Object Recognition Using Color Information by Optical Machine Learning Based on Single-pixel Measurement

Kazuya Sudo, Suguru Shimomura, and Yusuke Ogura

Graduate School of Informaiton Science and Technology, The University of Osaka
E-mail: k-sudou@ist.osaka-u.ac.jp

Abstract— In this study, we propose an object recognition system by optical machine learning based on single-pixel measurement, which utilizes objects' color information. By irradiating the objects with designed light patterns and weighting the response intensity for each color using the optimized coefficients, a high output value is obtained for a target object, leading to accurate recognition. We demonstrated that the proposed system enabled to classify the CIFAR-10 dataset with higher accuracy than the conventional one through numerical simulations and optical experiments.

Keywords— *Object Recognition, Optical Machine Learning, Color Information*

I. INTRODUCTION

Machine learning systems enable to classify and predict unknown data by learning known data. To achieve accurate prediction and recognition, large-scale learning models and a lot of training data are required, which involves huge computational cost. Optical machine learning systems can perform high-speed matrix operations using spatial parallelism with low energy consumption. An optical machine learning framework based on single-pixel imaging (MLSPI) performs matrix operations at the light-speed and allows for object recognitions using incoherent light [1]. In MLSPI, a target object is recognized based on optical responses reflecting the shape of the object. If another objects' characteristic is extracted from the optical responses, the recognition accuracy can be improved.

In this study, we propose an MLSPI system using color information. By weighting intensity values for each color channel, the output depending on object's color is obtained. The proposed method utilizes information about not only the shape but also color of the object, and more accurate object recognition can be achieved. To evaluate the performance of the proposed method, we assessed the classification accuracy for color images as objects.

II. OBJECT RECOGNITION USING COLOR INFORMATION

Fig. 1 shows a schematic diagram of the MLSPI using color information. In the MLSPI, an incoherent light pattern $\boldsymbol{W}$ is irradiated to an object $\boldsymbol{X}$. The light from the object is collected by a lens to obtain it using a single-pixel detector. The obtained response intensity y is given by Equation (1).

$$y = \boldsymbol{W}\boldsymbol{X}. \tag{1}$$

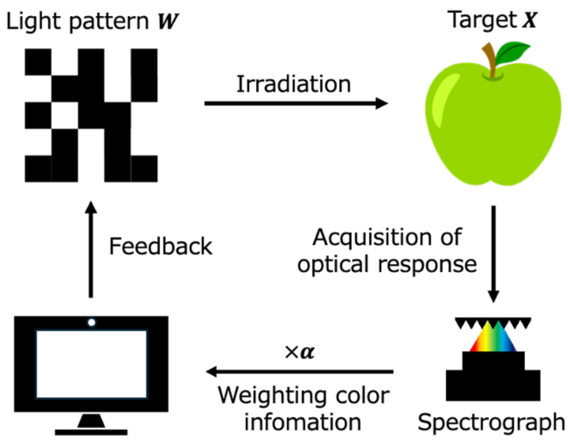

Figure 1. MLSPI using color information.

The light pattern $\boldsymbol{W}$ is optimized so that the obtained value becomes $y = 1$ when the object is the target one and $y = -1$ when it is not. The MLSPI recognizes the target object when the measured output y is high. To utilize color information, the object $\boldsymbol{X}$ is represented as a matrix $\boldsymbol{X} = [\boldsymbol{X}_R, \boldsymbol{X}_G, \boldsymbol{X}_B]^T$, based on the information in each RGB channel. We measure the optical response $[\boldsymbol{W}\boldsymbol{X}_R^T, \boldsymbol{W}\boldsymbol{X}_G^T, \boldsymbol{W}\boldsymbol{X}_B^T]$ for each channel by irradiating the object with the light pattern $\boldsymbol{W}$. The light pattern $\boldsymbol{W}$ is constructed as a grayscale pattern, and different optical response is obtained for each color channel although the object is irradiated with the same light pattern. The output $\tilde{y}$ is given by

$$\tilde{y} = \alpha_R \boldsymbol{W}\boldsymbol{X}_R^T + \alpha_G \boldsymbol{W}\boldsymbol{X}_G^T + \alpha_B \boldsymbol{W}\boldsymbol{X}_B^T, \tag{2}$$

where the coefficients $\boldsymbol{\alpha} = [\alpha_R, \alpha_G, \alpha_B]$ are the weights corresponding to the response intensities for each channel. In the training process, the light pattern $\boldsymbol{W}_i$ and the weight vector $\boldsymbol{\alpha}_i$ for class i are updated by the gradient descent so that the weighted output $\tilde{y}_i$ is the largest when the measured object belongs to class i. The update equations to optimize the light pattern $\boldsymbol{W}_i$ and weight vector $\boldsymbol{\alpha}_i$ are given by

$$\boldsymbol{W}_i = \boldsymbol{W}_i + r_1\left(T - \tilde{y}_i\right)\boldsymbol{\alpha}_i\boldsymbol{X}, \tag{3}$$

$$\boldsymbol{\alpha}_i = \boldsymbol{\alpha}_i + r_2\left(T - \tilde{y}_i\right)\boldsymbol{W}_i\boldsymbol{X}^T, \tag{4}$$

where r_1 and r_2 represent the learning rates for each update step, and $T \in \{1, -1\}$ denotes the target output corresponding to the object. From Equations (3) and (4), light patterns and weight vectors can be optimized to recognize the target object according to its shape and color characteristics.

