Non-destructive Visible/NIR Spectroscopy for Differentiation of Fresh and Frozen-thawed Fish

Musleh Uddin, Emiko Okazaki, Sandor Turza, Yamashita Yumiko, Munehiko Tanaka, and Yutaka Fukuda

ABSTRACT: Nondestructive visible/near-infrared (NIR) spectroscopy was evaluated to investigate whether fish has been frozen-thawed. Fresh or frozen-thawed red sea bream Pagrus major (n = 108) were scanned using a NIRSystems 6500 spectrophotometer equipped with a surface interactance fiber-optic accessory then discriminated by soft independent modeling of class analogy (SIMCA) and linear discriminant analysis (LDA) based on principal component analysis (PCA) scores. The major effect of freeze-thawing treatment involves a gross change in total reflectance after freezing and thawing; this arises from changes in light scatter presumably arising from alterations in the physical structure of at least the surface layer of fish. Untreated original absorbance spectra achieved much better (100%) classification accuracy for the prediction samples while the same figures for multiplicative scatter correction (MSC) treated spectra are considerably worse, indicating that scattering is the major information that makes classification work. No Incorrect type of classification at all and also there are no samples classified to both groups either. This faster technique has the potential to differentiate fresh and frozen-thawed fish and could be applied for online or at-line processing control.

Keywords: visible/NIR, fresh or frozen fish, LDA, PCA, SIMCA

Introduction

Given the perishable nature of fish, extension of its shelf life is a requirement of normal trading. However, frozen fish usually have a much lower market price than fresh fish; therefore, the substitution of frozen-thawed for fresh fish is a significant authenticity issue. According to the Food and Agricultural Organization (FAO) and the Japan Agriculture Standard (JAS) regulations, labeling should state that the fish has been frozen-thawed must not be re-frozen (FAO 1982; JAS 2000). Fresh fish is, indeed, understood as having been freshly caught or having been chilled and stored for a short period at normal refrigeration temperature before purchase or use. For storage over longer periods, freezing is normally used. However, although frozen storage is effective for protecting against microbiological deterioration of fish meat, its physicochemical and organoleptical properties suffer (Uddin and Okazaki 2004). The consumer perception of such fish is inferior to that of the fresh material and this is reflected in the price it realizes. In practice, a considerable number of frozen fish are thawed in fish shops, stored on ice, and sold as unfrozen fish without being labeled as such. For the benefit of the consumer and prevention of unfair competition in the trade of fishery products, correct labeling of frozen-thawed fish or fillets is desirable. Consequently, control of labeling is possible only if there is any rapid and reliable methods that allow food control authorities to distinguish between fresh and frozen-thawed fish or fillets.

To differentiate fresh and frozen-thawed fish, measurement of the electric properties of fish tissues (Sakaguchi and others 1989), visual inspection of the eye lens (Love 1956), judgments of the electric properties of fish tissues (Sakaguchi and others 1989), or fillets.

The method offers the possibility of measuring physical and chemical properties. Once calibrated, the NIR spectrometer is simple to operate (Buning-Pfaue 2003). It has been widely used in the food industry and is based on the electromagnetic absorption of organic compounds (Misra and others 2000; Uddin and others 2002; Blazquez and others 2004).

From the aforementioned reasons, a need exists for a method that is capable of differentiating between fresh and frozen-thawed fish or fillets. Ideally, any such method should be nondestructive and rapid as well as reliable. In our earlier report, dry extract spectroscopy by infrared reflection (DESIR) of fresh and frozen-thawed fish was performed on the extracted meat juices then discriminated (Uddin and Okazaki 2004). In the present study, nondestructive visible/NIR spectroscopy was applied to investigate whether fish has been frozen-thawed. Compared with DESIR, no extractions are needed and no wastes are produced in visible/NIR spectroscopy using a fiber-optic probe, which would be an eco-friendly instrumental technique. It has demonstrated that the potential for addressing some authenticity issue in foods which known to be a rapid instrumental technique (Ding and Xu 1999; Cozzolino and others 2002; Downey and others 2003; Arnalds and others 2004).
Visible-NIR for fresh/frozen-thawed fish . . .

