

Optimal Vision System Design for Characterization of Apples Using US/VIS/NIR Spectroscopy Data

Sara Sharifzadeh¹, Line H. Clemmensen³, Bjarne K. Ersbøll⁴

Department of Mathematics and Computer Science
Technical University of Denmark
Copenhagen-Denmark
{sarash, lkhc, bker@dtu.dk}

Mabel V. Martinez Vega²

Department of Plant and Environmental Science
University of Copenhagen
Copenhagen-Denmark
mmar@life.ku.dk

Abstract— Quality monitoring of the food items by spectroscopy provides information in a large number of wavelengths including highly correlated and redundant information. Although increasing the information, the increase in the number of wavelengths causes the vision set-up to be more complex and expensive. In this paper, three sparse regression methods; lasso, elastic-net and fused lasso are employed for estimation of the chemical and physical characteristics of one apple cultivar using their high dimensional spectroscopic measurements. The use of sparse regression reduces the number of required wavelengths for prediction and thus, simplifies the required vision set-up. It is shown that, considering a tradeoff between the number of selected bands and the corresponding validation performance during the training step can result in a significant reduction in the number of bands at a small price in the test performance. Furthermore, appropriate regression methods for different number of bands and spectrophotometer design are determined.

Keywords—Sparse regression, spectroscopy, lasso, elastic-net, fused lasso

I. INTRODUCTION

In food industry, the vision-based techniques such as spectroscopic measurements are widely used methods for quality monitoring of the food items. They acquire changes in the chemical and physical composition as factors of quality [1, 2]. For instance, the optical characteristics such as reflectance or absorbance measured by UV/VIS/NIR spectroscopy can represent the pigmentation and structural tissue changes in the plant organs.

There are different types of spectrophotometers used for spectroscopy and their spectral resolution (provided by monochromator) is an important characteristic showing the range of wavelengths they support [1, 3]. However, not all the wavelengths are equally important for characterization of food items. Usually the data in adjacent wavelengths are highly correlated and many of them are redundant, whereas other wavelengths may not carry relevant information for the problem at hand. Therefore, choosing a proper set of wavelengths carrying relevant information will help to simplify the vision system.

The aim of this paper is to solve such problems by employing sparse regression methods on UV/VIS/NIR spectroscopic data (306-1130 nm) of an apple cultivar. Two quality parameters, the sugar content called soluble solid content (SSC) and firmness of the apples [2] were predicted using their spectroscopic data. Sparse regression methods assist to reduce the number of wavelengths [4] and can simplify the vision set-ups used in food quality control [8]. We compared three sparse regression techniques; least angle shrinkage and selection operator (lasso) [4], elastic-net (EN) [4] and fused Lasso (FL) [5]. The data set was divided into different training and test sets four times and the average results are considered. A 10-fold cross validation (CV) was employed for training the prediction models. However, using the model parameters corresponding to the minimum validation error resulted in the use of a considerable number of wavelengths. In order to reduce the number of wavelengths even more, two strategies were investigated in the training phase. First, the one standard error rule was used [4]. In addition, manual selection of the proper number of wavelengths corresponding to an acceptable performance compared to the optimal point was performed. Results showed that both methods reduced the number of wavelengths significantly for all methods. However, this reduction was more considerable for firmness than SSC. In addition, the second strategy decreased the number of required wavelengths more and achieved better performance than the first one. Finally, a relation between the statistical methods and the design of the vision setups were determined.

II. DATA DESCRIPTION

The apple cultivar that was used in this paper is called “Rajka”. Spectroscopic measurements were performed on both sides, exposed and non-exposed to the sun, in 825 wavelengths (306-1130 nm) and the average results were considered. There were 185 data points (apple samples) in total. In addition, the SSC (%Brix) and the firmness (N) values for each apple were available from laboratory measurements. Figure 1 shows the spectroscopic data in UV/VIS and NIR wavelengths as well as the corresponding

This work was financed by the Centre for Imaging Food Quality project which is funded by the Danish Council for Strategic Research (contract no 09- 067039) within the Program Commission on Health, Food and Welfare.

sorted SSC and firmness signals. A slight trend can be seen in spectroscopic data in both UV/VIS and NIR bands that follows the corresponding sharp changes in SSC and firmness. In order to form the training and test sets, the samples were ranked in ascending order according to the SSC or firmness level. Then, from every 4 fruit samples, one was chosen as test (unseen data during training) and the rest as training. This was repeated 4 times by changing the number of test samples (1,2,3,4) to prepare 4 training and test sets ($X_{tr}, Y_{tr}, X_{ts}, Y_{ts}$).

