Identification of Goose Down and Duck Down using Infrared Spectroscopy and Multivariate Analysis

Miyuki Funahashi*1,2, Norio Yoshimura*1, and Masao Takayanagi*1

*1 Tokyo University of Agriculture and Technology, United Graduate School of Agricultural Science, 3-5-8, Saiwai-cho, Fuchu, Tokyo, 183-8509, Japan
*2 Tokyo Laboratory, Nissenken Quality Evaluation Center, 2-16-11, Kurame, Taito, Tokyo 111-0051, Japan

Abstract: Down feathers, which are used as a filling material for duvets and winter clothes because of their excellent heat retention, are obtained mainly from domestic geese and ducks, but some are obtained from wild eider ducks. It is important to determine from which bird down was taken because the quality and price of down vary greatly depending on the bird species. In this study, infrared spectroscopy instead of the conventional microscopic observation was used to identify the bird species from which down was obtained. Goose down and duck down could be accurately identified by selecting an appropriate wavenumber region of the infrared (IR) absorption spectrum obtained using the attenuated total reflection (ATR) method and the partial least squares discriminant analysis (PLS-DA) method. Score plots based on principal component analysis (PCA) were found to be effective for identifying eider duck down and non-eider duck down.

(Received 14 August, 2021; Accepted 3 October, 2021)

1. Introduction

Feathers are epithelial structures composed of keratin proteins. Bird feathers can be roughly divided into two types, namely feathers and down, depending on their shape. Feathers have a hard quill, a quill shaft, and barbs, whereas down is composed of a small quill point and soft barbs that are derived from the tip of a quill point. Down does not have a firm shaft and is fluffier than feathers. Down has excellent heat retention and thus is used as a filling material for duvets and winter clothes.

Down is found on the chest of waterfowl. Land birds do not have down. Down used as the filling for duvets and clothes is mostly from domestic geese and ducks. Generally, geese have larger down feathers than those of ducks. Therefore, goose down is considered to be of higher quality and tends to be more expensive. Moreover, down from certain bird species is more expensive. The down of eider ducks (*Somateria mollissima*), which live in polar regions such as Iceland, has a particularly high value. Wild eider hens pull the down from their chests and spread it in their nests. After the birds leave, the down is collected from the nests for human use. Therefore, the amount of down that can be obtained is highly limited compared to that obtained from domestic waterfowl. In addition, eider duck down, which has hook-shaped branches, is strongly entangled, and exhibits high heat retention. Eider duck down is considered to be the highest-quality down because of its rarity and high performance.

Because the performance and price of down depend on the bird species and breed, the type of down used in duvets and clothes is extremely important for manufacturers, dealers, and consumers. Traceability has not been fully established for down. Cases of suppliers disguising the bird species and origin have been often reported. Therefore, it is important to independently identify the bird species and production areas of down. It is difficult to differentiate down from different bird species whose down has almost the same appearance and chemical composition. Identification is mainly performed by observing the morphological differences of down among bird species using a microscope. Down has barbs (fibrous materials) that grow from the quill point. From the barbs, smaller barbs and barbules...
emerge. For waterfowl, there are nodes on the barbule, whose shape and density differ depending on bird species [1‒4]. The subtle differences in these nodes observed through a microscope are used to identify down from different bird species. It takes a long time for testers to acquire the necessary skills for this subjective method. The identification results may depend on the tester. Furthermore, for natural materials such as down, individual differences in samples make identification even more difficult. Eider duck down is easy to distinguish from goose down and duck down because of its characteristic color and entanglement strength. However, the problem of individual differences in samples remains. A more objective and reliable identification method for down is thus strongly needed.

DNA analysis [5‒9] and isotope analysis of the elements that make up amino acids [10] are usually applied to determine the bird species and origins of down. However, these analysis methods are destructive. It is difficult to apply them in practice because about 50 to 100 down samples used in a product are inspected to determine the bird species and mixing ratios.

Spectroscopy has been applied as a non-destructive and simple identification method. Infrared (IR) spectroscopy and near-infrared (NIR) spectroscopy are widely used for the non-destructive material identification and quality evaluation of agricultural products, foods, textile products and other daily necessities [11‒13]. Zhou et al. [14] reported that identification of goose down and duck down can be performed using NIR spectroscopy and multivariate analysis. However, the reported models for the identification seem not reliable enough due to insufficient consideration of their construction and validation procedures.

