A Novel Color Temperature Prediction Algorithm by Machine Learning

Yi-Ting Chung, Kun-Yi Lin, Ding-Teng Shih, Hsin-Ting Chan, Chui-Yi Chou
Novatek Microelectronics Corporation, Hsinchu, Taiwan

Abstract

In this paper, we propose using a machine learning algorithm to predict correlated color temperature (CCT) with in-cell ambient color sensors.

Machine learning algorithms can more effectively identify the relationship between sensor data and color temperature values. Using this method, we can reduce the maximum prediction error from 250 to 150 across five light sources ranging from CCT 2300K to 6500K.

Author Keywords

Ambient color sensor, Machine learning, correlated color temperature, In-cell technology.

1. Introduction

With the advent of advanced technology, full-screen displays on phones and tablets have become a trend. To enhance the aesthetic appeal and integration convenience of screens, various sensors, including ambient light/color sensors, need to be integrated into the panel. [1]

Even without considering the hardware integration challenges, the integration of ambient light/color sensors into the panel is complicated. As shown in Figure 1, the prediction objective functions and parameters for light source color temperature vary due to different cover glass and color filter glass factors. Identifying these relationships quickly is a highly complex task.

However, machine learning surpasses basic formula derivation and programming by generating outputs through the analysis of complex data. The goal of machine learning is to enable machines to analyze large amounts of data using statistical models to identify patterns and produce results, thereby allowing the machine or system to automatically learn and improve from experience. [2]

This study will employ machine learning models, such as AdaBoost[3] and GBDT[4], and compare them with traditional methods across multiple color temperatures and an illuminance range of 800 lux to 6500 lux.

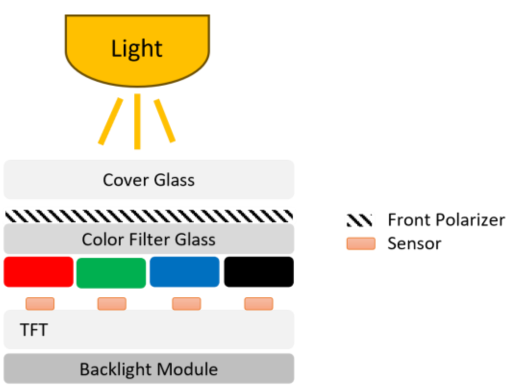

Figure 1. The structure of in-cell integrated ambient/color sensor.

2. CCT Predict Algorithm with Machine Learning

The training flow using machine learning (ML) is shown in Figure 2.

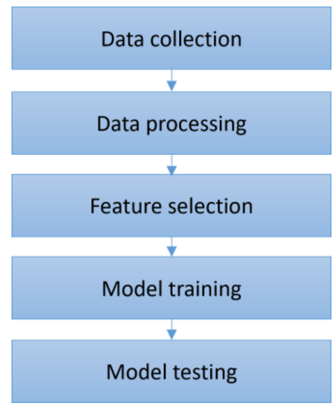

Figure 2. The training flow.

Data Collection:

1. Sensor data collected under different light sources and brightness levels.
2. Measurement of color temperature and illuminance values.
3. Noise sensor data collected without a light source

6500K	Data 1	Data 2	...	CCT
0500lux (01)	1072	949		6216.33
0500lux (...)	1070	944		6216.33
0500lux (50)	1067	941		6216.33
0600lux (01)	1120	965		6269.18
...				
6500lux (50)	3152	1530		6372.49
Without light	923	902		NAN

Table 1: Sample Data Collection on D65

Data Processing:

1. Remove saturation data, as illustrated in Figure 3
2. Noise reduction: Data = Raw – Base
3. Total energy: Sum of all sensor data
4. Color energy ratio1: Color1/<sum of all color data>
5. Color energy ratio2: Color1/without CF data
6. Refer to training base raw: Data = Raw – Base + Training Base
7. Generate polynomial and interaction features.

ISSN 0097-996X/25/5601-0286-$1.00 © 2025

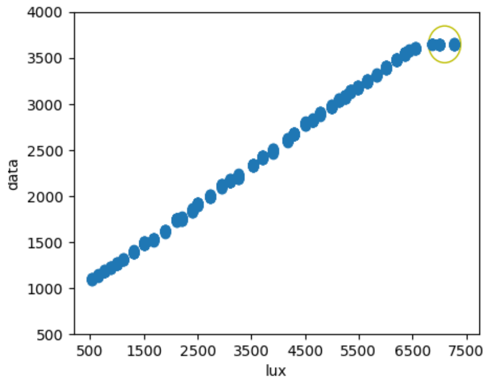

Figure 3. Data1 Distribution Chart

Feature selection:

After data preprocessing, a total of 22 features are obtained. Select the most relevant features for each ML model.