III. STATISTICAL ANALYSIS

Since the number of samples $N=185$ was much smaller than the number of wavelengths $P=825$, the parsimonious regression methods were employed.

A. Lasso

Considering the general regression problem with $(X^i, y_i), i = 1, 2, \dots, N$, where $X^i = (x_{i1}, \dots, x_{iP})^T$ are the predictor variables and y_i are the responses, the lasso regression method estimates the regression coefficients $\hat{\beta}_{lasso}$ by minimizing the residual sum of squares so that, the L_1 norm of the coefficients is penalized [4]:

$$\hat{\beta}_{lasso} = \underset{\beta}{\operatorname{argmin}} \left\{ \frac{1}{2} \sum_{i=1}^N (y_i - \beta_0 - \sum_{j=1}^P x_{ij} \beta_j)^2 + \lambda \sum_{j=1}^P |\beta_j| \right\}$$

In $P > N$ case, the lasso selects at most N variables before it saturates [4]. Moreover, it does not support a grouping effect. That means that, if there is a group of variables with high pairwise correlations, then the lasso tends to select only one variable from the group. In this paper, the lasso algorithm implementation based on least angle regression (LAR) algorithm formed in [6] was used. The training involved a 10 fold CV to determine the number of non-zero coefficients varying from 1 to N .

B. EN

EN is a sparse regression method in which the regression coefficients $\hat{\beta}_{EN}$ are calculated based on both L_1 and L_2 norms penalty [4]:

$$\hat{\beta}_{EN} = \underset{\beta}{\operatorname{argmin}} \left\{ \frac{1}{2} \sum_{i=1}^N (y_i - \beta_0 - \sum_{j=1}^P x_{ij} \beta_j)^2 + \lambda_1 \sum_{j=1}^P |\beta_j| + \lambda_2 \sum_{j=1}^P |\beta_j|^2 \right\}$$

It can be used in ill-posed conditions where $N < P$. In addition, it has the grouping effect that helps to design a vision system with groups of close wavelengths. The EN implementation from [6] was used in this paper. Training was performed by a 10 fold CV with loops for selection of the

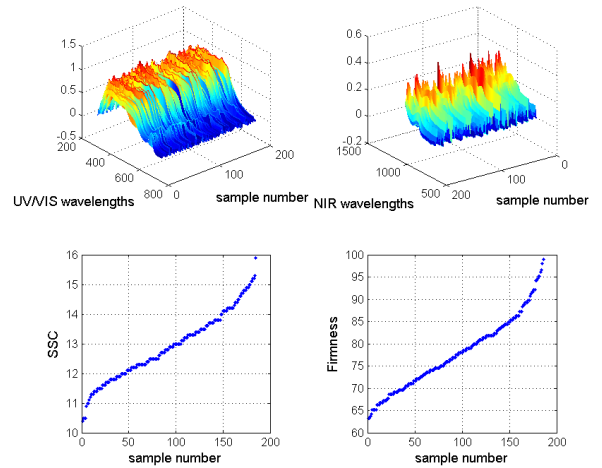


Fig. 1. The UV/VIS and NIR wavelengths sorted spectroscopic data and the corresponding SSC and firmness signals.

norm 2 penalty λ_2 and the number of non-zero coefficients.

C. FL

FL is a generalized version of lasso that encourages sparsity by means of the L_1 norm penalty on both regression coefficients and their successive differences [5]:

$$\hat{\beta}_{FL} = \underset{\beta}{\operatorname{argmin}} \left\{ \frac{1}{2} \sum_{i=1}^N (y_i - \beta_0 - \sum_{j=1}^P x_{ij} \beta_j)^2 + \lambda_1 \sum_{j=1}^P |\beta_j| + \lambda_2 \sum_{j=2}^P |\beta_j - \beta_{j-1}| \right\}$$

The fused lasso is especially useful for the $N < P$ cases, since it sets many coefficients to zero and finds groups of close features. The first penalty term encourages sparsity in the coefficients and the second one encourages sparsity in their differences. Therefore, with this solution we expect to find groups of adjacent wavelengths. In this paper, the implementation of this algorithm from the SLEP package [7] was used. A 10 fold CV with two loops for the choice of the two penalty terms were used for training the models. In addition, for each trained model the number of bands were calculated in each case and used for making decision about the best model parameters.

D. Model Selection

As mentioned in section 1, the one-standard error rule [4] was used as the first strategy for reducing the number of bands and making the models more parsimonious. The one-standard error rule picks the simplest model within one standard error from the minimum error point. Suppose that, we have P number of variables and M folds. The standard error of the error matrix $err_{P \times M}$ at point p is computed as follows:

$$se_p = \frac{std(err(p,:))}{\sqrt{M}}$$

Computing this value at the minimum point, the best model parameter can be found at one se_p distant from the minimum point in the direction where less bands are chosen. The second selection strategy is simply a manual selection by comparing the error at the minimum point and the points with less number of bands. In fact, a tradeoff was made between the reduction in number of bands and the increase in error.