In this study, IR spectroscopy was applied to identify goose down and duck down. In IR spectra, much more bands that can be assigned to specific molecular vibrations of samples are observed than in NIR spectra. The assigned bands are helpful for the identification. For example, it is well known that higher-order structures of proteins can be analyzed using IR spectroscopy [15‒17]. Therefore, it is expected that important information on the distinction of down with different properties can be obtained. Goose down and duck down were identified by analyzing the obtained IR spectra using score plots based on principal component analysis (PCA) and the partial least squares discriminant analysis (PLS-DA) method [18,19]. Furthermore, eider duck down and the down of other duck varieties (non-eider duck down) were identified using principal component score plots. These methods are expected to provide objective and less variable identification results because they are based on the molecular vibrations of sample components.

PCA score plots are widely used for discriminant analysis based on near-infrared and infrared spectra [11‒13]. It is also often used to identify fibers and fabrics [20‒22], Zhou et al. [14] used PCA to identify goose down and duck down. PLS-DA has become also widely applied for near-infrared and infrared spectra as a simple and effective discriminant analysis method. However, there have been few applications related to fibers yet [23‒25]. PLS-DA has not been used to identify the type of down yet.

2. Experimental

2.1 Down samples

Samples of goose down and duck down, whose types and bird breeds are known, received from Toyo Feather Co., Ltd. were used. Commercially available down was also collected. The bird species were identified using the conventional method before the down samples were subjected to analysis.

For the identification of goose down and duck down, 80 samples of each type were used. Of these 80 samples, 60 samples were used to construct discriminant models and 20 samples were used to validate the constructed models. Not all breeds of goose and duck from which down was collected were known. Therefore, the exact breeds were not identified.

For identification of eider duck down and down from other ducks (Bulgarian mulard duck, American Peking duck, and two Chinese ducks whose exact breeds were unknown), 50 samples of each were used. Of these 50 samples, 30 samples were used to construct the discriminant models and 20 samples were used to validate the constructed models.

2.2 Spectral measurement

IR absorption spectra (measurement range: 4000‒600 cm⁻¹, resolution: 4 cm⁻¹, accumulation: 32) were obtained using the attenuated total reflection (ATR) method (Jasco, ATR PRO ONE with a diamond prism) with a Fourier transform (FT)-IR spectrophotometer (Jasco, FT/IR-6800) in a laboratory with a constant
temperature of 20 °C and a constant humidity of 65 %.
The samples were left in this environment for at least 4 hours before their spectral measurements. The composition of a sample may vary depending on the measurement position. Therefore, all measurements were performed at the barbs, and spectra obtained at three different positions for given sample were averaged and used for analyses.

2.3 Analyses

For the obtained spectra, ATR correction, second-derivative calculation using the Savitzky-Golay (SG) method, and standard normal variate (SNV) conversion were performed as preprocessing. In the SG method, quartic function fitting was employed with a window width of 61 points. The order of the fitting function and the window width were optimized by trial and error. Derivative-calculation processing can remove the skewed baseline of a spectrum. Moreover, because derivative calculation emphasizes the intensity change of a spectrum, it is effective for capturing slight differences in the spectrum. After a wavenumber region suitable for analysis was extracted from the second-derivative spectra, SNV conversion was performed. In the SNV conversion, a spectrum was normalized by (1) subtracting the average of the observed data in the target wavenumber region from the observed value at each wavenumber and then (2) dividing by the standard deviation of the observed data in the region. The normalized spectra were mean-centered by subtracting the average spectrum, i.e., the centers of the intensity variations were aligned. After the data structure was simplified by this preprocessing, PCA and PLS-DA were performed to effectively capture and discriminate the characteristics of the spectra.

