Calibration Color Patch Reduction for Electrophotography*

Yan-Fu Kuo[a], George T.-C. Chiu[b], Yuehwern Yih[c], and Jan P. Allebach[d]; a: Department of Bio-industrial Mechatronics Engineering, National Taiwan University, Taiwan ROC, b: School of Mechanical Engineering, c: School of Industrial Engineering, d: School of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana USA

Abstract

This paper presents results of reducing color patches printed during a calibration using a variable selection approach for color electrophotography (EP). During a calibration, primary color patches at different halftone levels are printed on a belt and measured using on-board sensors. Regression models are used to predict primary color tone values on output media from these on-board sensor measurements. Then necessary adjustments using tone correction or other control variables can be applied to compensate tone deviation. Laying down and measuring color patches consumes time and toner. It is desired to determine a minimal set of color patches for the regression models while ensuring their prediction accuracy. This work proposes a procedure that reduces color patches in regression model development. The procedure applies variable selection algorithms to identify the appropriate number of color patches and their associated halftone levels. Results demonstrate that the proposed method can use 45% fewer color patches for the regression models while maintaining prediction accuracy effectively the same as that of models using all the available color patches.

Introduction

A color electrophotographic (EP) printing system typically uses four primary colors – cyan, magenta, yellow, and black. Calibrations are performed to maintain consistent color reproduction under different operating conditions [1]. During a calibration, multiple color patches at different halftone levels of the same primary color are printed on a transfer belt inside the printer and are measured with on-board sensors. Calibration models then are used to predict the primary color tone values on output media from these on-board sensor measurements. Printing and measuring color patches consume time and toner and are not economically preferred. In this study, our aim is to reduce the number of color patches used for the calibration models while maintaining a desired tone prediction accuracy of the models for color EP systems.

The calibration models are developed with data collected in printer life tests. During these tests, color patches are printed on output media immediately following a calibration and are measured with spectrophotometers. Calibration models are then developed as a mapping of the on-board sensor measurements to output tone values measured on media [2]. When a calibration is performed while the product is in use in the field, on-board sensor measurements are taken to predict tone values with the calibration models. Appropriate tone correction is then performed by adjusting bias voltages or modifying the tone correction mapping. Since laying down and measuring color patches consumes time and toner, it is crucial to use a minimal set of color patches for the calibration models while ensuring their reliability.

Recent research [3] has shown improved prediction accuracy of the calibration models using a principal component regression (PCR) approach. The work included all the available on-board sensor measurements as explanatory variables of the calibration models to improve their prediction accuracy, and used a principal component regression (PCR) approach to tackle the numerical challenges introduced by the multicollinearity among the on-board sensor measurements. While this approach significantly improved model prediction accuracy, the issues of redundant on-board sensor measurements were not addressed. Some of the color patches may provide very trivial contribution to model accuracy improvement. Not printing and measuring these color patches can reduce calibration process time and cost without losing significant prediction accuracy.

In this work, we propose a procedure to select a set of color patches and their associated halftone levels that are significant to model prediction accuracy using the *least absolute shrinkage and selection operator* (*LASSO*) [4]. The *LASSO* method provides a trade-off between model prediction accuracy and color patch economy. To illustrate the utility of the proposed approach, a first-order linear calibration model for an off-the-shelf in-line color EP printer is developed using existing life test data. Cross-validation results demonstrate a 45% reduction in the number of color patches required in calibration model development while a desired prediction accuracy of the calibration models can be maintained.

Method

A calibration model is a mapping from a set of on-board sensor measurements to a set of tone values obtained during and immediately after a calibration. Let $\boldsymbol{x} = [x_j] \in \Re^{1\times p}$ denote a set of on-board sensor measurements from color patches printed on the transfer belt at halftone levels $\boldsymbol{h} = [h_j] \in \Re^{1\times p}$, and $\boldsymbol{y} = [y_j] \in \Re^{1\times q}$ denote a set of tone values from color patches printed on the output media at halftone levels $\boldsymbol{g} = [g_j] \in \Re^{1\times q}$. Note that these halftone levels $\boldsymbol{h}$ and $\boldsymbol{g}$ are usually fixed after the development stage of the product. In this work, the calibration model is

* Research supported by the Hewlett-Packard Company. Research partially conducted while Yan-Fu Kuo was a graduate student at Purdue University.

formulated as a linear transformation relating the tone values on the output media to the on-board sensor measurements, i.e.,

$$\boldsymbol{y} = \boldsymbol{x}\boldsymbol{B}, \tag{1}$$

where $\boldsymbol{B} \in \Re^{p\times q}$ is the calibration model matrix. Without loss of generality, it is assumed that the tone values are standardized columnwise. Hence no intercept vector is required in the formulation. Note that, typically, a calibration model is developed separately for each primary color since the primary colors are reproduced independently in in-line single-pass color EP printers.