IV. EXPERIMENTAL RESULTS

In order to evaluate the regression methods the root mean square error and the R-square criteria were used:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}, \quad R^2\% = \left(1 - \frac{RSS}{TSS}\right) \times 100\%,$$

$$RSS = \sum_{i=1}^n (y_i - \hat{y}_i)^2, \quad TSS = \sum_{i=1}^n (y_i - \bar{Y})^2$$

where y_i and \hat{y}_i are the original and estimated response values and \bar{Y} is the mean value of all target values.

In figure 2 the minimum error points as well as the points selected by the one standard error rule and also manual selection are shown for the three methods. Since the dimension of the resulting average validation error map from the 10 fold CV loop varies for the three regression methods, the three plots do not represent an equal number of points. In case of lasso, the number of bands was the only model parameter and thus the average validation error was an N length vector. For EN model, besides the number of bands, the norm 2 penalization coefficient λ_2 was the second parameter. Therefore, the illustration was performed by reducing the dimension into one so that, just the minimum error for each lambda as well as the points found by the two selection strategies are illustrated. Finally, for FL, the two varying parameters were λ_1 and λ_2 that created a 2D average error map. However, in this work we are interested to make the decision based on the number of wavelengths. Therefore, a 2D matrix of the same length showing the average number of wavelengths over the 10 folds was also calculated. Then, these two matrixes were vectorized and shown verses each other. In all of these cases, the selection direction of the new points was defined in a way that the simplest and most parsimonious models could be selected. The density of selection of different wavelengths for the four data sets in SSC estimation is shown in figure. 3. As can be seen, the UV, VIS and NIR regimes are selected by all three methods.

Finally, the average results of the four training and test sets from all the three regression methods for the SSC and firmness are presented in tables 1 and 2. As can be seen, in both cases moving from the minimum point toward a new point selected manually or by one standard error rule significantly reduced the average number of wavelengths N_w

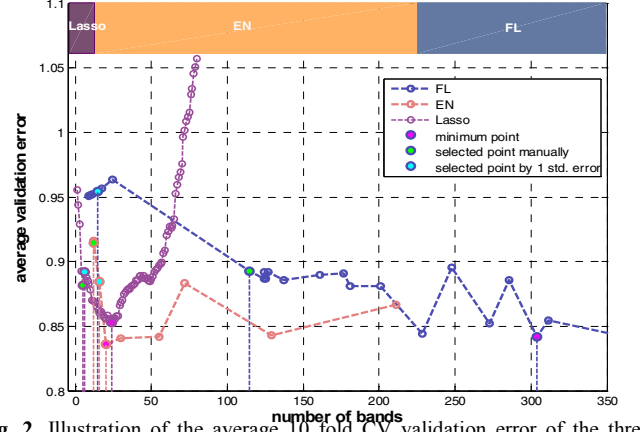


Fig. 2. Illustration of the average 10 fold CV validation error of the three regression methods for set 1.

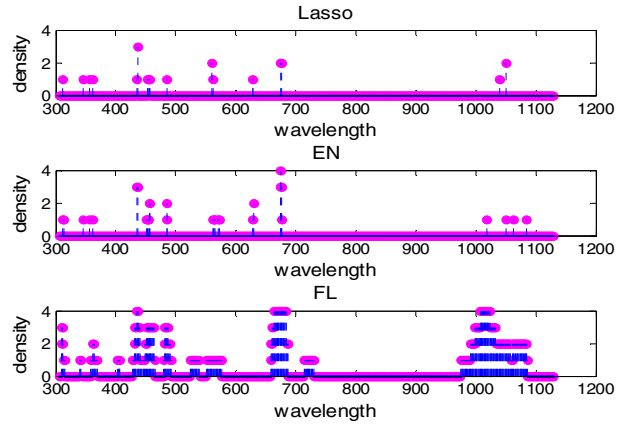


Fig. 3. Density of band selection for SSC estimation using the three regression methods. Model parameters were obtained by manual selection for 4 data sets.

and also reduced the over-fitting by decreasing the difference between the training and test performances. Moreover, the manual selection reduced the number of bands more than the one standard error rule.