AdaBoost:

List the feature importance proportions in Figure 4. The following 7 features are selected:

- R_ratio1: red color filter on color energy ratio1
- B_ratio1: blue color filter on color energy ratio1
- RWinter: R data and W data interaction
- R_diff: R data without noise
- GWinter: G data and W data interaction
- W_raw: W data
- BWinter: B data and W data interaction

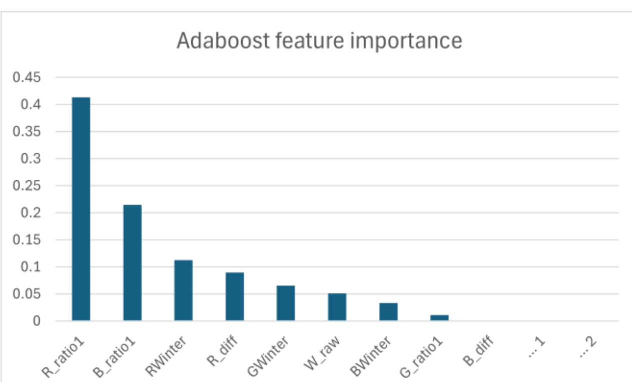

Figure 4. Features importance in AdaBoost model

GBDT:

List the feature importance proportions in Figure 5. The following 6 features are selected:

- B_ratio1: blue color filter on color energy ratio1
- R_ratio2: red color filter on color energy ratio2
- R_ratio1: red color filter on color energy ratio1
- B_ratio2: blue color filter on color energy ratio2
- B_diff: B data without noise
- B_train_raw: refer to training raw

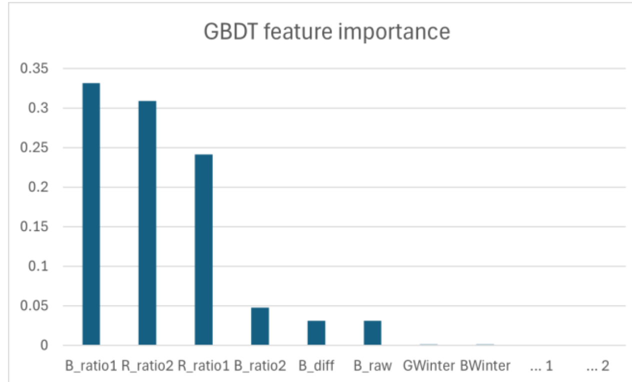

Figure 5. Features importance in GBDT model

Model training:

The model training process involves selecting the optimal hyperparameters for the learning algorithm. These hyperparameters guide the algorithm to learn the optimal parameters that accurately map the input features to the labels or targets.

The collected data will be divided into 75% for training and 25% for model testing.

Experiment Result:

Integrate the trained model into the system and conduct tests with various color temperatures to evaluate the error range.

To simplify the training process and reduce costs, only the data from In-cell Module 1 will be used for training. However, to verify performance and achieve scalability for mass production, not only Module 1 but also other modules will be used for validation.

3. Training Result

Collect three types of color temperature data and apply the training flow. The Model training result in Figure 6 indicates a maximum error is 62.54.

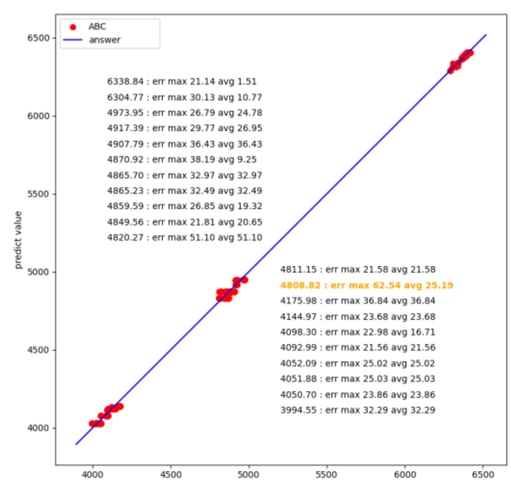

Figure 6. ML train and test result with 3 CCTs

Collect five types of color temperature and apply the training flow. The Model training result in Figure 7 shows that the maximum error is 129.75.

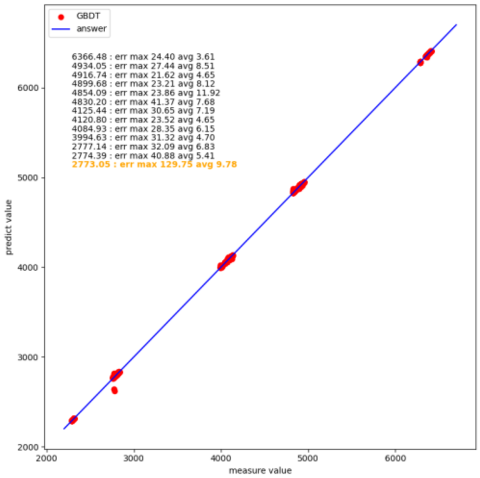

Figure 7. ML train and test result with 5 CCTs

As shown in Figure 6 and 7, only the instances where the maximum error exceeds 21 are displayed, and not all data is presented in text form.