Calibration Model Development

Calibration models are developed using regression analysis. Suppose that $n \in \boldsymbol{N}$ repetitions of calibration are made in the life test. Denote $x_{ij} \in \Re$ and $y_{ij} \in \Re$ as the j^{th} on-board sensor measurement and the j^{th} tone value measurement, respectively, in the i^{th} repetition. The regression analysis can be performed separately for an individual tone value associated with a specific halftone level. Denote $\boldsymbol{x}^{(j)} = [x_{ij}] \in \Re^{n\times 1}$ as all the on-board sensor measurement repetitions associated with the halftone level h_j in a form of a column vector, and $\boldsymbol{y}^{(j)} = [y_{ij}] \in \Re^{n\times 1}$ as all the tone value repetitions associated with the halftone level g_j in a form of a column vector. Let $\boldsymbol{\beta}^{(j)} = [\beta_{ij}] \in \Re^{p\times 1}$ represent the j^{th} column of the calibration model matrix $\boldsymbol{B}$, i.e.,

$$\boldsymbol{B} = [\boldsymbol{\beta}^{(1)}\ \boldsymbol{\beta}^{(2)} \ldots \boldsymbol{\beta}^{(q)}]. \tag{2}$$

The calibration model column $\boldsymbol{\beta}^{(j)}$ represents the regression model coefficients to predict a tone value associated with a halftone level g_k. The loss function to be minimized in the regression is

$$\begin{aligned}\boldsymbol{\beta}^{(j)} = \beta_{ij} &= \arg\min\left\{\boldsymbol{y}^{(j)} - [\boldsymbol{x}^{(1)}\boldsymbol{x}^{(2)} \ldots \boldsymbol{x}^{(p)}]\boldsymbol{\beta}^{(j)}\right\} \\ &= \arg\min\left\{\sum_{i=1}^{n}\left(y_{ij} - \sum_{k=1}^{p} x_{ik}\beta_{ij}\right)^2\right\}.\end{aligned} \tag{3}$$

The calibration model is the concatenation of the individual columns $\boldsymbol{\beta}^{(j)}$ obtained by solving Eq. (3).

Since there exists a high degree of multicollinearity among the on-board sensor measurements of the same primary color at different halftone levels [3], some on-board sensor measurement columns $\boldsymbol{x}^{(j)}$s in Eq. (3) that are associated with halftone levels h_js may provide redundant information. Removing these on-board sensor measurements $\boldsymbol{x}^{(j)}$s in the regression may not significantly reduce the model prediction accuracy. The objective of this work is to find a minimal common set of halftone levels $\overline{\boldsymbol{h}} \subset \boldsymbol{h}$ and their associated on-board sensor measurement columns $\boldsymbol{x}^{(j)}$s, where $h_j \in \overline{\boldsymbol{h}}$, that should be retained in the regression to maintain the prediction accuracy of the model. Once the selected halftone levels are determined, printing the color patches associated with the remaining halftone levels can be skipped in a calibration and hence time and consumable usage can be reduced.

LASSO Calibration Model

The *LASSO* method is applied to achieve color patch reduction and halftone level selection. The *LASSO* formulates the regression as a problem of the loss function in Eq. (2) in addition to the constraint that the L[1]-norm of $\boldsymbol{\beta}^{(j)}$ be no greater than a given tuning parameter $0 \le t \in \Re$, i.e.,

$$\left\|\boldsymbol{\beta}^{(j)}\right\|_1 = \sum_{i=1}^{p}\left|\beta_{ij}\right| \le t\,. \tag{4}$$

When the tuning parameter t is small enough, some of the β_{ij}s can be shrunk to zero. This constrained least-squares formulation is useful in this application due to its tendency to prefer solutions with fewer nonzero parameter values, effectively reducing the number of on-board sensor measurements used in the regression. The *LASSO* problem can be solved using quadratic programming.

A two-step color patch reduction procedure is proposed in this work. In step I, the *LASSO* is performed separately for each individual tone value to obtain a set of on-board sensor measurements and their associated halftone levels to be included in the calibration model for that tone value. In step II, a common set of on-board sensor measurements to be used in the final calibration model is then determined as a trade-off between prediction accuracy and color patch economy.

