

# COLOR RECOGNITION FOR RUBIK'S CUBE ROBOT

Shenglan Liu, Dong Jiang, Lin Feng, Feilong Wang, Zhanbo Feng, Xiang Liu, Shuai Guo, Bingjun Li

School of Computer Science and Technology, Dalian University of Technology, Dalian 116024, China

## ABSTRACT

In this paper, we proposed three methods to solve color recognition of Rubiks cube, which includes one offline method and two online methods. Scatter balance & extreme learning machine (SB-ELM), a offline method, is proposed to illustrate the efficiency of training based method. We also point out the conception of color drifting which indicates offline methods are always ineffectiveness and can not work well in continuous change circumstance. By contrast, dynamic weight label propagation is proposed for labeling blocks color by known center blocks color of Rubiks cube. Furthermore, weak label hierarchic propagation, another online method, is also proposed for unknown all color information but only utilizes weak label of center block in color recognition. We finally design a Rubiks cube robot and construct a dataset to illustrate the efficiency and effectiveness of our online methods and to indicate the ineffectiveness of offline method by color drifting in our dataset.

**Index Terms**— Rubiks Cube, color drifting, color recognition, label propagation, cube dataset

## 1. INTRODUCTION

Rubiks cube puzzle has continually been as a hot topic in intelligence competition for child/adult. While in service robot fields, efficient solution of Rubiks cube puzzle is a challenge for computer vision. A software scheme to solve Rubiks cube puzzle includes detection, color recognition and solve method of a randomly scramble cube. Rubiks cube puzzle can be also considered as a sequential manipulation problem for service robot [1]. For example, optical time-of-flight pre-touch sensor are used for grasp Rubiks cube to achieve a high precise sequential manipulation.

In recent years, excellent research works devote in proposing quick algorithms for solving Rubiks cube puzzle. Kociemba [2] proposed close to Gods algorithm to solve the problem. Korf [3] [4] involved tree search and graph search to enhance the efficient performance of Rubiks cube puzzle. Rokicki [5] et al. proposed "Rubiks Cube Group Is Twenty", which is a ground-breaking results in Rubiks cube puzzle. The above algorithms focus on solution of cube for human or robots. However, many procedures need to be completed before the solution which is the last procedure in ser-

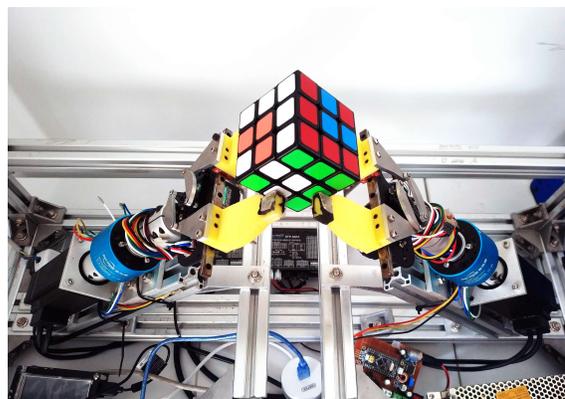


Fig. 1. Cube-solving Robot

vice robot for Rubiks cube puzzle. In cube service robot, we first detect the location of the cube by camera. Then, color recognition method need to achieve after get the color block of cube, which is one of the most important problem in the whole cube service robot procedures. The whole sequential manipulation of cooperating service robots should get correct color information of each surface in Rubiks cube. In fact, we will meet an important issue that incorrect color information make a wrong initialization for algorithm of Rubiks cube puzzle. However, a few works focus on color recognition of Rubiks cube puzzle to our best knowledge. Cezary [6] et al. introduce a full robot project of Rubiks cube including grasp the cube, color recognition, solving algorithm and mechanical arm controlling, which uses HSV color space to realize the color block distribution and illustrates the importance of color recognition in Rubiks cube solving robot. In computer vision, many image processing methods can be referred to color recognition of Rubiks cube (CRRC) problem [7] [8] [9]. Although we can get numerous images of cubes, it is hard to generalize all the situations caused by complex environment, e.g. illumination changing, cube material and abrasion by long time using. Therefore, most pre-trained models can not achieve a satisfactory performance on CRRC problem, which can be denoted as *color drifting*.