Materials and Methods

Fish sample

One hundred eight live red sea bream, *Pagrus major*, were purchased from Kanazawa Prefecture, Kanazawa, Japan, and transported with sea-water to the Natl. Research Inst. of Fisheries Science, Yokohama, Japan. The fish used in this study were between 416 g to 1307 g (average 903 g) and fork length between 23.2 cm and 34.9 cm (average 28.5 cm). The fish were killed in ice-cold water shock in big plastic buckets, which allowed us to keep them injury-free within 30 min. Fish were divided into 2 equal groups and used for further evaluation. For fresh or unfrozen fish, 54 samples were used soon after being killed, whereas the 2nd lot of 54 fish was kept overnight at 5 °C. After 30 d of frozen storage, fish were removed and thawed for quality evaluation of fish samples. Excess moisture was removed from the fish immediately before scanning by means of a paper tissue. Before spectra were measured on the fish, a reference spectrum was obtained by measuring the reflected radiation from a 5-cm-dia white ceramic plate. Spectra were recorded at a wavelength range of 400 to 1100 nm at 2-nm intervals; above 1100, the fiber material became a significant absorber of near-infrared radiation. Operation of the spectrophotometer and the collection of spectra were performed using the VISION software packages (Version 3.20, NIRSystems, Md., U.S.A.). The spectra were stored in optical density units log (1/T), where T represents the percentage of energy transmitted.

NIR measurement

The fish samples were scanned using a NIRSystems 6500 spectrophotometer (Perstrop Analytical Inc., Silver Spring, Md., U.S.A.) equipped with a surface interactance fiber-optic accessory. Within a 4-cm square probe face, 7 quartz windows (1 × 20 mm) are fitted; windows are alternatively light exit (n = 4) and collection (n = 3) ports. The red sea bream were measured at a location just behind the dorsal fin, midway on the epaxial part. Earlier works (Downey 1996; Wold and Isaksson 1997) indicated this as the optimal area for NIR measurements for quality evaluation of fish samples. Excess surface moisture was removed from the fish immediately before scanning by means of a paper tissue. Before spectra were measured on the fish, a reference spectrum was obtained by measuring the reflected radiation from a 5-cm-dia white ceramic plate. Spectra were recorded at a wavelength range of 400 to 1100 nm at 2-nm intervals; above 1100, the fiber material became a significant absorber of near-infrared radiation. Operation of the spectrophotometer and the collection of spectra were performed using the VISION software packages (Version 3.20, NIRSystems, Md., U.S.A.). The spectra were stored in optical density units log (1/T), where T represents the percentage of energy transmitted.

Multivariate data analysis

Spectral data were analyzed with The Unscrambler software (Version 8.05, Camo, N.J., U.S.A.). Of the 108 samples in total, 54 were fresh and 54 were frozen-then-thawed and were divided into a modeling set and a prediction set. The modeling set contained 35 samples for the fresh and 35 for the frozen fish. Twenty-seven of those samples were picked as every odd-numbered sample in the order of recording, and the remaining 8 samples were selected randomly. Thus, 19 samples for both fresh and frozen fish were allocated to the prediction set. Sample spectra for both sets were treated in exactly the same way with 2nd derivative or multiplicative scatter correction (MSC), or with no treatment at all. MSC was originally developed to correct light scattering variations in reflectance spectroscopy and was 1st proposed by Martens and others (1983). MSC is a simple technique to estimate both offset a and scale b, using linear regression of spectral variables versus the average spectrum and correct multiplicative and additive scatter effects using the calculated a and b values.

\[ x_k = a + b\bar{x}_k + e_k \]

\[ x_{k,\text{corrected}} = (x_k - a)/b \]

where \( x_k \) is the spectral data at wavelength \( k \), and \( \bar{x}_k \) is the mean spectrum in the sample set.

There are many ways to explore data structures, recognize patterns, and classify samples according to some distance measure, as described in an easy-to-understand manner book by Esbensen (2000). For our purposes, we used the classification method called Soft Independent Modeling of Class Analogy (SIMCA) as defined by Wold (1976) and linear discriminant analysis (LDA) using PCA (principal components analysis) scores (McLachlan 1992). The former method is based on disjoint PCA models; for each group, an independent PCA model is constructed which is then used to classify new, unknown samples. Details of PCA can be found in Martens and Martens (2001). The later one, in our case, uses the so-called scores values of PCA results as input variables to the LDA. By performing PCA 1st, we reduce the number of variables and make them independent; by doing so, only a small fraction of information is lost. This is important because LDA requires the number of samples to be considerably higher than the number of variables to have a statistically meaningful classification.