As shown in Figure 8, the maximum error for the color temperature 2773.05K exceeds 100, but this single data point can be considered as noise interference.

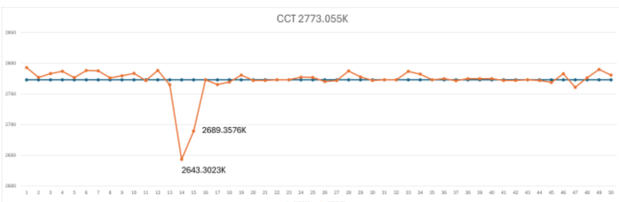

Figure 8. The detail of 2773.05K CCT prediction

Based on the above results, the average error is less than 55, indicating that this ML model is applicable. The next steps are to compare the ML model with traditional methods and test its performance on another module.

Compared to traditional methods, the errors in machine learning are relatively smaller, as shown in Figure 9.

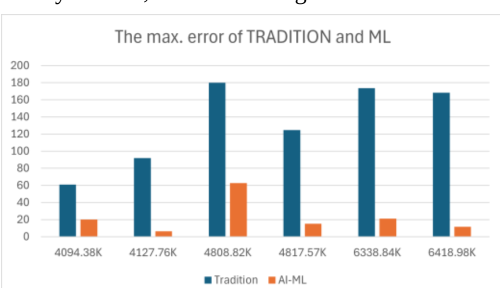

Figure 9. Different methods result

4. Experiment Result

Import the ML model obtained in step three into other in-cell modules and evaluate whether the color temperature error exceeds 200.

Figure 10 shows AdaBoost model with feature R_ratio1, B_ratio1, RWinter, R_diff, GWinter, W_raw, and BWinter in another in-cell module.

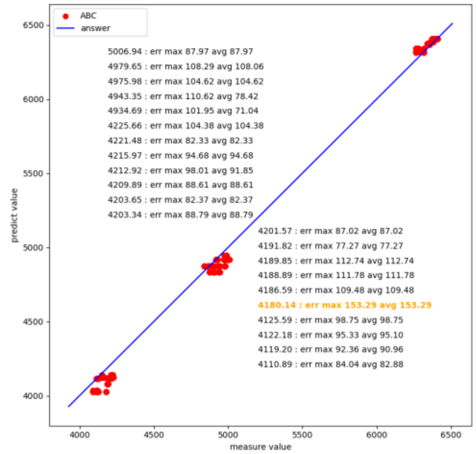

Figure 10. Other module prediction CCT with AdaBoost

Figure 11 shows GBDT model with feature B_ratio1, R_ratio2, R_ratio1, B_ratio2, B_diff, and B_train_raw in another in-cell module.

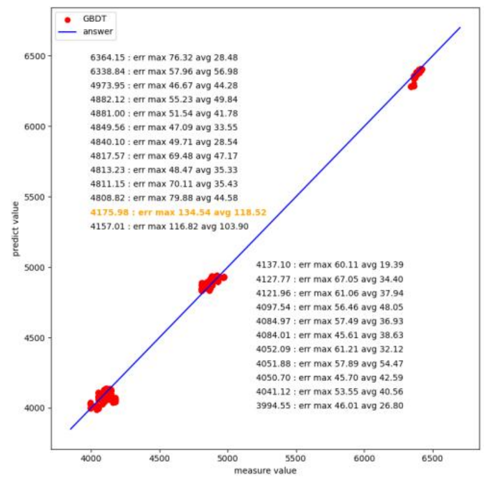

Figure 11. Other module prediction CCT with GBDT

5. Conclusion

Due to manufacturing processes or cost considerations, the in-cell ambient light/color sensor may exhibit significant variations in the data it collects, which can affect its performance. Machine learning can be used to quickly identify the correlation between the data and the target values. Furthermore, the model established on one module can be applied to other modules, enabling rapid training and scalability.

This paper presents two case studies of CCT predictions using ML on in-cell sensors. The experimental results demonstrate that the proposed method can improve accuracy compared to traditional methods.

6. References

1. Huang, C., Bu, H., Liu, S., Song, J., Niu, X., Tian, C., et al. (2022). A newly Integrated Ambient Light sensor with full in-cell structure. SID Symposium Digest of Technical Papers, 53(S1), 1039-1040
2. Ray, S. (2019). A Quick Review of Machine Learning Algorithms. 2019 International Conference on Machine

Learning Big Data Cloud and Parallel Computing (COMITCon), 35-39.

3. Cao, Y., Miao, Q.-G., Liu, J.-C., & Gao, L. (2014). Advance and prospects of AdaBoost algorithm. Acta Automatica Sinica, 39(6), 745-758.

4. Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. Annals of Statistics, 29, 1189–1232.