In this paper, we collect different illumination and material images and label the images using perceptual visual interactive learning (PVIL) method [10] of Rubiks cubes as a dataset (RC dataset) which contains 348 images and offers

two HSVs features. To solve CRRC problem, we introduced one model and proposed two models for CRRC problem including offline and online approaches. The offline method using supervised dimensionality reduction and classification methods calls scatter balance & extreme learning machine (SB-ELM) to enhance the robustness and considers different utilizing environment for CRRC problem. The online methods fall into two approaches as follows.

(1) Weak label hierarchic propagation (WLHP) base on weak supervised hypothesis [11] is proposed in this paper for a special problem which indicates nine color blocks on each face of Rubiks cube.

(2) Dynamic weight label propagation (DWLP), which is motivated by idea of KNN, designs known center block color of each surface on Rubiks cube. Here, we use orange and blue center block surfaces face to the camera as original status on robot arm (See Fig. 1).

The rest of this paper is organized as follows. The offline and online methods of color recognition on Rubiks cube is proposed in section 2. Our new Rubiks cube dataset for color recognition is introduced and analyzed in section 3. In section 4, the experimental results are listed and analyzed. The conclusion is the last section.

## 2. PROPOSED METHOD

### 2.1. Overview

In this section, we propose two HSVs features to represent color blocks of Rubiks cube. Based on HSV's features, one offline approach and two online approach are introduced as follows. The offline method using supervised dimensionality reduction and classification methods calls scatter balance & extreme learning machine (SB-ELM). The online methods are Weak label hierarchic propagation (WLHP) and Dynamic weight label propagation (DWLP). WLHP takes the fact that none of the six central color blocks of Rubik's cube belong to the same color. While the process of DWLP adopts specific colors of six central blocks.

### 2.2. Features

Before color recognition, it is necessary to select suitable features to represent color blocks. RGB color space is often used in computers to represent color. While compared to the RGB color space, HSV color space makes it easier to recognize colors because the HSV color space simulates human perception of color [12] [13] [14]. We propose two features built in the HSV color space in this paper, one is 3DHSV, the other one is 16DHSV.

**3DHSV:** Since the pixels in one color block are concentrated in HSV color space, we only need to find the coordinates of the pixel with the highest frequency in histogram as the feature. As Fig. 2 shows, for each dimension in HSV color space, 3DHSV calculates the histogram on the color

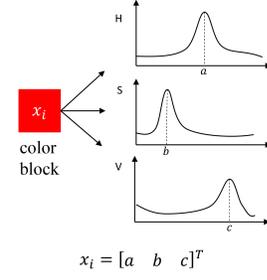


Fig. 2. Description of the 3DHSV

block and takes the value of the max bin in histogram as one dimension.

**16DHSV:** 3DHSV might ignore key information in certain situations, such as in the phenomenon of reflection. To improve 3DHSV, we design an uneven histogram feature based on the distribution of color blocks [6]. According to the color distribution characteristics, uneven histogram statistics are performed on the color blocks to obtain 16DHSV features.

### 2.3. Scatter balance & extreme learning machine

We use a novel supervised dimensionality reduction method ALDE (Angle Linear Discriminant Embedding) [15] to extract the features, with ELM (Extreme Learning Machine) [16] as the classifier. ALDE redefines the within-class scatter matrix and between-class scatter matrix by introducing CO-Angle measurement. Thus ALDE is not inclined to neither within-class scatter nor between-class scatter. ALDE can be transformed to the following optimization problem:

$$\begin{aligned} & \max_W \text{tr} [W^T M W] \\ & M = \frac{1}{d} I - S'_w + S'_b \\ & \text{s.t. } W^T W = I \end{aligned} \quad (1)$$

where  $S'_w$  and  $S'_b$  are the new defined within-class scatter matrix and between-class scatter matrix,  $I$  is the identity matrix, and  $W$  is the transformation matrix.  $S'_w$  and  $S'_b$  are showed as below:

$$\begin{aligned} S'_w &= \frac{1}{N} \sum_{i=1}^c \sum_{j=1}^{N_i} (x_{ij} - \mu_i)_e (x_{ij} - \mu_i)_e^T \\ S'_b &= \frac{1}{N} \sum_{i=1}^c N_i (\mu_i - \mu)_e (\mu_i - \mu)_e^T \end{aligned} \quad (2)$$

where  $c$  denotes the number of classes,  $N$ ,  $N_i$  denote the number of all samples and the number of samples in class  $i$ . In addition,  $\mu$ ,  $\mu_i$  denote the mean of all samples and the

mean of samples in class  $i$ . While  $e$  denotes the process of unitizing.