Results and Discussion

For a classification to be successful, 2 things are needed. First, samples belonging to the same group should be as similar as possible, and second, the number of samples used in the modeling set should be considerably higher than the number of variables. This is important because LDA requires the number of samples to be considerably higher than the number of variables to have a statistically meaningful classification.

Figure 1—(a) Average original absorbance spectra of fresh (black line) and frozen (red line) red sea bream (RSB) sample spectra in the 900- to 1098-nm wavelength range and (b) 2nd derivative spectra of some fresh and frozen RSB samples in the same wavelength range.
Water absorbs strongly in specific wavelengths, which is expected and usually exhibits a broad band because of H-bonding interactions with itself and with other components in the meat (Figure 1a). In visible-NIR spectroscopy, the regions from 740 to 760 nm and 960 to 980 nm are related to O–H bond of the water in the sample (Murray and Williams 1990; Shenk and others 1992; Osborne and others 1993; Byrne and others 1998). In this figure, only the 900- to 1098-nm region is displayed since we cut off the visible region from 400 to 898 nm and excluded it from any calculations. This range was further optimized, and we found that the 900- to 1098-nm spectral interval with original absorbance spectra provided the best separation, as seen in Figure 2. We have tested other spectral intervals also, but the distance of the 2 groups was the biggest and the spread of data points within each group the smallest when the 900- to 1098-nm interval with original absorbance spectra was used. The PCA score plot clearly shows us that the fresh (right side) and the frozen-thawed (left side) samples are well separated (Figure 2). For this model, only 1 factor was enough to separate the 2 groups. As can also be noticed, the frozen-thawed samples have a more compact structure, that is, data points are closer to each other, whereas in the fresh samples, the group is not that well defined (larger spread of data points). Similar separation was also observed in DESIR analysis of fresh and frozen-thawed fish was performed on the meat juices (Uddin and Okazaki 2004). Using the results of this exploratory stage for all spectral treatments applied, 2 independent PCA models (900- to 1098-nm wavelength range with original absorbance spectra) were generated with the modeling sets and then used to build SIMCA models. There are several powerful advantages of the SIMCA approach compared with methods such as cluster analysis. First, SIMCA is not restricted to situations in which the number of objects is significantly larger than the number of variables as is invariably the case with classical statistical techniques. Not so with the present bilinear methods, which are stable with respect to any significant imbalance in the ratio objects/variables, be it either (very) many object with respect to variables, or vice versa. Because of the score-loading outer product nature of bilinear models, the entire data structure in a particular data matrix will be modeled well, even in a case in which 1 dimension of the data matrix is (very) much smaller than the other, within reasonable limits of course. Another advantage is that all the pertinent results can be displayed graphically with exceptional insight regarding the specific data structure behind the modeled patterns (Esbensen 2000).

SIMCA models were applied to the prediction set and results of the prediction can be best visualized by plotting the sample-to-model distances for all samples as shown in Figure 3 and 4. These plots are called Coomans plots, which show orthogonal (transverse) distances from all new objects (samples) to 2 selected models (classes) at the same time. In Figure 3, the 2 groups are well defined and separated. All prediction samples are much closer to the group that they should belong to. However, not every sample is within membership limits for both the modeling and the prediction samples. As can be seen, some samples are located in the upper right quadrant, indicating that they belong to none of the defined models. No sample is in the lower left quadrant, meaning that no sample was classified to both groups simultaneously. The upper left and lower right quadrants define samples that belong to 1 group. In Figure 4, however, where sample spectra were subjected to MSC transformation, modeling and classification seem much more uncertain. The 2 groups are very close; in fact, they almost overlap even at the modeling stage. This means that the MSC transformation removed information, that is, scattering, on which the previous model is based; therefore, models are not that far apart. However, the units in Figure 4 are by 1 to 2 magnitudes smaller compared with those of Figure 3, which explains why samples of the 2 groups are closer and more scattered. The distance, which is by a magnitude smaller compared with previous
Visible-NIR for fresh/frozen-thawed fish . . .

model, between groups is important in terms of reliability or robustness of the model. This does not necessarily mean that classification accuracy is worse, as indicated in Table 1, in which results are summarized for models with original absorbance and MSC transformed spectra. Looking at the figures, we can see that they are the same. The same proportion of samples is classified correctly, to none and to both groups, meaning the classification accuracy is the same for both models; however, the model using original absorbance spectra has higher reliability. This is also an important model feature because the model is more stable against random errors or interferences from any source.

