

Technical University of Denmark



## Design of characteristics of optical filter set for prediction and visualization of fat content in raw beef cuts

**Kobayashi, Ken-ichi; Nishino, Ken; Dissing, Bjørn Skovlund; Mori, Masaaki; Toyota, Toshihiro; Nakauchi, Shigeki**

*Published in:*

Scandinavian Workshop on Imaging Food Quality 2011

*Publication date:*

2011

*Document Version*

Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

*Citation (APA):*

Kobayashi, K., Nishino, K., Dissing, B. S., Mori, M., Toyota, T., & Nakauchi, S. (2011). Design of characteristics of optical filter set for prediction and visualization of fat content in raw beef cuts. In Scandinavian Workshop on Imaging Food Quality 2011: Ystad, May 27, 2011 - Proceedings (pp. 23-28). Kgs. Lyngby, Denmark: Technical University of Denmark (DTU). (IMM-Technical Report--2011; No. 15).

## DTU Library

Technical Information Center of Denmark

---

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# Design of characteristics of optical filter set for prediction and visualization of fat content in raw beef cuts

Ken-ichi Kobayashi<sup>1</sup>, Ken Nishino<sup>1</sup>, Bjørn Skovlund Dissing<sup>2</sup>, Masaaki Mori<sup>3</sup>,  
Toshihiro Toyota<sup>1</sup>, and Shigeki Nakauchi<sup>1</sup>

<sup>1</sup> Department of Computer Sciences, Toyohashi University of Technology,  
1-1 Hibarigaoka, Tenpaku-cho, Toyohashi 441-8580, Japan  
{kobayashi09,nishino06,toyota}@vpac.cs.tut.ac.jp, nakauchi@tut.jp  
<http://www.vpac.cs.tut.ac.jp/>

<sup>2</sup> Department of Informatics and Mathematical Modelling,  
Technical University of Denmark, DK-2800 Lyngby, Denmark  
bdi@imm.dtu.dk

<sup>3</sup> Mie Prefecture Livestock Research Institute,  
1444-1 Ureshino-cho, Matsusaka 515-2324, Japan  
morim05@pref.mie.jp

**Abstract.** Quantification of specific compounds in a food-matrix is a very important factor for an overall quantification of the quality. Near infrared (NIR) hyperspectral imaging is a powerful technique to quantify specific constituents as well as its spatial distribution of the food-matrix. Hyperspectral imaging is however very expensive. We propose a way to design a simple measurement system consisting of a NIR sensitive monochrome camera together with a small set of optical filters to estimate and visualize a specific food compound without requiring a full hyperspectral device. Based on a set of hyperspectral measurements of beef and physical and chemical analysis of the fat within the beef, we propose a method to design a set of ideal Band Pass Filters (BPF), as small as possible while still maintaining predictability of fat content. The results show that 2 filters is a suitable amount of filters for prediction.

**Keywords:** NIR hyperspectral imaging, Optical filter, Beef, Content

## 1 Introduction

Traditionally quality evaluation of food has been done using visual inspection, chemical measurements or sensory testing. These methods are destructive, time-consuming and/or subjective, which calls for other quantification methods. Recently non-destructive methods for evaluation of food quality as well as visualization of the spatial distribution of constituents, by using (NIR) hyperspectral information have emerged [1]. Although hyperspectral image data is very versatile and contains much information, the measurement system is extremely expensive to install in a food factory. As another approach based on hyperspectral data, a method designing the optical transmission for the optical filter to modulate a

RGB camera’s spectral sensitivity and to highlight an object’s spectral features is proposed [2][3].

We propose a simple measurement system consisting of a NIR monochrome camera together with a small set of optical filters to estimate and visualize a specific food compound without use of a hyperspectral device. We use the fat content in raw beef as the target. Currently in Japan, the quality evaluation of beef carcasses is performed manually by a grader. In this grading, only visual inspection is used. Marbling, which is the amount and distribution of fat in the meat is the most important factor. Based on a set of hyperspectral measurements and physical and chemical analysis of fat within the beef, we propose a method to design a set of optical Filters, which accurately is able to predict the amount and distribution of this fat.