Furthermore, ELM is a feed-forward neural network with a hidden layer, which has extreme fast training speed and high recognition accuracy. ELM can be mathematically modeled as:

$$\begin{aligned} & \min_{\beta} \frac{1}{2} \|\beta\| + \frac{1}{2} C \sum_{i=1}^N \xi_i^2 \\ & \text{s.t. } \sum_{i=1}^{\tilde{N}} \beta_i g(a_i \cdot x_j + b_j) - t_j = \xi_j, j = 1, 2, \dots, N \end{aligned} \quad (3)$$

where  $x_j$  denote the samples,  $t_i$  denote their labels, and  $g(x)$  is the activation function. While  $a_i$  denote input weights,  $b_i$  denote bias parameters, and  $\beta_i$  are output weights. A sigmoid function is used as the activation function:

$$g(x) = \frac{1}{1 + \exp(-x)} \quad (4)$$

## 2.4. Weak label hierarchic propagation

Weak label hierarchic propagation (WLHP) takes advantage of the sociality in identifying the cube state to get data with weak labels, and the labels are used by hierarchical propagation to solve CRRC problem.

Actually, to solve CRRC problem, it only needs to know the distribution information of each cube block, so CRRC problem can be simplified. All we have to do is just distinguishing the belonging surfaces of cube blocks, without knowing their specific colors. And it is a fact that none of the six central color blocks of the cube belong to the same color. So WLHP takes the features of six central color blocks as start points of propagation, then find other color blocks that belong to the same face with the corresponding central color block by the hierarchic propagation.

The implementation of the hierarchic propagation is shown in Alg. 1, where  $X$  and  $L$  is the set of features and labels for cube color blocks respectively.  $x_i \in X, l_i \in L$ , the subscript  $i$  of  $x_i$  or  $l_i$  donates the index of cube color blocks.  $M_i$  and  $N_i$  record the intermediate results of nearest neighbor index. At the initial time,  $X$  contains color block features of a cube obtained online, there a total of 54 features. Hierarchical propagation is to deal with the situation where the feature distribution of the central color block deviates from the center of the color category.

## 2.5. Dynamic weight label propagation

Label Propagation is a semi-supervised machine learning algorithm that assigns labels to previously unlabeled data points [17]. Dynamic weight label propagation uses dynamic

---

### Algorithm 1 Weak label hierarchic propagation.

---

**Input:** The set of 54 features for cube color blocks,  $X$ ;

**Output:** The recognition results,  $L$ ;

```

1: for  $i = 0; i < 6; i ++$  do
2:    $j = 9i + 5$ ;
3:    $l_j = i$ ;
4:   remove  $x_j$  from  $X$ ;
5:   get 2 nearest neighbors index of  $x_j$  in  $X$  by KNN,
   stored in  $M_i$ ;
6:   for  $m \in M_i$  do
7:      $l_m = i$ ;
8:     remove  $x_m$  from  $X$ ;
9:   end for
10: end for
11: for  $i = 0; i < 6; i ++$  do
12:   for  $m \in M_i$  do
13:     get 3 nearest neighbors index of  $x_m$  in  $X$  by
     KNN, stored in  $N_i$ ;
14:     for  $n \in N_i$  do
15:        $l_n = i$ ;
16:       remove  $x_n$  from  $X$ ;
17:     end for
18:   end for
19: end for
20: return  $L$ ;

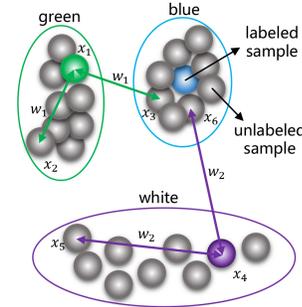
```

---

weights which base on labels. For CRRC problem, the size of the data is very small (54 features), so KNN is used as a measure of label propagation rather than constructing graph. Weight matrix is used to adjust the distance between two samples, which can be described as follows:

$$d_{jk} = \|w_i I(x_j - x_k)\| \quad (5)$$

where  $d_{jk}$  is the distance between  $x_j$  and  $x_k$ ,  $x_j$  represents the sample with the label of  $i$ ,  $x_k$  represents the sample without labels,  $w_i$  denotes the weight matrix of label  $i$ .



