patterns with those using the full spectra. High accuracies (98.1–100%) were also obtained for the VNIR reflectance and fluorescence data using the linear, quadratic, and cubic SVMs. Overall, the VNIR reflectance and fluorescence data provided the best performance for classifying the fish species. The accuracies using the Raman data were slightly lower and the SWIR reflectance data generally gave the lowest accuracies.
Fig. 3. Mean spectra of red snapper fillets as received (AR) and after two freeze-thaw cycles (FT1 and FT2): (a) VNIR reflectance, (b) fluorescence, (c) SWIR reflectance, and (d) Raman. Selected bands for fish freshness classifications are marked on each spectrum of the AR red snapper fillet.

Fig. 4. Species classification accuracies for fillets from six types of fish by 24 machine learning classifiers using (a) full spectra, (b) first ten components of PCA, and (c) selected bands by sequential feature selection method.
3.4 Fish Freshness Classifications

The freshness classification results for the red snapper fillets are summarized in Fig. 5. For the full spectra (Fig. 5a), the highest classification accuracy is 99.9%, which was achieved by the subspace discriminant classifier on the SWIR reflectance data. High accuracies (98.1–99.0%) were also obtained for the VNIR reflectance data when the linear and quadratic discriminant classifiers and the quadratic and cubic SVMs were used. The first ten components of PCA for the VNIR reflectance spectra generally gave higher accuracies than the other three types of the spectra for most of the 24 classifiers (Fig. 5b), with the highest accuracy obtained by the cubic SVM at 97.4%. The accuracies using the selected bands from the VNIR reflectance and fluorescence spectra (Fig. 5c) were generally lower than those using the full spectra and the PCA data. The SWIR reflectance data outperformed the other three types of the data, with the highest accuracy obtained by the quadratic SVM at 95.3%. The performance of the fluorescence spectroscopy was moderate and the Raman data generally yielded the lowest accuracies (<80%). These results demonstrated that the VNIR and SWIR reflectance modes seem more suitable for the fish freshness classification than the fluorescence and Raman modes. Considering fish species and freshness classifications together, the VNIR reflectance spectroscopy technique coupled with selected machine learning classifiers demonstrated strong performance for both tasks.

Fig. 5. Freshness classification accuracies for as-received and frozen-thawed red snapper fillets by 24 machine learning classifiers using (a) full spectra, (b) first ten components of PCA, and (c) selected bands by sequential feature selection method.

4. CONCLUSIONS

This study presented multimode hyperspectral imaging techniques to inspect substitution and mislabeling for fish fillets. Four types of spectra extracted from hyperspectral images created sufficiently large datasets to train and validate machine learning classifiers for fish species and freshness classifications. Results from different combinations of machine learning classifier, spectral type, and data subset provided an intuitive way to compare their performances and identify the best combination. The highest classification accuracies were achieved using selected machine learning classifiers to differentiate the fish species and evaluate the fish freshness using full reflectance spectra in the visible and near-infrared region and the short-wave infrared region, respectively. The reduced spectral datasets by principal component analysis and sequential feature selection method generally yielded lower classification accuracies than the full datasets. The reflectance spectroscopy technique in visible and near-infrared region demonstrated its potential for simultaneous inspection of the fish species and freshness.
REFERENCES


