

Color constancy using color edge moments and regularized regression in anchored neighborhood

Wu Meng Luo Kai

(School of Marine Science and Technology, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract: To improve the accuracy of illumination estimation while maintaining a relative fast execution speed, a novel learning-based color constancy using color edge moments and regularized regression in an anchored neighborhood is proposed. First, scene images are represented by the color edge moments of various orders. Then, an iterative regression with a squared Frobenius norm (F-norm) regularizer is introduced to learn the mapping between the edge moments and illuminants in the neighborhood of the anchored sample. Illumination estimation for the test image finally becomes the nearest anchored point search followed by a matrix multiplication using the associated mapping matrix which can be precalculated and stored. Experiments on two standard image datasets show that the proposed approach significantly outperforms the state-of-the-art algorithms with a performance increase of at least 10.35% and 7.44% with regard to median angular error.

Key words: color constancy; color edge moments; anchored neighborhood; nearest neighbor

DOI: 10.3969/j.issn.1003-7985.2016.04.006

The object colors captured by regular digital cameras or videos may be shifted under varying illumination conditions. The displacement of the observed color from the genuine color usually causes problems in identifying the scene and the objects within it. In contrast, the human visual system (HVS) possesses the ability to constantly perceive the actual color of an object, regardless of the external illuminant, which is called color constancy^[1]. Removing the effect of incident light from the color-biased image is of importance for many computer vision applications, such as object tracking and object recognition. One of the solutions is to estimate the scene illuminant and then calibrate the color-biased image with the illuminant estimate.

Based on these two steps, many color constancy methods have been proposed, see Ref. [1] for recent reviews.

Received 2016-06-19.

Biography: Wu Meng (1984—), female, doctor, lecturer, wumeng@nwpu.edu.cn.

Foundation items: The National Natural Science Foundation of China (No. 61503303, 51409215), the Fundamental Research Funds for the Central Universities (No. G2015KY0102).

Citation: Wu Meng, Luo Kai. Color constancy using color edge moments and regularized regression in anchored neighborhood[J]. Journal of Southeast University (English Edition), 2016, 32(4): 426–431. DOI: 10.3969/j.issn.1003-7985.2016.04.006.

According to whether a training phase is needed, most existing color constancy methods can be classified into two groups: static methods and learning-based methods^[2].

A large proportion of static methods utilize the statistics or physical properties of images for illumination estimation. For example, the well-known grey-world algorithm^[3] estimates the illuminant by computing the mean in each color channel of image, based on the assumption that the average reflectance in the scene is achromatic. Some other grey-based methods include white-patch^[4], grey-edge^[5], etc. Stemming from these grey-based methods, moments of the zeroth-order or first-order derivatives of images are exploited for illumination estimation by using the central moments^[6] and the weighted moments^[7]. Besides, other more complex static methods include local surface reflectance statistics (LSRS)^[8] and double-opponency-based color constancy (DOCC)^[9]. Although these methods have the advantages of simple implementation and fast computation, the diverse reflectance distributions of scenes do not always accord with the actual situations^[11].

Learning-based methods use the training data to learn the model that maps the statistics or related features of an image to the corresponding illuminant. Representative examples include the Bayesian approach^[10], supported vector regression (SVR)^[11], exemplar-based method^[2], principal component analysis (PCA) correction^[12] and thin plate spline (TPS)^[13], etc. Based on an underlying analysis of the spatial characteristics in a scene, the exemplar-based color constancy^[2] combines illuminants over surfaces which are the nearest to the query surfaces. In addition, others try to find the best or a combination of simple static algorithms for each image using extracted features^[14]. Although they can produce better performance than those aforementioned static methods, these learning-based methods put more emphasis on developing sophisticated features about the surface or intricate models to estimate the illuminant, which require a high computational cost. Recently, Finlayson^[15] proposed a simple corrected moment-based approach upon grey-based methods with a matrix transform, achieving competitive performance compared with the sophisticated methods^[2]. However, there is much room for performance improvement since the mapping matrix in Ref. [15] is estimated by least squares on the whole training set.

In this paper, we propose a novel learning-based method using color edge moments and regularized regression

in an anchored neighborhood to learn a finer model for improving the accuracies without consuming too much prediction time. Motivated by the success of edge moments in earlier works^[5-7, 15], we first represent the scene images as color edge moments of various orders to characterize the intrinsic structures of images. We then learn the mapping between the edge moments and illuminant ground truth. By incorporating the concept of anchored neighborhood regression (ANR)^[16], we propose an iterative regression with a squared Frobenius-norm (F-norm) regularization to precompute the mapping matrix in the anchored neighborhood. The illumination estimation for the test image can be finally resolved by searching for its nearest anchored point, followed by the mapping of the color edge moments to the illuminant using the stored mapping matrix. Experiments verify that our method significantly outperforms existing approaches on the standard datasets.