**Fig. 3.** Dynamic weight label propagation

As we can see in Fig. 3, during the propagation process of green labels, if we directly calculate the distance from  $x_1$  to other samples without weight matrix  $w_1$ , it is hard to choose

$x_2$  and  $x_3$ . In DWLP, distance is calculated in the form of Eq. 5, it will be easy to recognize  $x_2$  and  $x_3$ . The matrix  $w$  is dynamically determined by the label during the propagation process. For example,  $x_2$  is used in the propagation process of white label. In this way, the distribution characteristics of the data are well utilized, and a high recognition rate can be obtained.

### 3. RUBIK'S CUBE DATASET

#### 3.1. Overview

To evaluate performance of the color classification algorithm, we have collected a set of Rubik's cube images under different environmental circumstances. In this section, we first introduce the collection process of the Rubiks cube dataset, then explain how to label the images with perceptual visual interactive learning (PVIL) method [10].

#### 3.2. Samples collection

We put the cube in the hands of the robot and use the camera to take pictures in different situations. Images are collected in groups. Each group represents the state of one cube, which consists of three images. And each image in one group contains two faces of the cube, as Fig. 4(a) shows.

In order to facilitate the extraction of color features, it is necessary to separate the cube form the background, so we manually mark the six corners of the cube then use perspective transformation to convert irregular surfaces into regular squares, as Fig. 4(b) shows.

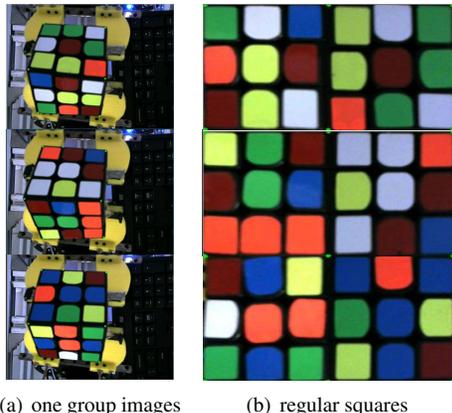


Fig. 4. Separate the cube surfaces from background

The entire dataset consists of 348 images with the same resolution ( $640 \times 480$ ), a total of 18792 color block information. The dataset contains the following 5 circumstances: cube A in bright, cube A in dark, cube B in bright, cube B illuminated from right, cube B illuminated from above. A and

B are cubes with two different backgrounds. Fig. 5 shows examples of different circumstances.

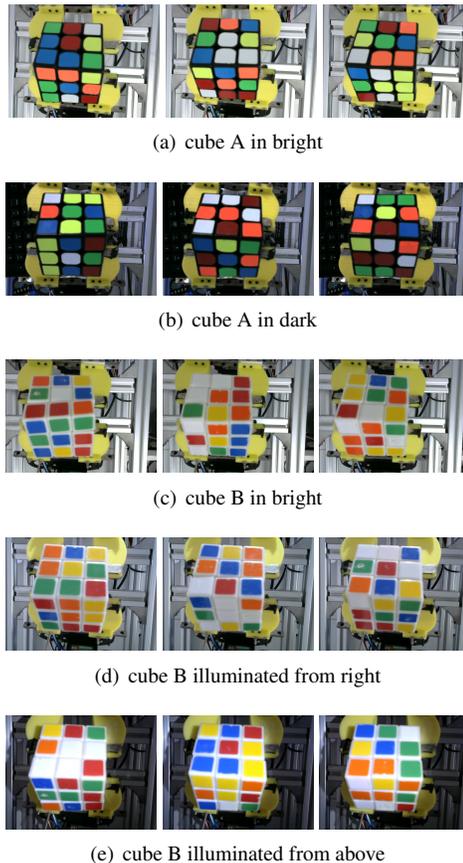


Fig. 5. Examples in Rubik's Cube Dataset

#### 3.3. Label samples

To get labels of the Rubiks cube dataset, we adopt a visual labeling method called PVIL labeling [10]. Compared with original handcraft labeling method, the PVIL method is stable in accuracy and efficient in time cost. The process of PVIL includes the following 4 steps as shown in Fig. 6. (a) mapping the origin data into a feature subspace. (b) Visualizing feature vector. (c) To judge and select the data which could be fallen into one group. (d) The labeling results is obtained.

