Comparison of Interpolation Techniques for Characterisation of a Half-Tone Printer

Shankhya Debnath

September 2, 2024

1 Introduction

The accuracy of various interpolation techniques for the characterization of a half-tone printer using the IT8.7/4 test target and the FOGRA39 characterization dataset has been presented. Characterization involves predicting color outputs (in CIELAB space) based on given inputs (in CMYK space). This is essential for ensuring color consistency and accuracy in printing applications.

2 Methodology

To evaluate the accuracy of different interpolation techniques, we used the FOGRA39 dataset, which provides both CMYK and corresponding LAB color values for the IT8.7/4 test target (containing 1617 color patches). The dataset was split into 80% training data and 20% testing data. Several interpolation techniques were then applied to predict the LAB values from the CMYK inputs in the test set. The techniques evaluated include:

- Distance-Weighted Interpolation
- Nearest Neighbor Interpolation
- Linear Interpolation
- 2nd Degree Polynomial Fitting
- 3rd Degree Polynomial Fitting
- Optimized Neural Network

3 Mathematical Background

3.1 Distance-Weighted Interpolation

Distance-Weighted Interpolation estimates the LAB values for a given test CMYK sample by assigning weights to each training sample based on the inverse of the distance between the test sample and the training samples in CMYK space. The weights are calculated as:

$$\omega_i = \left(\frac{1}{\text{distance}_i + \epsilon} \right)^p \tag{1}$$

where:

- distance_i is the Euclidean distance between the test sample and the i-th training sample in CMYK space.
- ϵ is a small constant added to avoid division by zero.
- p is a parameter that controls the influence of the distance.

The estimated LAB values are then computed as a weighted average:

$$\hat{L}, \hat{a}, \hat{b} = \frac{\sum_{i=1}^{N} \omega_i \cdot (L_i, a_i, b_i)}{\sum_{i=1}^{N} \omega_i} \tag{2}$$

3.2 Nearest Neighbor Interpolation

In Nearest Neighbor Interpolation, the LAB value of the nearest training sample in CMYK space is directly assigned to the test sample. This method does not involve any weighting or averaging.

3.3 Linear and Polynomial Interpolation

Linear Interpolation estimates LAB values using a linear combination of nearby training samples. Polynomial Interpolation, on the other hand, fits a polynomial of a specified degree (2nd or 3rd) to the training data in CMYK space:

$$\hat{L}, \hat{a}, \hat{b} = \text{PolynomialRegression}(\text{train\_CMYK}, \text{train\_LAB}) \tag{3}$$

where the polynomial regression model is used to fit the relationship between CMYK and LAB values.

3.4 Neural Network

A neural network model was used to capture the non-linear relationship between CMYK and LAB values. The network consists of multiple layers of neurons with ReLU activation functions and incorporates techniques like dropout and batch normalization to prevent overfitting and improve generalization.

4 Results and Observations

The performance of each technique was evaluated by calculating the mean CIELAB color difference (ΔE) between the predicted and actual LAB values in the test set. The results are summarized in Table 1.

Method	**Mean ΔE**
Distance-Weighted Interpolation	2.28
Nearest Neighbor Interpolation	3.97
Linear Interpolation	0.14
2nd Degree Polynomial Fitting	1.83
3rd Degree Polynomial Fitting	0.62
Neural Network	3.95

Table 1: Mean CIELAB color difference (ΔE) for each method.

4.1 Significance of Results

The results demonstrate that Linear Interpolation and 3rd Degree Polynomial Fitting achieve the lowest mean ΔE values, indicating the highest accuracy in predicting LAB values from CMYK inputs. Linear Interpolation, in particular, performs exceptionally well, with a mean ΔE of just 0.14, suggesting near-perfect accuracy for most samples.

Distance-Weighted Interpolation and 2nd Degree Polynomial Fitting also perform well, though they show slightly higher mean ΔE values, indicating less accuracy compared to the Linear and 3rd Degree Polynomial methods.

The Neural Network, despite being optimized with advanced techniques such as dropout and batch normalization, does not perform as well as expected, achieving a mean ΔE of 3.95. This suggests that while neural networks are powerful tools, their performance is highly dependent on the architecture, training process, and data distribution.

5 Conclusion

The results indicate that Linear Interpolation and 3rd Degree Polynomial Fitting are the most accurate methods for this task. The Neural Network, while capable, did not outperform the simpler interpolation methods in this case.

These findings suggest that for the task of predicting LAB values from CMYK inputs in printer characterization, traditional interpolation methods can be highly effective and, in some cases, may outperform more complex models such as neural networks. However, since this work was carried out using the characterization dataset, the results can vary dramatically if real world color data is used.