In PCA, explanatory variables are decomposed into loadings and scores based on their variability to summarize the multivariate data. Here, spectra of down samples were explanatory variables. A spectrum is decomposed into loadings, which are appropriate components to explain the fluctuation of the spectra of samples, and the scores corresponding to the coefficients for loadings are obtained. Using these scores, each spectrum can be represented as a score plot corresponding to one point in a multidimensional (usually two-dimensional) space. Because the distance on the score plot corresponds to the similarity of the spectra, samples whose plots are close to each other are similar and those whose plots are far away from each other are significantly different. From the loadings and the distance of the score plots, the difference between samples can be evaluated.

PCA is an unsupervised discriminant model construction method, whereas PLS-DA is a supervised method [18, 19]. In PLS-DA, a discriminant model was constructed with the spectral data as explanatory variables and the types of samples (goose down or duck down) as the objective variables; values of 1 and -1 were assigned to the spectra of goose down and duck down. In this discrimination method, the PLS regression method is used to construct a model that maximizes the covariance between the explanatory variables and the objective variables. Because the model is constructed so that the two groups to be discriminated can be best distinguished, it is possible to construct discriminant models that are more effective than those obtained using PCA.

Preprocessing and analyses (construction and validation of discriminant models) were carried out using Microsoft Office Excel 2016 and programs developed by the authors in FreeMat 4.0 (http://freemat.sourceforge.net/).

3. Results and Discussion

3.1 Identification of goose down and duck down

Figs. 1 (a) and 1 (b) shows the IR absorption spectra of goose down and duck down, respectively. Each spectrum is the average spectrum of the 60 samples used for model construction. The baselines do not coincide with zero because the averages were calculated for the spectra after SNV conversion. Goose down and duck down, both of which are mainly composed of protein, have similar spectra and are thus difficult to distinguish visually.

Figs. 2 (a) and 2 (b) show the second-derivative spectra of goose down and duck down, respectively, in the fingerprint region; the region that is expected to show the largest spectral differences. The second-derivative calculation with smoothing removed the baseline fluctuation and emphasized some fine structures that appeared in the original spectra. Nevertheless, the difference between the second-derivative spectra of goose down and duck down was still too small to visually distinguish the type of down. However, the difference spectrum (Fig. 2(c)) obtained by subtracting the goose down spectrum (Fig. 2 (a)) from the duck down spectrum (Fig. 2 (b)) showed...
distinct peaks.

Intense peaks (positive or negative) were observed at approximately 1610, 1500, and 1470 cm\(^{-1}\). It should be noted that the peak frequencies in Fig. 2 (c) do not coincide with the frequencies of the observed spectra, because the spectrum in Fig. 2 (c) is a difference spectrum. The band at 1470 cm\(^{-1}\) can be assigned to the bending vibration of a methyl or methine group and those at 1610 and 1500 cm\(^{-1}\) are due to amide I (1700‒1600 cm\(^{-1}\)) and amide II (1575‒1480 cm\(^{-1}\)), respectively. Amide I and amide II are the strongest and most important absorption bands related to the vibrations of a protein backbone such as the C=O stretching of peptide bonds, N-H bending vibration, and C-N stretching vibration. In particular, the amide I band reflects the secondary structure of the protein [15‒17]. There might be differences in the structures of proteins of goose down and duck down, resulting in the distinct peaks in the difference spectrum.

Any distinct difference between the spectra of goose down and duck down may make it possible to identify these types of down using PCA. Identification using PCA was thus tried. The spectral region of 1800–1400 cm\(^{-1}\), where intense signals due to amide I and amide II were observed in the second-derivative spectra was used for PCA. Some intense signals were observed in other regions. The peaks due to the CH stretching vibrations observed in the 3000–2800 cm\(^{-1}\) region show less clear dependences on molecular structure than those observed below 1800 cm\(^{-1}\). It is difficult to use the strong absorption above 3000 cm\(^{-1}\) attributed to NH stretching vibration because it overlaps with the absorption due to moisture. Therefore, the region above 2800 cm\(^{-1}\) was not used for PCA. Although a characteristic band, namely amide III band, is generally observed around 1240 cm\(^{-1}\) in the IR absorption spectra of proteins, no distinct signal appears at this wavenumber in the difference spectrum in Fig. 2 (c). Therefore, the wavenumber region below 1400 cm\(^{-1}\) was not used for PCA. PCA was additionally conducted using the
spectra in regions other than 1800–1400 cm\(^{-1}\); however, the results were not good as those presented below.