## 2 Materials and methods

### 2.1 Samples and measurements

A total of 126 meat samples consisting of various parts from three 25-month-old Japanese black cattle were collected. After about 60 days of ageing at  $0 - 5^{\circ}\text{C}$ , the beef carcasses were kept at  $-25^{\circ}\text{C}$  to maintain the fat properties during storage and transportation.

The fat content used for reference values was analyzed by physical and chemical method. Automated Soxhlet extraction equipment (Soxtherm416, Gerhardt, Germany) was used to obtain the fat percentage.

The hyperspectral measurements were performed by a NIR hyperspectral imaging system consisting of a linear image sensor (Spectral Camera SWIR; SPECIM Spectral Imaging Ltd, Finland), a linear slide table and halogen light sources (MH-M15, 150 W; Hataya Ltd, Japan). The hyperspectral camera works in the wavelength range of 970-2500 nm with a bandwidth of 6.3 nm at a resolution of 320 pixels (X-axis). We acquired samples at a resolution of  $380\mu\text{m}/\text{pixel}$  over a rectangular region of  $120 \times 130\text{mm}$  by moving the slide table. The exposure time was 3.0 ms.

### 2.2 Calculation of filter transmission intensity

The MATLAB 7.5 (R2007b; The MathWorks Inc., Natick, MA, USA) software package was used to analyze the hyperspectral image data. Optical filters were designed as ideal (rectangle-shaped) BPF and an assumption was made that a measurement using an optical filter consists of three images; a dark current image ( $I_{Dark}$ ), a white standard image ( $I_{White}$ ) and a sample image ( $I_{Sample}$ ). To remove the effects of dark current, spectral features produced by the light source, and flat field inhomogeneities, we use  $I_R$  as a model parameter calculated from measured images or hyperspectral images by

$$I_R = \frac{I_{Sample} - I_{Dark}}{I_{White} - I_{Dark}} = \int_{\lambda_{short}}^{\lambda_{long}} \frac{I_{Sample}(\lambda) - I_{Dark}}{I_{White}(\lambda) - I_{Dark}} d\lambda$$

Where  $\{I_{Dark}, I_{White}(\lambda), I_{Sample}(\lambda)\}$  is hyperspectral data,  $\lambda_{short}$  and  $\lambda_{long}$  are the wavelength edges of the BPF. When calculating  $I_R$ , the spectra  $\{I_{Sample}(\lambda), I_{White}(\lambda)\}$  were interpolated by cubic spline to 1,000 wavelength points between  $\lambda_{short}$  and  $\lambda_{long}$ .

### 2.3 Design of optical filter property

The filter properties were modeled by using the center wavelength ( $\lambda_c$ ) and the half-bandwidth ( $w_h$ ). We limited the minimum bandwidth of BPF to 50 nm, because too narrow BPFs cannot obtain enough luminance, which will cause reduction of the signal-to-noise ratio. The maximum bandwidth was limited to 1,000 nm, because very wide BPFs are hard to implement as a real optical filter. The wavelength range was also limited from 1,000 nm to 2,300 nm, because shorter/longer wavelength ranges of hyperspectral data could not provide sufficient intensity. With a spectral resolution of about 6.3 nm, meaning a total number of wavelength-points of 206. Even if the edges of BPFs are limited to these wavelength-points, every possible combination of  $n$  BPFs is  $\simeq 10^{4n}$ . Therefore “brute-force search” is not suitable for more than 2 or 3 filters in terms of searching time.

Multiple Linear Regression (MLR) was used to estimate parameters for linear models using filter transmission intensities as variables. To create and evaluate the estimation models, samples were divided into calibration and validation sets. Calibration samples were selected randomly ( $N_c = 84$ ) and remainder were used as validation samples ( $N_v = 42$ ). These sample sets were fixed to compare the results of different feature selection method.