As regards LDA, the results are much more clear-cut as seen in Table 2. To perform modeling and classification at the same wavelength range (900 to 1098 nm), spectral transformation and prediction samples were used for LDA analysis as well. It is clear from the table that the model using original absorbance spectra achieved much better (100%) classification accuracy for the prediction samples. The same figures for MSC-treated spectra are considerably worse, indicating again that scattering is the information that makes classification work. We think that for fresh fish, the cellular structure is intact. When light enters the fresh fish, cells not only absorb the light but change its direction until the light reaches the next cell. This process may continue until all light is absorbed or until the light emerges at the other side of the sample. This multiple change in the direction of light is called scattering, which increases the distance the light travels from the entry point to the exit point of the sample. This increase results in increased absorbance as seen in Figure 1a. On the other hand, when freezing and thawing is done, the cell membranes get damaged leaking the intracellular contents into the extracellular space. Thus, there is a much smaller number of cells that can scatter light as it travels through the sample, reducing the distance the light has to cover. As a result, light interacts with a smaller number of molecules, which in turn results in a decrease of absorbance (Figure 1a). It is interesting to note that frozen-thawed samples were a little better classified than fresh ones, but as the number of samples is not so great, making conclusions is too farfetched. This method maximizes the ratio of between-class variance to the within-class variance in any particular data set thereby guaranteeing maximal separability.

Several techniques were proposed to differentiate fresh and frozen-thawed fish (Yoshioka 1983; Vincenzo and others 1985; Uneo and others 1988; Rehbein and Cakli 2000); however, those are either empirical or time-consuming. Visible/NIR spectroscopy could minimize those aforementioned limitations. In a preliminary study, we have applied this technique to identify fresh and frozen-thawed horse mackerel (Trachurus japonicus) with a large number of sam-

---

**Table 1**—Discrimination results between fresh and frozen-thawed prediction red sea bream samples using the soft independent modeling of class analogy (SIMCA) method

<table>
<thead>
<tr>
<th>Spectral transformation</th>
<th>Sample</th>
<th>Kind of correctly</th>
<th>Classified to none</th>
<th>Classified to both</th>
<th>Nr of PCs</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>Fresh</td>
<td>63</td>
<td>37</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Frozen</td>
<td>84</td>
<td>16</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>73</td>
<td>27</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>MSC</td>
<td>Fresh</td>
<td>63</td>
<td>37</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Frozen</td>
<td>84</td>
<td>16</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>73</td>
<td>27</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

*aProportion (%) of samples that were classified to the correct model at 5% significance level.
*bProportion of samples which were classified to none of the models at 5% significance level.
*cProportion of samples which were classified to both models at 5% significance level.
*dNumber of principal components (PCs) that were used for making class model.
*eProportion of fresh and frozen samples combined.
*fMultiplicative scatter correction.

**Table 2**—Discrimination results between fresh and frozen-thawed prediction red sea bream using linear discriminant analysis (LDA) with principal component analysis (PCA) scores as input variables

<table>
<thead>
<tr>
<th>Spectral transformation</th>
<th>Nr correct in groups</th>
<th>Group proportion correct</th>
<th>Overall proportion correct</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fresh</td>
<td>Frozen</td>
<td>Fresh</td>
</tr>
<tr>
<td>None</td>
<td>19</td>
<td>19</td>
<td>100%</td>
</tr>
<tr>
<td>MSC</td>
<td>15</td>
<td>16</td>
<td>79%</td>
</tr>
</tbody>
</table>

*aThe number of correctly classified samples out of 19 prediction samples for the fresh and frozen-thawed red sea bream groups, respectively.
*bMultiplicative scatter correction.
ple sets (n = 162) by which other groups were identified with similar accuracy. Better accuracy and reliability are possible to achieve using a custom-made fiber or improved sample presentation. Nevertheless, results are promising that a fast measurement method can be developed to detect fraud such as when frozen-thawed fish are sold as fresh.

Conclusions

The applicability of visible/NIR technique has been successfully demonstrated to differentiate between fresh and frozen-thawed red sea bream, although we also pointed out places where improvements should be made and are possible. The technique uses the fact that fish muscle absorbs and reflects light in different ways during storage and thawing. By spectroscopically measuring raw materials using known characteristics, models can be developed that can again be used to estimate the characteristics for unknown specimens. Once we have the models, fresh and frozen-thawed fish could be differentiated in seconds.

Acknowledgments

The authors would like to acknowledge Dr. Nakamura Kouji (JIRCAS) for his support and technical assistance.

References