III. NUMERICAL SIMULATION OF MLSPI USING COLOR INFORMATION

To verify that the use of color information improves recognition performance, we evaluated the classification performance through numerical simulation. We employed the CIFAR-10 dataset, which consists of color images from 10 classes. By supervised learning using 20,000 randomly selected training images, light patterns $\boldsymbol{W}_i$ and weight vectors $\boldsymbol{\alpha}_i$ to recognize class i were designed in a computer. For each test image $\boldsymbol{X}$, the weighted outputs were calculated for all class-specific light patterns $\boldsymbol{W}_i$ using the corresponding weight vectors $\boldsymbol{\alpha}_i$. After the optimization, class of 1,000 images were predicted. Fig. 2 shows the confusion matrixes representing the relationship between the predicted labels and the true labels. While the classification accuracy of the conventional MLSPI was 26.40%, the proposed method achieved 32.50%. In particular, the recognition accuracy for images in the "airplane" class (corresponding to label 0) was improved significantly from 14.89% to 56.57%. These results confirm that the use of color information can enhance recognition performance of MLSPI.

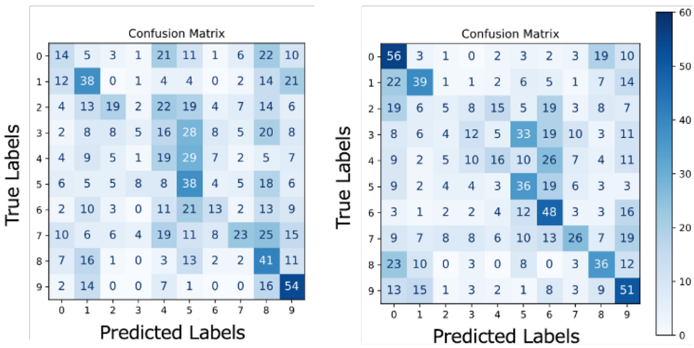

Figure 2. Confusion matrixes between the true labels and predicted labels for CIFAR-10 images in the numerical simulation with (left) conventional and (right) proposed MLSPI.

IV. OPTICAL EXPERIMENT FOR BINARY CLASSIFICATION

To evaluate the performance of the proposed method experimentally, we conducted binary classification using the CIFAR-10 dataset with an optical system. Three types of classes (cat, deer and frog) were selected, and six test images from each class were used for the prediction. Fig. 3 shows the optical system to irradiate the object with the light patterns and obtain optical responses. Among the 10 types of the light patterns designed, the light patterns corresponding to the target classes were used for the binary classification. Color images printed on a paper were irradiated with two types of optimized light patterns corresponding to the target classes. The light associated with optical response from the paper was then focused through a lens. The response intensities for each color were measured by a spectrometer. The spectral data were converted into RGB data, and the output for each light pattern were obtained by weighting the response intensity at each channel. Among these outputs, the class corresponding to the light pattern which provides the higher output value was assigned as the predicted label. Table 1 shows the relationship between the binarized classes and the classification accuracy in the MLSPI and the proposed method. In the classification between the cat and deer classes, the proposed method achieved an accuracy of 66.7%, outperforming the 41.7% accuracy of MLSPI. This result shows that utilizing color information enables more accurate object recognition compared to conventional MLSPI which relies on shape information only. On the other hand, for deer and frog classification, the accuracy of the proposed method was 66.7% while the conventional MLSPI was 75.0%. The differences are attributed to the loss of color information during the light propagation. By calibrating the optical setup and correcting spectral data, the recognition performance will be improved.

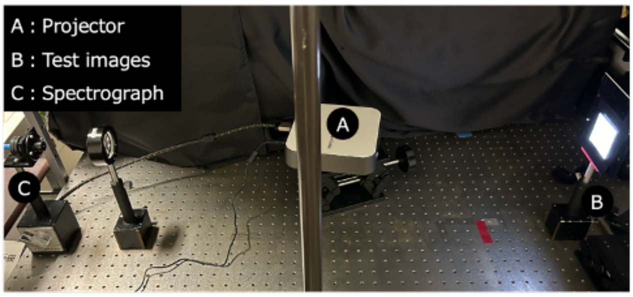

Figure 3. Optical setup for MLSPI using color information.

TABLE 1. CLASSES AND ACCURACY FOR EACH METHOD

Classified class		Accuracy (%) (MLSPI)	Accuracy (%) (Proposed method)
Cat	Deer	41.7	66.7
Deer	Frog	75.0	66.7

V. CONCLUSION

In this study, we proposed an optical machine learning system that recognizes objects by utilizing their color information. Future works include improvement of recognition performance by optimizing the illumination patterns and weights based on learning through optical experiments.

This research is supported by KAKENHI 25K21339.

REFERENCES

[1] Jiao Shuming *et al.*, "Optical machine learning with incoherent light and a single-pixel detector," Optics Letters, Vol. 44, Issue 21, pp. 5186-5189 (2019).