Filter feature selections were done using leave-one-out cross validation, to minimize the root mean square error of cross-validation ( $RMSECV$ ) given by

$$RMSECV = \sqrt{\frac{\sum (y_c - \tilde{y}_c)^2}{N_c}}$$

where  $y_c$  is the reference value, and  $\tilde{y}_c$  is the predicted value of the calibration-set in cross validation. Furthermore the standard error of calibration ( $SEC$ ), the root mean square error of calibration ( $RMSE_c$ ) and the standard error of prediction ( $SEP$ ) were calculated as

$$SEC = \sqrt{\frac{\sum (y_c - \hat{y}_c)^2}{N_c - n - 1}}, \quad RMSE_c = \sqrt{\frac{\sum (y_c - \hat{y}_c)^2}{N_c}}, \quad SEP = \sqrt{\frac{\sum (y_v - \hat{y}_v)^2}{N_v}}$$

where  $\hat{y}_c$  is the predicted value of the calibration-set using the model,  $n$  is the number of filters,  $y_v$  is the reference value of the validation-set, and  $\hat{y}_v$  is the predicted value of the validation-set using the model.

We compared the following three feature selection methods.

**Stepwise random selection** In this method one needs to define the number of filters. A scoremap and a countmap is maintained for each filter which is used for deciding the final properties for the corresponding filter:

1. Generate the  $\{m, (m + 1), (m + 2), \dots, n\}$ -th filters randomly.
2. Calculate  $n$  filter outputs of each calibration sample.
3. Make a MLR model by using the calculated filter outputs and the corresponding reference values of the calibration-set.
4. Calculate the  $RMSECV$  for the calibration-set.
5. Add the  $RMSECV$  value to the  $n$  points in the scoremap. Also add 1 to the  $n$  points in the countmap.

4

6. Repeat step 1 to 5 sufficiently many times.

The scoremaps and the countmaps are made for each  $m$ -th filter individually, and have coordinate points corresponding to every filter feature  $(\lambda_c, w_h)$ .

7. Remove 0-count points both in the  $m$ -th countmap and the  $m$ -th scoremaps. Then make a mean scoremap by dividing the  $m$ -th scoremap by the  $m$ -th countmap.
8. Choose a point that minimizes the mean scoremap. This is then fixed as the  $m$ -th filter  $(\lambda_{c_m}, w_{h_m})$
9. Repeat steps 1 to 8  $n$  times with the  $\{1, 2, \dots, (m - 1)\}$ -th fixed filters.

Finally, local optimization (constrained nonlinear optimization) is performed using the result as a starting guess, minimizes

$$\min \left( \sqrt{\frac{\sum_i^{N_c} (y_i - \sum_j^n f(\lambda_{short}(j), \lambda_{long}(j), x_i))^2}{N_c}} \right), f(\lambda_{short}, \lambda_{long}, x) = \int_{\lambda_{short}}^{\lambda_{long}} x d\lambda$$

*s.t.*     $(1000 \leq \lambda_{short}, \lambda_{long} \leq 2300), \quad (50 \leq \lambda_{long} - \lambda_{short} \leq 1000)$

where  $x$  is reflectance (hyperspectral data) of a sample. In this method, a filter feature which has small average error in many trials of various feature combinations, is assumed to contain useful information for estimation.

**Forward selection method** Inspired by the classic variable selection technique also known as forward selection [4], the forward selection method calculates *RMSECV* for all possible filters and selects the filter  $(\lambda_{C_1}, w_{h_1})$  which minimizes this metric. This filter  $(\lambda_{C_1}, w_{h_1})$  is fixed at the found feature and the procedure is then repeated for next filter  $(\lambda_{C_2}, w_{h_2})$ . The  $m$ -th filter  $(\lambda_{C_m}, w_{h_m})$  is chosen with  $(m - 1)$  fixed filters. The wavelength of the BPF edges are discrete values.

At the  $m$ -th filter selection step, this method consider the combination of  $(m - 1)$  filters already fixed and  $m$ -th filter. Contrary to this the stepwise random selection method considers the combination of  $n$  filters at every step.

**Forward selection with local optimization** After each forward selection step, local optimization (same to stepwise random selection method) is performed using the result of the forward selection method as a starting guess.

### 3 Results and discussion

Table 1 shows the statistics of results for the calibration and validation samples found by the physical and chemical analyses. The samples have a rather wide distribution.