Step I – On-board Sensor Measurement Selection for Individual Tone Values

The *LASSO* formulation of Eqs. (3) and (4) is applied to determine a subset of on-board sensor measurements and their associated halftone levels to be retained in the model for each tone value. Consider a tone value y_j associated with halftone level g_j is examined. In the process, the tuning parameter t is gradually reduced from the sum of the absolute coefficients $\Sigma_{i=1}^{p}\left|\beta_{ij}\right|$ of a full model down to zero. Along with the reduction of the tuning parameter t, more and more coefficients in $\boldsymbol{\beta}^{(j)}$ are shrunk to zero. The on-board sensor measurement subsets associated with different numbers of non-zero coefficients β_{ij} are recorded.

After the *LASSO* solutions are obtained, a tenfold cross-validation (CV) is performed to estimate the prediction accuracies of the models using different numbers of on-board sensor measurements as explanatory variables. In the process, the data is split into roughly ten equal-sized parts. Parts of the data are used in the regression to obtain the model coefficients $\boldsymbol{\beta}^{(j)}$, and a different part of the data is used to test the model accuracy. Errors between prediction and true values and their aggregated root-mean-squared values (RMSE) are computed. Standard errors of the individual RMSE for each of the ten parts are also calculated. Then a "one-standard error" rule [5] is used to choose the subset of on-board sensor measurements. Typically, the most parsimonious model whose error is no more than one standard error above the RMSE of the most accurate model is chosen.

Step II – Common On-board Sensor Measurement Selection for the Calibration Model

Frequencies of the *LASSO* selected on-board sensor measurements $\boldsymbol{x}^{(j)}$ and their associated halftone levels h_j for all the tone values are listed. The frequencies are used to prioritize the significance rank for these on-board sensor measurements. After that, accuracies of calibration models using different numbers of on-board sensor measurements as explanatory variables are calculated using CV. Then the final calibration model is selected based on specified criteria.

The proposed color patch reduction procedure is performed on an off-the-shelf single-pass color EP printer to develop the calibration model. The printer prints and measures seventeen calibration patches at different halftone levels for each primary color during a calibration, i.e., $p = 17$. These halftone levels are labeled as h_j, $j = 1\ldots17$, from light to dark. Calibration patches identical to those printed in a calibration are printed on output media for each primary color immediately following a calibration, i.e., $\boldsymbol{g} = \boldsymbol{h}$. Their tone value measurements are made with spectrophotometers (X-Rite® DTP-70). In this study, a tone value is defined as the Euclidian distance (ΔE) in CIE L*a*b* space between the color point of a primary color printed at a particular halftone level and the substrate appearance color. A commercial white paper (Xerox® 4200 Business) is used as the output media. The experiment is performed on five printers with several consumable sets under a wide range of environmental conditions. The temperature ranges from 15 to 30°C, and the relative humidity ranges from 10 to 80%. Totally four hundred and twenty observations, i.e., $n = 420$, are collected. The models are identified following the proposed procedure using Matlab®.

Figures 1(a) and 1(b) shows the halftone levels of the on-board sensor measurements to be retained in the calibration model and the corresponding CV RMSEs, respectively, for the tone value associated with the halftone level g_9 for Magenta. In this example, the model with five explanatory variables – h_2, h_7, h_{14}, h_{15}, and h_{16} – is chosen. Note that during the shrinkage, some coefficients of on-board sensor measurements at different halftone levels may be reduced to zero simultaneously (for example, see in Fig. 1(a) that h_7 and h_{16} are reduced to zero at the same time). Figure 2 shows the frequencies of the *LASSO* chosen halftone levels for all the tone values. It is shown that for some primary colors, more than half of the on-board sensor measurements are not chosen. This verifies the potential to reduce the number of color patches printed during a calibration.