### 4. EXPERIMENT

In this section, we evaluate proposed methods on the Rubiks Cube dataset introduced above. Methods are divided into an offline method and online methods. SB-ELM is an offline method, and 16DHSV features are used in SB-ELM. WLHP and DWLP are online methods based on 3DHSV features. In our experiment, we use A to E to represent the five circumstances in the Rubiks cube dataset. Circumstances are

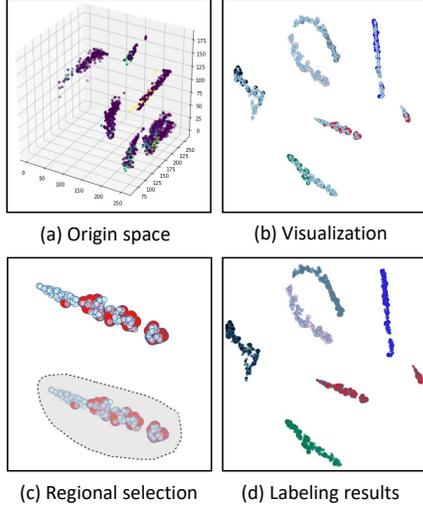


Fig. 6. The process of PVIL

arranged according to the order from A to E as follows: cube A in bright, cube A in dark, cube B in bright, cube B illuminated from right, cube B illuminated from above.

#### 4.1. Accuracy of offline method

Table 1. Accuracy of SB-ELM with 16DHSV

samples	A	B	C	D	E	total
50	85.75	96.06	82.03	78.07	76.15	84.54
100	97.76	97.88	<b>82.39</b>	79.17	76.86	84.47
150	96.94	98.88	78.52	78.64	79.04	87.02
200	96.96	98.98	82.34	80.34	78.42	87.22
250	<b>98.80</b>	<b>99.13</b>	82.34	<b>81.37</b>	<b>79.64</b>	85.03
300	97.54	94.56	81.33	80.10	79.18	<b>87.23</b>

We test the accuracy of SB-ELM in single circumstance and mutple circumstances respectively. The accuracy of SB-ELM under a single environmental circumstance is high, while the accuracy under mutple circumstances is relatively low. The reasons of the above situation can fall into two aspects: (1) It is hard to generalize all the circumstances caused by complex environment. (2) When the brightness of the light changes, most of features of orange and red in the color space will shift, which result in the overlap distribution of orange and red colors. As we can see in Tab. 1, the accuracy of SB-ELM in circumstance A and B trained by 250 samples can reach around 99%. At the same time, the accuracy in mutiple circumstances is between 84% and 88%.

#### 4.2. Accuracy of online methods

We test the accuracy of online methods in single circumstance and allover circumstances respectively. The accuracy of KNN is used as a baseline. Compared with KNN, WLHP has a higher recognition rate because it considers that the center color block feature is not distributed in the date center. To explain more specific, if KNN is used to select 8 neighbors of the center color block at a time, some of the 8 neighbors may be misidentified. While WLHP reduces the number of neighbors selected, effectively avoiding the occurrence of misjudgment. At the same time, the neighbors of neighbors are used as the remaining color blocks to ensure the correct number of identifications. It can be seen from Tab. 2 that WLHP can get better results in all circumstances than KNN.

Table 2. Accuracy of online methods with 3DHSV

Methods	A	B	C	D	E	total
KNN	95.26	95.11	85.19	79.01	77.68	86.45
WLHP	96.22	95.85	87.27	77.70	80.70	87.55
WLHP*	97.63	<b>99.85</b>	<b>99.61</b>	89.86	94.15	96.22
DWLP	<b>100.0</b>	99.26	96.84	<b>93.88</b>	<b>98.15</b>	<b>97.63</b>

The optimized WLHP is another version of online method, which is denoted as WLHP\*. The process is as follows: Find the white color block in center color blocks based on the saturation dimension of 3DHSV, then identify other white color blocks by the white center color block. When identifying other color blocks, the weight of the hue dimension is increased when calculating the distance between two features. The shift of color distribution are more sufficiently considered by increasing the weight, base on which the discrimination between colors is more obvious. The experiment results in Tab. 2 show that the accuracy of WLHP\* is 8.67% better than WLHP on average.

Inspired by WLHP\*, DWLP makes full use of color information. DWLP utilizes the distribution characteristics of the six colors in HSV color space to increase the degree of color differentiation. Color-varying weight matrix is used to calculate the distance between two features when performing different label propagation. After the above improvements, DWLP achieves a good recognition effect in the experiment. As we can see in Tab. 2, the accuracy of DWLP (97.63%) is the highest across the entire dataset.

The accuracy of online methods higher than the offline method because the relative distribution of the data is stable regardless of light changes, for example, the brightness of the orange is always higher than the red one. In CRRC problem, the experiment results has illustated that the online methods, especially DWLP, can perform better in varied environment. As shown in Tab. 1 and Tab. 2, the accuracy of SB-ELM across entire dataset is only around 87%, while the accuracy

of DWLP is around 97%.