The results of PCA are shown as a score plot in Fig. 3, where the scores of the third principal component (PC3) are plotted against those of PC2. Along the PC2 axis, the plots of goose down and duck down tend to be in positive and negative regions, respectively. However, both goose-down and duck-down plots appear in the range of -0.5 to 0.5 of PC2. Therefore, this score plot cannot be used to identify the type of down.

No information that could be used for distinguishing between goose down and duck down was contained in the score along PC3. The variation of spectra captured by PC3 was fine fluctuations without wavenumber dependence and seemed mostly due to noise. The score along PC1 also did not contain useful information for the identification. The loading spectrum of PC1 mainly captured the variation of amide I band, which seems to depend on the individuality of each sample rather than the bird species. PCA, which is an unsupervised discriminant learning technique, did not construct a model that loading could sufficiently identify goose down and duck down. Although the score along PC2, which captured mainly the variation of amide II band, contains information on the difference between goose down and duck down, reliable identification of the type of down could not be done; the extent of spectral variation from sample to sample for a given bird type was the same as that between the two bird types, making accurate identification of down difficult.

PLS-DA, a supervised discriminant modeling method, was next employed to construct a discriminant model. The second-derivative spectra in the 1800–600 cm\(^{-1}\) region were used for the model construction. As done for PCA, discriminant models were also constructed using spectra in narrower wavenumber regions. The results were not as good as those obtained with the spectra in the 1800–600 cm\(^{-1}\) region. In the PLS-DA scheme, values of 1 and -1 were assigned to the spectra of goose down and duck down, respectively. The model for predicting the value from a spectrum was constructed using the PLS method. If the predicted value was positive (negative), the sample corresponding to the spectrum was determined to be goose (duck) down.

The results of PLS-DA are shown in Fig. 4. The number of latent variants for PLS was determined to be 7 by trial and error. With this number of latent variants, the best result, shown in Fig. 4, was obtained. The discriminant model was constructed using the same samples of goose down and duck down as those used in Fig. 3. The number of latent variants for the discriminant model was 7.
used for PCA (60 samples for each type of down). The constructed model was tested with 20 samples for external validation. In Fig. 4, the horizontal axis denotes the sample number, and the vertical axis denotes the predicted value for each sample. Both the samples for model construction (Fig. 4 (a)) and the samples for model validation (Fig. 4 (b)) were correctly identified; all predicted values for goose down were positive and those for duck down were negative.

The plots of test samples scattered more than those of model samples. The samples used for the analysis were goose down and duck down of various breeds from various regions of the world. There may have been differences in how the samples were cleaned and processed. In addition to differences in bird species, differences in the origin and processing might affect identification results.

Because the number of samples used for the model construction was relatively small, overfitting might have occurred. By increasing the number of model samples, more robust discriminant models that can be applied to a wider variety of samples could be constructed. Alternatively, constructing a discriminant model for each breed, each place of origin, each processing method, and so on would improve identification performance. However, discriminant models constructed in this way would be less versatile; they can only be used for specific targets.

3.2 Identification of eider duck down and non-eider duck down

Identification of down from eider ducks and that from other types of ducks was attempted. Here, goose down was excluded from the identification. IR absorption spectra of eider duck down and non-eider duck down are, respectively, shown in Figs. 5(a) and 5(b). Each spectrum is the average spectrum of the 30 samples used to construct a discriminant model. Because the average spectra were obtained after SNV conversion, the baseline do not coincide with zero. No clear visual difference was observed between spectra of eider duck down and non-eider duck down.

Fig. 6 (a) shows the second-derivative spectrum of eider duck down overlaid on that of non-eider duck down in the 1800‒600 cm\(^{-1}\) region after SNV conversion. Each spectrum is the average spectrum of 30 samples. By normalizing and overlaying spectra, some differences were clearly observed for spectra of eider duck down and non-eider duck down. The differences around 1400 and 1000 cm\(^{-1}\) were especially distinct. There was a particularly large difference in

![Fig. 5](image1)

![Fig. 6](image2)

**Fig. 5** IR absorption spectra of (a) eider duck down and (b) non-eider duck down. Each spectrum is the average spectrum of the 30 samples for model construction.