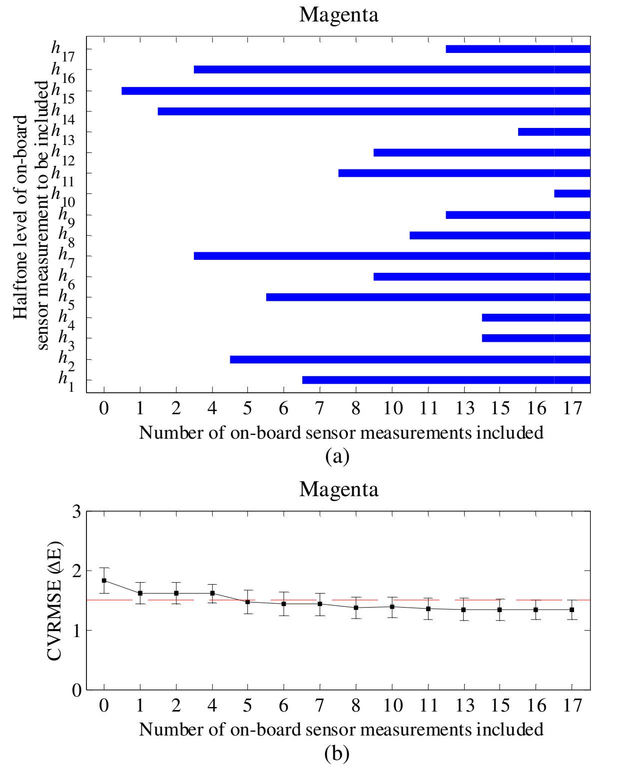

Figure 1. *(a) The LASSO selected halftone level subsets and (b) their cross-validation root-mean-squared errors (CVRMSE) for tone values associated with the halftone level g_9 for Magenta. In Fig. 1(b), the error bars represent the standard errors of the CV RMSEs. The red dashed line represents the one standard error above the RMSE of the most accurate model, which is the model with seventeen on-board sensor measurements in this example.*

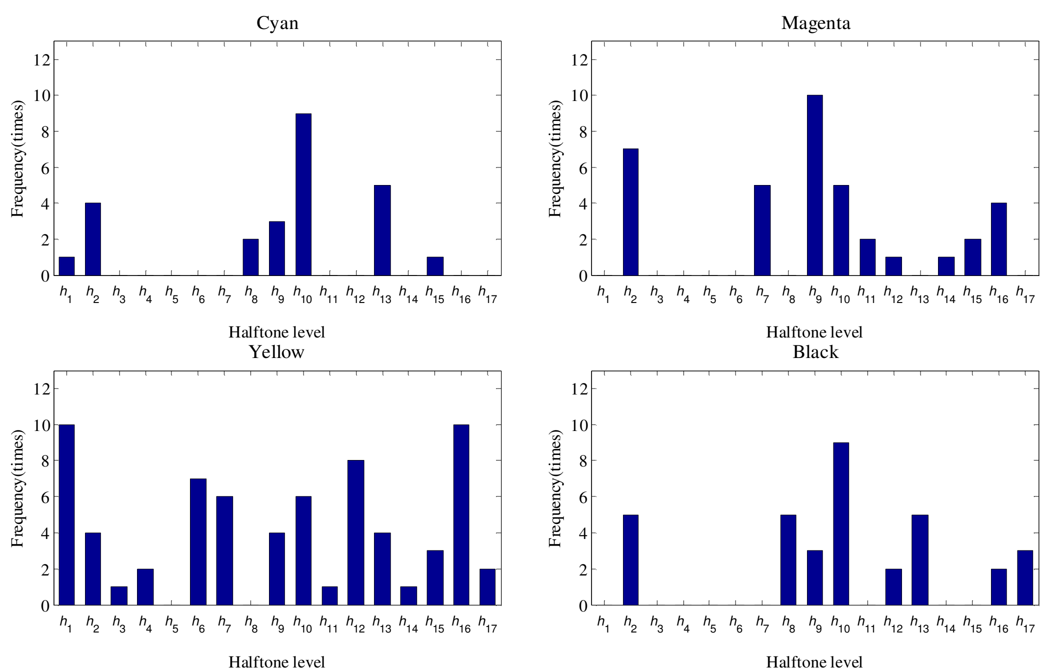

Figure 2. *Frequencies of the LASSO selected on-board sensor measurements at different halftone levels (h_j) for each primary color.*

Model Performance Comparison

The *LASSO* models using all the selected on-board sensor measurements shown in Fig. 2 as the explanatory variables are developed through the proposed procedure and are compared with existing models and full models. In the existing models, each tone value is regressed with a single on-board sensor measurement printed at the same halftone level. In the full models, each tone value is regressed with all the available on-board sensor measurements using the PCR approach [3]. The *LASSO* models use a total of thirty-nine on-board sensor measurements, compared to sixty-eight measurements for the full models. Figure 3 shows the CV RMSEs of the three models. Statistical t-tests are conducted to compare the CV squared errors of the *LASSO* and full models. The results show that there exists no significant difference in their CV squared errors at all halftone levels for all colors at 95% confidence level.

Conclusion

A color patch selection procedure using the *LASSO* approach is proposed in this work to reduce the number of color patches printed during a calibration while maintaining the desired prediction accuracy for the calibration model. The effectiveness of the proposed method is verified with experimental data collected under different environmental conditions and consumable usage levels from several single-pass in-line color EP printers of the same model. The performance of the proposed *LASSO* models is compared to that of the full models which use all on-board sensor measurements as the explanatory variables. The accuracy of the *LASSO* models effectively is the same as that of the full models based on statistical t-tests while it uses 45% fewer color patches.