1 Proposed Method

1.1 Color edge moments

A majority of grey-based methods that utilize the statistical moments have been unified into a single framework^[7]:

$$\left(\int \left| \nabla^n \rho_\sigma(\mathbf{x}) \right|^p d\mathbf{x} \right)^{1/p} = k e^{n,p,\sigma} \quad (1)$$

where $\rho_\sigma(\mathbf{x}) = \mathbf{I}(\mathbf{x}) \otimes \mathbf{G}_\sigma(\mathbf{x})$ is the convolution of the color image $\mathbf{I}(\mathbf{x})$ and a Gaussian filter $\mathbf{G}_\sigma(\mathbf{x})$ with the kernel parameter σ ; $\nabla^n(\cdot) = \partial^n(\cdot)/\partial \mathbf{x}^n$, n is the grey-edge order; p is the Minkowski-norm order; and k is a normalization factor. This framework actually incorporates the zeroth-order methods (e.g., grey-world and white-patch), first-order methods (i.e., grey-edge), as well as other higher-order methods.

It is proved that spatial derivatives of images (e.g., edges) are correlated with the illumination direction^[5, 7]. Meanwhile, since the statistical color moments of images provide useful information for computational color constancy, several approaches that utilize the color moments for illumination estimation have achieved competitive performance^[6, 15]. In concert with these earlier works, we represent the scene images by the higher-order color edge moments to capture the inherent structures of images, regardless of illuminant changes.

Given an image $\mathbf{I}(x, y) = \{R(x, y), G(x, y), B(x, y)\}^T$ defining RGB triplets for image position (x, y) , a spatial-domain operator $f(\cdot)$ is applied to obtain a transformed image $\mathbf{I}_f(x, y) = f(\mathbf{I}(x, y)) = \{R_f(x, y), G_f(x, y), B_f(x, y)\}^T$. Here, $f(\cdot) = \partial^n(\cdot)/\partial \mathbf{x}^n$ is the derivative operator. Following Ref. [17], we define the generalized transform moment by

$$M_{rq}^{abc} = \iint x^r y^q [R_f(x, y)]^a [G_f(x, y)]^b [B_f(x, y)]^c dx dy \quad (2)$$

where M_{rq}^{abc} is the generalized transform moment of order $r + q$ and degree $a + b + c$. Considering that the spatial location of moments provides little information for color constancy, only the generalized transform moments of M_{00}^{abc} are taken into account. It is self-evident that the left side of Eq. (1) can be regarded as a special case of M_{00}^{abc} with degree p on each color channel for the transformed image $\nabla^n \rho_\sigma(\mathbf{x})$. Similar to Ref. [15], it is believed that cross-channel effects are beneficial and moments of higher degree are involved. In this work, only generalized transform moments up to the third degree with order 0 are considered. For different edge order n of $\nabla^n \rho_\sigma(\mathbf{x})$, we can have various generalized transform moment features which we name color edge moments. Note that the feature vectors are normalized by l_2 -norm.

1.2 Mapping matrix computation using regularized regression in anchored neighbourhood

Following the learning-based methods, the problem of color constancy is formulated as the inference of the illuminant color from the edge moment feature of an image. Given a set of m training images $\{\mathbf{I}_i\}_{i=1}^m$ and the corresponding illuminants $\mathbf{E} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_m\} \in \mathbf{R}^{3 \times m}$, we calculate the edge color moments of images and acquire the feature matrix $\mathbf{X} = \{x_1, x_2, \dots, x_m\} \in \mathbf{R}^{d \times m}$, where d is the dimension of the edge moments feature. The model of illumination estimation is formulated as

$$\mathbf{E} = \mathbf{P}\mathbf{X} \quad (3)$$

where $\mathbf{P} \in \mathbf{R}^{3 \times d}$ is the mapping matrix.

Instead of the least squares estimation adopted in Ref. [15], we resort to a regularization-based regression to reduce the variance in attaining \mathbf{P} in Eq. (3). To this end, we employ a squared F-norm regularizer which makes the regression weights not to vary too much. The optimization problem becomes

$$\hat{\mathbf{P}} = \arg \min_{\mathbf{P}} \|\mathbf{P}\mathbf{X} - \mathbf{E}\|^2 + \lambda \|\mathbf{P}\|_{\text{F}}^2 \quad (4)$$

where λ is the regularization parameter. The closed-form solution of Eq. (4) is given by

$$\hat{\mathbf{P}} = \mathbf{E}\mathbf{X}^T (\mathbf{X}\mathbf{X}^T + \lambda \mathbf{I})^{-1} \quad (5)$$

Given the edge moments feature of a test image $\mathbf{y} \in \mathbf{R}^d$, the illuminant can then be estimated using \mathbf{P} on the whole training dataset,

$$\hat{\mathbf{e}} = \hat{\mathbf{P}}\mathbf{y} \quad (6)$$

Since the mapping matrix $\hat{\mathbf{P}}$ can be computed offline, it means that we only need to multiply the precomputed mapping matrix with the test feature vector to derive the illuminant. This formulation is referred as global regression (GR).

However, a global solution via the entire training set

does not guarantee that an optimal estimate for a suitable solution should enforce relevant information about the light source color and disregard irrelevant information. Since it is generally assumed that images with similar distributions share similar illumination conditions^[14], the reasonable option is to estimate the illuminant using the nearest neighbors of the test image. Instead of the entire training set, we consider the neighborhood that consists of the nearest neighbors of the test sample to compute the mapping matrix.