A. Discussion

Based on the results, the methods that are suitable for vision set-ups with different number of bands are illustrated on top of figure 2 by different colors. Besides the number of bands, the width of the regimes is important in spectrophotometer design. Considering figure 2 and 3 and the two tables, lasso is suitable when a few individual narrow bands (less than 10 bands) can be provided by e.g. a few LEDs. EN is suitable when more bands (up to 200) in narrow regimes can be supplied. A monochromator capable of selecting a few narrow regimes of laser light suits this case. Finally, FL is the best choice when a lot of bands (e.g. more than 200) in broad regimes of laser light are available. The monochromator does not need to provide high resolution in this case.

Table 1. The average results of the three regression methods for SSC.

SSC		Minimum point	Manual Selection	1 Std. Error Rule
FL	$R_{ts}^2\%$	41.12	39.40	37.80
	$R_{tr}^2\%$	55.70	42.68	43.98
	$rmse_{ts}$	0.86	0.87	0.88
	$rmse_{tr}$	0.76	0.87	0.86
	N_w	532.5	142.75	158.75
EN	$R_{ts}^2\%$	41.28	41.32	37.86
	$R_{tr}^2\%$	53.32	44.92	39.59
	$rmse_{ts}$	0.86	0.86	0.88
	$rmse_{tr}$	0.78	0.85	0.89
	N_w	30.25	11.50	19.25
LASSO	$R_{ts}^2\%$	41.55	40.87	35.16
	$R_{tr}^2\%$	55.46	44.39	37.02
	$rmse_{ts}$	0.86	0.86	0.90
	$rmse_{tr}$	0.76	0.85	0.91
	N_w	16.0	6.25	5.0

Table 2. The average results of the three regression methods for firmness.

Firmness		Minimum point	Manual Selection	1 Std. Error Rule
FL	$R_{ts}^2\%$	47.33	41.53	36.52
	$R_{tr}^2\%$	67.41	42.89	46.35
	$rmse_{ts}$	5.74	6.05	6.29
	$rmse_{tr}$	4.71	6.24	5.99
	N_w	795.0	22.25	179.0
EN	$R_{ts}^2\%$	47.50	39.96	44.75
	$R_{tr}^2\%$	68.46	39.87	53.72
	$rmse_{ts}$	5.73	6.14	5.88
	$rmse_{tr}$	4.63	6.39	5.60
	N_w	758.5	16.0	307.5
LASSO	$R_{ts}^2\%$	45.17	41.97	33.20
	$R_{tr}^2\%$	68.87	44.12	36.43
	$rmse_{ts}$	5.86	6.03	6.47
	$rmse_{tr}$	4.60	6.17	6.56
	N_w	36.75	4.0	4.75

I. CONCLUSION

In this paper, three regression methods; lasso, EN and fused lasso were used for estimation of the SSC and firmness level of an apple cultivar using their spectroscopic data. By manual selection of a new point or using the one standard error rule instead of the minimum error point, we could significantly reduce the number of required wavelengths for training the prediction model at a price of a small increase in the test error. Finally, the proper regression methods with different number of bands and types of spectrophotometer design were defined in the discussion.

REFERENCES

- [1] B. Herold, S. Kawano, B. Sumpf, P. Tillmann, and K. B. Walsh, *Book Chapter: VIS/NIR spectroscopy*, CRC Press, London, 2009.
- [2] B. Nicolai, K. Beullens, E. Bobelyn, A. Peirs, W. Saeys, K. Theron and J. Lammertyn, "Non-destructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review," *Postharvest Biology and Technology*, Elsevier, pp. 99-118, 2007.
- [3] D. Wen Sun, *Hyperspectral imaging for food quality analysis and*

control, Elsevier academic press, United Kingdom, 2010.

- [4] T.Hastie, R. Tibshirani and F. Jerome, *The Elements of Statistical Learning*, Springer, New York, 2008.
- [5] R. Tibshirani and M. Saunders, "Sparsity and smoothness via the fused lasso," *Journal of royal statistical society (67)*, Blackwell Publishing, United Kingdom, pp. 91-108, 2005.
- [6] K. Sjöstrand, "Regularized Statistical Analysis of Anatomy," *PhD thesis*, Department of Informatics and Mathematical Modeling, Technical University of Denmark, 2007.
- [7] J. Liu, S. Ji, and J. Ye. "SLEP: Sparse Learning with Efficient Projections," Arizona State University, 2009.
- [8] S. Sharifzadeh, J. L. Skytte, O. H. A. Nielsen, B. K. Ersbøll, L. K. H. Clemmensen, "Regression and Sparse Regression Methods for Viscosity Estimation of Acid Milk From its SLS Features," *Proceedings: IWSSIP 2012, 11-13 April 2012*, pp. 58-61, Vienna, 2012.