## 5. CONCLUSION

CRRC is a very important issue in Rubiks cube robot which also can be treated as a sequential manipulation problem. Therefore, CRRC should be deep analyzed. In this paper, we point out the color drifting problem of CRRC. Furthermore, we construct a dataset to compare our proposed offline and online methods, which illustrates the effectiveness of online methods. Meanwhile, to verify our viewpoint, we design a Rubiks cube robot which works in real-world environment. The experimental results are satisfactory by most users who unknowns the operational principle of our Rubiks cube robot.

## 6. REFERENCES

- [1] W. Kasprzak, W. Szynkiweicz, and L. Czajka, "Rubiks cube reconstruction from single view for service robots," *Machine Graphics & Vision*, vol. 15, no. 3-4, pp. 451–9, 2006.
- [2] Herbert Kociemba, "Close to gods algorithm," *Cubism for Fun*, vol. 28, no. April, pp. 10–13, 1992.
- [3] Richard E Korf, "Depth-first iterative-deepening: An optimal admissible tree search," *Artificial intelligence*, vol. 27, no. 1, pp. 97–109, 1985.
- [4] Richard E Korf, "Linear-time disk-based implicit graph search," *Journal of the ACM (JACM)*, vol. 55, no. 6, pp. 26, 2008.
- [5] T. Rokicki, H. Kociemba, M. Davidson, and J. De-thridge, "The diameter of the rubik's cube group is twenty," *SIAM REVIEW*, vol. 56, no. 4, pp. 645–670, 2014.
- [6] Cezary Zielinski, Tomasz Winiarski, Wojciech Szynkiewicz, Maciej Staniak, Witold Czajewski, and Tomasz Kornuta, "Mrroc++ based controller of a dual arm robot system manipulating a rubik's cube," 12 2018.
- [7] Shenglan Liu, Jun Wu, Lin Feng, Yang Liu, Hong Qiao, Wenbo Luo, Muxin Sun, and Wei Wang, "Perceptual uniform descriptor and ranking on manifold: a bridge between image representation and ranking for image retrieval," *arXiv*, p. 14, 2016.
- [8] AM. Ferman, AM. Tekalp, and R. Mehrotra, "Robust color histogram descriptors for video segment retrieval and identification," *IEEE TRANSACTIONS ON IMAGE PROCESSING*, vol. 11, no. 5, pp. 497–508, 2002.
- [9] Gh. Liu and JY. Yang, "Content-based image retrieval using color difference histogram," *PATTERN RECOGNITION*, vol. 46, no. 1, pp. 188–198, 2013.
- [10] Shenglan Liu, Xiang Liu, Yang Liu, Lin Feng, Hong Qiao, Jian Zhou, and Yang Wang, "Perceptual visual interactive learning," *arXiv preprint arXiv:1810.10789*, 2018.
- [11] Martin Rajchl, Matthew C. H. Lee, Franklin Schrans, Alice Davidson, Jonathan Passerat-Palmbach, Giacomo Tarroni, Amir Alansary, Ozan Oktay, Bernhard Kainz, and Daniel Rueckert, "Learning under distributed weak supervision," 2016.
- [12] Ritendra Datta, Jia Li, and James Ze Wang, "Content-based image retrieval approaches and trends of the new age," in *ACM Sigm International Workshop on Multimedia Information Retrieval*, 2005, pp. 253–262.
- [13] Jing Huang, S Ravi Kumar, Mandar Mitra, Wei-Jing Zhu, and Ramin Zabih, "Image indexing using color correlograms," in *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on.* IEEE, 1997, pp. 762–768.
- [14] Liang Zheng, Shengjin Wang, Ziqiong Liu, and Qi Tian, "Packing and padding: Coupled multi-index for accurate image retrieval," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 1939–1946.
- [15] Shenglan Liu, Lin Feng, and Hong Qiao, "Scatter balance: An angle-based supervised dimensionality reduction," *IEEE transactions on neural networks and learning systems*, vol. 26, no. 2, pp. 277–289, 2015.
- [16] Guang-Bin Huang, Qin-Yu Zhu, and Chee-Kheong Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, no. 1-3, pp. 489–501, 2006.
- [17] Xiaojin Zhu and Zoubin Ghahramani, "Learning from labeled and unlabeled data with label propagation," 2002.