**Fig. 6** (a) Spectra of eider duck down and non-eider duck down in the 1800‒600 cm\(^{-1}\) region overlaid for comparison. (b) Magnified view of the 1150‒850 cm\(^{-1}\) region of spectra in (a).
the 1060‒950 cm⁻¹ region as shown in Fig. 6 (b), where the spectra are shown with expanded scales. However, the difference is so small that it is difficult to identify the bird species visually from the whole spectrum. Therefore, identification using a PCA score plot was examined.

PCA was performed for the spectra in the 1040‒950 cm⁻¹ region after mean-centering. The results are shown in Fig. 7 as a score plot. The score of the first PC was negative for all samples of eider duck down and positive for all samples of non-eider duck down; the scores for the two types of down are in the negative and positive regions on the horizontal axis, respectively. The samples for both model construction and model validation were separated. Therefore, a PCA score plot was found to be applicable for the identification of eider duck down and non-eider duck down.

The 1040‒950 cm⁻¹ region used for the analysis was determined by trial and error. As shown in Fig. 6, the spectra of eider duck down and non-eider duck down differed significantly in the 1060‒950 cm⁻¹ region. However, PCA that included the 1060‒1040 cm⁻¹ region yielded the worse score plots; the plots for the different types of down were not sufficiently separated for reliable identification. Moreover, identification of the two types of down based on the PCA of the spectra around 1400 cm⁻¹, where distinct differences were observed in Fig. 6, did not work well either.

Fig. 7 PCA score plot for eider duck down and non-eider duck down. PCA was applied to spectra in the 1040–950 cm⁻¹ region. Results of model construction samples and test samples are overlaid.

Fig. 8 (a) shows the difference spectrum in the 1040-950 cm⁻¹ region obtained by subtracting the second-derivative spectrum of eider duck down from that of non-eider duck down. (b) Loading spectrum of PC1 obtained using the PCA of spectra in the same region.

Goose down and duck down could not be identified using PCA. In contrast, eider duck down and non-eider duck down could be identified using
PCA. Eider duck down is known to be different from the down of other ducks and geese in appearance and performance; eider duck down differs from other types of down in terms of material components and thus its IR absorption spectrum is unique. Specifically, eider duck down and non-eider duck down might be different in terms of the amino acids that make up the protein of down, which allows for their identification. Eider ducks are wild birds, and the other ducks are domestic. Thus, they have different diets, which may also cause differences in down amino acid composition.

The difference in processing for eider duck down and non-eider duck down might be another cause of the difference in their spectra. Because white down is preferred in general, ordinary down is often bleached. Various treatments such as antibacterial and fire-retardant processing are also often employed. In contrast, eider duck down, which is brownish in color, is considered valuable for its color and as a natural luxury product. Eider duck down is rarely subjected to any processing or treatment other than washing. Therefore, the difference in the spectra of eider duck down and non-eider duck down might be caused by the presence or absence of processing. Although the differences in the types and degrees of processing might affect spectra, it is difficult to clarify the dependence of spectra on processing.

4. Conclusions

With the methods newly developed in the present study, goose down from duck down, and eider duck down from non-eider duck down could be successfully discriminated. A PCA score plot could not distinguish these two types of down. Their identification was found to be possible by applying PLS-DA to the second-derivative spectra in the 1800-600 cm⁻¹ region. Eider duck down and non-eider duck down could be distinguished using a score plot based on the PCA of the second-derivative IR absorption spectra in the 1040-950 cm⁻¹ region, where distinct differences were observed in the spectra. The plots of eider duck down and non-eider duck down were clearly separated along the PC1 axis. Of note, the difference between eider duck down and non-eider duck down is larger and more distinct than that between goose down and duck down.

Although the techniques developed in the present study is an indirect analysis method, if discrimination models are properly constructed, they enable discrimination of type of down with the same accuracy as the conventional DNA discrimination technique that is a time-consuming and destructive analysis method. The technique presented here is applicable to practical analyses that identify large numbers of samples quickly and non-destructively. Therefore, the DNA discrimination technique and analysis based on spectroscopy should be used properly according to the purpose.