**Table 1.** Statistics for reference values in each data set

	Number	Mean[%]	SD[%]	Min[%]	Max[%]
Entire set	126	30.92	14.66	5.25	71.23
Calibration set	84	30.36	13.99	5.25	63.50
Validation set	42	32.04	16.03	8.98	71.23

Figure 1 shows an example of the result of stepwise random selection for  $n=6$  with the area-averaged reflectance spectra from all 126 samples. The number of

repetitions for each selection step is  $10^5$ . Figure 3 shows the mean scoremaps. The minimum point (the selected condition) is indicated by a magenta mark in each of the maps. Figure 2 shows an example of the plot of the evaluation values for  $n= 1$  to 10. In this result, local optimization decreases the *RMSECV*, and increases the *SEP*.

Figure 4 shows the result of the forward selection method, while figure 5 shows a plot of the evaluation values. Although the *RMSE<sub>c</sub>* continues to decrease at least until 10 filters, the *SEP* begin to increase after 6 filters, and the *RMSECV* is not improved after 3 filters. Also the 4-7th filter is exactly similar, it's impractical to implement a real optical filter individually.

Figure 6 shows the modifications achieved by local optimization after each forward selection step. It appears that there are few or no modifications. Figure 7 shows a plot of the evaluation values. Local optimization decreases the *RMSECV* slightly, however the *SEP* is increased. This might indicate that the local optimization causes overfitting.

In summary, local optimization decreases the *RMSECV*, however, it does not necessarily mean the model's accuracy improves. In this case, 2 filters might be enough for prediction.

In previous work [1], which uses hyperspectral data and PLS1 regression, the *SEP* is 4.81. In that study, spectral correction (multiplicative scatter correction) was used. Because the purpose is to implement a real optical filter, we use raw reflectance spectra without any spectral correction. We achieve a minimum *SEP* around 5.0. This could indicate that our results demonstrate sufficient accuracy.

In this study, similar *RMSECV* curves but different *SEP* curves are obtained. Accuracy has not improved even we increase the number of filters to more than 2. This might mean that the prediction of the fat content is rather easy, because it was high in fat.

In future work, we will apply these methods to the estimation of fatty acid content, which is more difficult to predict than fat content because the content is much lower. Also, we will apply imaging to visualize the food composition. To improve the accuracy, we will add a combination of filter output values to the estimation model, and perform selection of them.

## References

1. Kobayashi, K., Matsui, Y., Maebuchi, Y., Toyota, T., Nakauchi, S.: Near infrared spectroscopy and hyperspectral imaging for prediction and visualisation of fat and fatty acid content in intact raw beef cuts. *J. Near Infrared Spectrosc.*,18,301–315(2010)
2. Nishino, K., Nakamura, M., Matsumoto, M., Tanno, O., Nakauchi, S.: Optical Filter for Highlighting Spectral Features Part I: Design and Development of the Filter for Discrimination of Human Skin With and Without an Application of Cosmetic Foundation. *Optics Express*, 19, 6020–6030(2011)
3. Nishino, K., Nakamura, M., Matsumoto, M., Tanno, O., Nakauchi, S.: Optical Filter Highlighting Spectral Features Part II: Quantitative Measurements of Cosmetic Foundation and Assessment of their Spatial Distributions under Realistic Facial Conditions. *Optics Express*, 19, 6031–6041(2011)
4. Wilkinson, L., Dallal, G.E.: Tests of significance in forward selection regression with an F-to enter stopping rule. *Technometrics.*,23,377–380(1981)

6

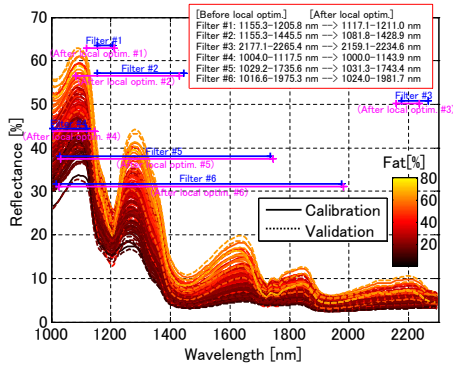


Fig. 1. An example of the result of stepwise random selection ( $n = 6$ )