Acknowledgements

We gratefully acknowledge the support from the Hewlett-Packard Company. We would like to especially thank Jeff Trask and George Kerby for their valuable guidance in this research.

References

[1] C. Yang, Y. Kuo, Y. Yih, G. T. Chiu, D. A. Abramsohn, G. R. Ashton, and J. P. Allebach, "Improving tone prediction in calibration of Electrophotographic printers by linear regression: environmental, consumables, and tone-level factors," J. Imaging Sci. Tech. 54: 050301 (2010).

[2] D. A. Johnson, "Calibration of printing devices," U. S. Patent 6,982,812 (2006).

[3] Y. Kuo, C. Yang, Y. Yih, G. T. Chiu, and J. P. Allebach, "Improving tone prediction in calibration of Electrophotographic printers by linear regression: using principal components to account for collinearity of sensor measurements," J. Imaging Sci. Tech., 54: 050302 (2010).

[4] R. Tibshirani, "Regression shrinkage and selection via the LASSO," J. Roy. Statist. Soc. B 58: 267 (1996).

[5] T. Hastie, R. Tibshirani. and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, (Springer-Verlag, New York, 1990).

Author Biography

Yan-Fu Kuo is an assistant professor in the Department of Bio-industrial Mechatronics Engineering at National Taiwan University. He obtained his Ph.D. degree from Purdue University. His research interests include dynamic system modeling and control, machine learning, and statistics applications.

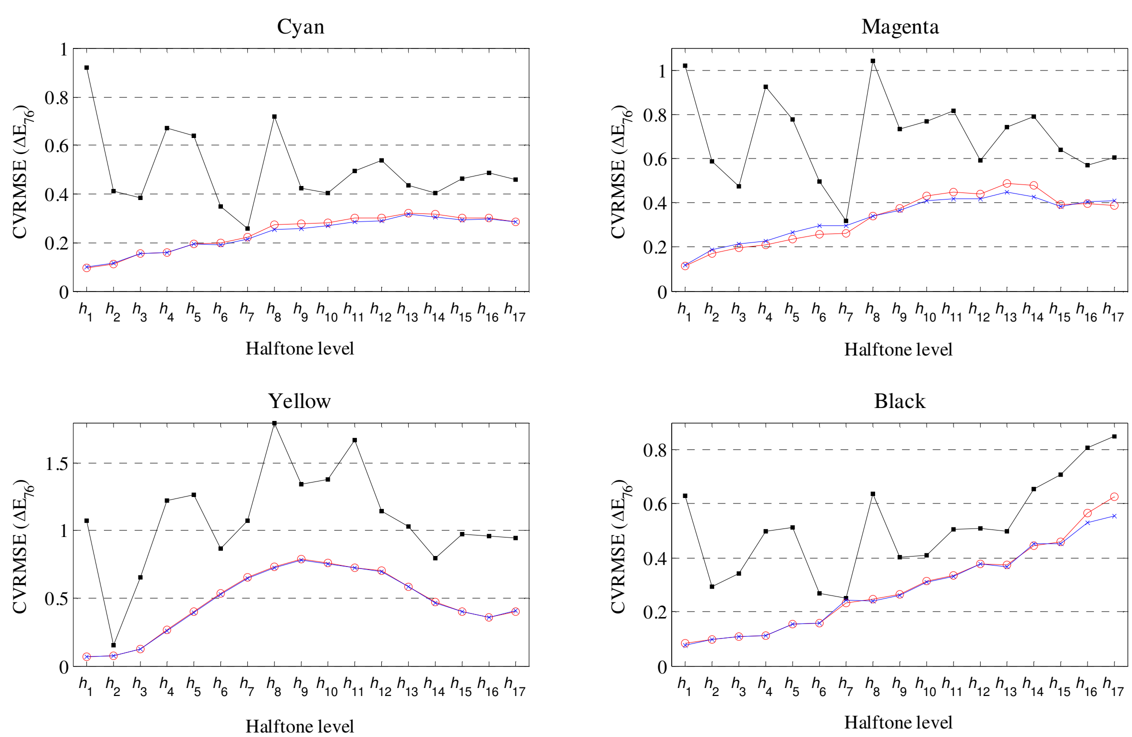

Figure 3. *Cross-validation root-mean-squared errors (CVRMSE) of the existing(■), LASSO(○), and full (x)models at each halftone level (h_j).*