The rapidly evolving data science techniques together with reliable samples make accurate discrimination possible, although the reason why the type of down can be clearly identified by infrared spectroscopy has not yet been clarified sufficiently. It is necessary to examine the principle of discrimination, however, to improve the reliability of discrimination and to enable more detailed discrimination.

A practical identification method should distinguish down of various grades obtained from birds of different origins, breeds, and maturities. By increasing the number and variety of samples for model construction, more detailed, precise, and robust discriminant models can be obtained. The application of NIR spectroscopy instead of IR spectroscopy is another issue because NIR spectroscopy is expected to enable more rapid, simpler, and nondestructive identification.

Acknowledgment

The authors would like to thank Toyo Feather Industry, Co., Ltd. for providing the down samples.

References

1. M. G. Day, *Journal of Zoology*, 148, 201 (1966).
2. T. G. Brom, *Bijdragen tot de Dierkunde*, 56, 181 (1986).
3. J. Robertson, C. Harkin, and J Govan, *Journal of the Forensic Science Society*, 24, 85 (1984).
4. D. Yildiz, E. U. Bozkur and S. H. Akturks, *Journal of Animal and Veterinary Advances*, 8, 2650 (2009).
5. P. D. N. Hebert, M. Y. Stoeckle, T. S. Zemlak and C. M. Francis, *PLoS Biology*, 2, e312 (2004).
6. F. E. Hogan, R. Cooke, C. P. Burridge and J. A. Norman, *Molecular Ecology Resources*, 8, 561 (2008).
7. C. F Speller, G. P. Nicholas and D. Y Yang, *Investigative Genetics*, 2, 16 (2011).
8. S. Boonseub, G. Johnston and A. Linacre, *Journal
of Forensic Sciences, 57, 1574 (2012).
9. M. Natonek-Wiśniewska, P. Krzyścin and M. Bugno-Poniewierska, European Food Research and Technology, 245, 335 (2019).
10. K. A. Hobson, Oecologia, 120, 314 (1999).
11. Y. Ozaki, W. F. McClure and A. A. Christy, “Near-Infrared Spectroscopy in Food Science and Technology”, John Wiley & Sons (2006). ISBN: 978-0471672012
12. Y. Ozaki, C. Huck, S. Tsuchikawa and S. B. Engelsen, “Near-Infrared Spectroscopy: Theory, Spectral Analysis, Instrumentation, and Applications”, Springer (2021). ISBN: 978-9811586477
13. D.-W. Sun, “Infrared Spectroscopy for Food Quality Analysis and Control”, Academic Press (2009). ISBN: 978-0123741363
14. Y. Zhou, H. R. Xu and Y. B. Ying, Transactions of the ASABE, 51, 1493 (2008).
15. J. T. Pelton and L. R. McLean, Analytical Biochemistry, 277, 167 (2000).
16. A. Barth, Biochimica et Biophysica Acta, 1767, 1073 (2007).
17. M. Carbonaro and A. Nucara, Amino Acids, 38, 679 (2010).
18. M. Sjöström, S. Wold and B. Söderström, “Pattern Recognition in Practice”, 461, Elsevier (1989).
19. D. Ballabio and V. Consonni, Analytical Methods, 5, 3790 (2013).
20. P. Peets, I. Leitoa, J. Pelt and S. Vahur, Spectrochimica Acta Part A, 173, 175 (2017).
21. H. Yan and H. W. Siesler, Journal of Near Infrared Spectroscopy, 26, 311 (2018).
22. J. Zhou, L. Yu, Q. Ding and R. Wang, AUTEX Research Journal, 19, 201 (2019).
23. L.-P. Xin, X.-S. Chai, D. Barnes, C.-X. Chen and R.-Q. Chen, Journal of Near Infrared Spectroscopy, 22, 347 (2014).
24. J. Manheim, K. C. Doty, G. McLaughlin and I. G. Lednev, Applied spectroscopy, 70, 1109 (2016).
25. H. Chen, Z. Lin and C. Tan, Microchemical Journal, 144, 489 (2019).
26. S. Hilterhaus-Bong and H. Zahn, International Journal of Cosmetic Science, 9, 101 (1987).
27. S. Ogawa, K. Fijii, K. Kaneyama and K. Arai, Sen’i Gakkaishi, 64, 137 (2008).