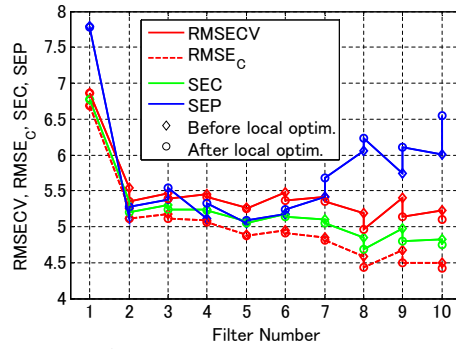


Fig. 2. An example of individual error measures (for  $n=1-10$ ) for stepwise random selection

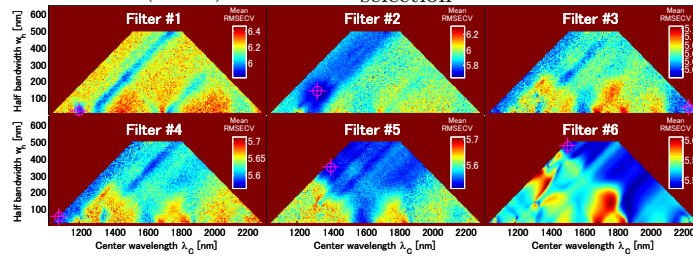


Fig. 3. An example of an evaluation map of stepwise random selection ( $n = 6$ )

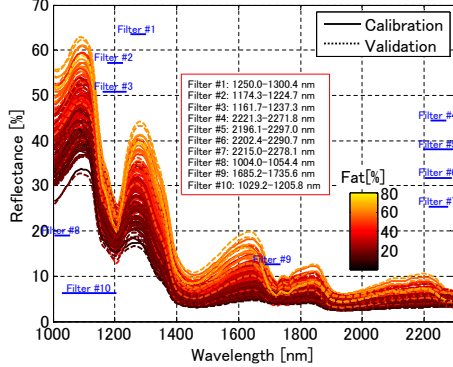


Fig. 4. The result of forward selection ( $n=10$ )

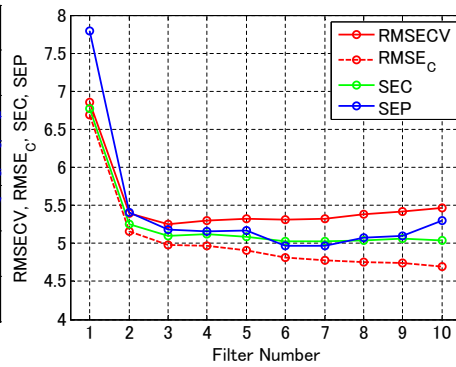


Fig. 5. Error measures for forward selection

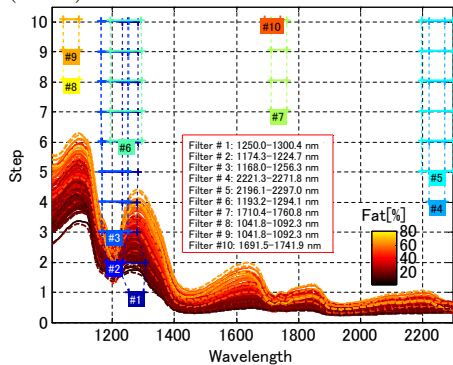


Fig. 6. Modifications of the BPFs found by local optimization after each forward selection step ( $n=1-10$ )

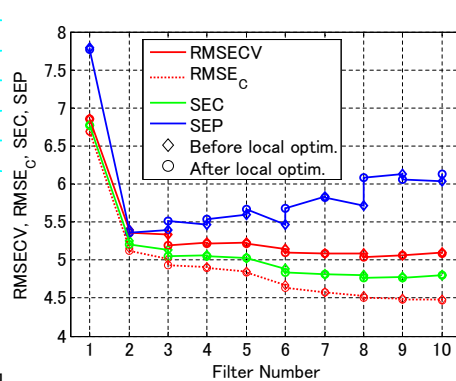


Fig. 7. Error measures for forward selection with local optimization