Although a straightforward K -nearest neighbor (KNN) can find the most similar images of the test sample, it will cost much time caused by recalculating the mapping matrix for each test sample. To overcome this, we incorporate the concept of the anchored neighborhood^[16] into Eq. (4), which can dramatically improve the execution speed during the test phase by precalculating and storing transformations in the training phase. Specifically, regarding each training sample as the anchored point, its K nearest neighbors from the training set are retrieved, forming its anchored neighborhood. Eq. (4) is then rewritten as

$$\hat{P}_i = \arg \min_{P_i} \|P_i X_i - E_i\|^2 + \lambda \|P_i\|_F^2 \quad (7)$$

where X_i and P_i are the anchored neighborhoods and the associated mapping matrix of x_i , while E_i is the corresponding illuminant matrix of X_i . The distance measure used for the nearest neighbor search is the correlation expressed as the inner product. Once the neighborhoods are defined, we can calculate a separate mapping matrix \hat{P}_i for each training sample, based on its own neighborhood rather than the entire data.

Considering the individual contribution of each sample to the objective by adding a diagonal scalar matrix D_i , Eq. (7) is reformulated as

$$\{\hat{P}_i, \hat{D}_i\} = \arg \min_{P_i, D_i} \|P_i X_i D_i - E_i\|^2 + \lambda \|P_i\|_F^2 \quad (8)$$

where D_i is a $K_i \times K_i$ diagonal matrix with positive elements, and K_i is the number of points in X_i .

The optimization in Eq. (8) is not straightforward. However, we notice that the objective function is convex if D_i or P_i is fixed. For a simpler approach, we adopt an alternative optimization procedure by minimizing the objective function with respect to one variable while fixing another. Specifically, we initialize the scalar matrix D_i to be the identity matrix and then alternate between updating the mapping matrix $P_i^{(t+1)} = E_i (X_i D_i^{(t)})^T [X_i D_i^{(t)} \cdot (X_i D_i^{(t)})^T + \lambda I]^{-1}$ while fixing D_i and updating the scalar matrix $D_i^{(t+1)} = \text{diag}((X_i P_i^{(t+1)})^\dagger E_i)$ while fixing P_i , where $(\cdot)^\dagger$ is the pseudo inverse and $\text{diag}(\cdot)$ turns the input matrix into a diagonal matrix.

At each iteration, the new values of D_i or P_i are obtained for the next iteration. Each of these steps is guaranteed to lower the bound, and iterating them alternative-

ly is guaranteed to find at least a local minimum with respect to D_i and P_i .

1.3 Illumination estimation by ANR

By incorporating the concept of the anchored neighborhood, the mapping matrix $\{\hat{P}_i\}_{i=1}^m$ associated with the anchored point can be computed offline during training stage. For the given test sample y (edge moment feature), the illumination estimation process turns into the search for the nearest neighboring training sample (anchored point) and multiplication with the associated mapping matrix \hat{P}_m ,

$$\hat{e} = \hat{P}_m y \quad (9)$$

2 Experiments

To evaluate our approach, we have conducted the experiments on two standard real-world datasets in color constancy, i. e., the ColorChecker dataset^[10] and the GreyBall dataset^[18]. Since the images within the same video clip from the GreyBall dataset are highly correlated, it is suggested that around 600 images be extracted from the full set^[10]. In our implementation, we manually select 450 images of the most representative after an even sampling of the full set in the temporal order.

For comparison, the existing methods considered are classified into two groups: 1) Static methods including grey-world^[3], white-patch^[4], grey-edge^[5], grey-CCAM (denoted grey-world-CCAM and grey-edge-CCAM)^[6], weighted grey-edge^[7], LSRS^[8], and DOCC^[9]; 2) Learning-based methods including NIS^[14], SVR^[11], PCA correction^[12], TPS^[13], and moment correction^[15].

2.1 Performance measure and experimental setup

The generally employed angular error is chosen as the error metric, since it correlates reasonably well with the human perceived quality of the color-corrected images^[1]. The angular error ε is defined as

$$\varepsilon = \cos^{-1} \left(\frac{\hat{e} \cdot e}{\|\hat{e}\| \cdot \|e\|} \right) \quad (10)$$

where \hat{e} and e are the estimated and ground-truth illuminants, respectively. Beside the commonly used measure of median angular error, we also report the measure of mean and trimean for a more comprehensive comparison. Note that the rg chromaticities of the estimated and true illuminants are converted to three-vector for angular error computation when necessary.

As to the experimental setup, we randomly divide each image dataset into three equal parts for training/validation/testing. Each experiment is repeated 10 times. For fair comparison, all the methods are tuned using the same experimental setup. For those methods which do not need tunable parameters, the validation subset is discarded. We set edge order $n = 1$. Note that we adopt edge moments feature for TPS^[13].

2.2 Experimental results

Tab. 1 lists the accuracies of different methods on the ColorChecker dataset in terms of the mean, median and trimean of angular errors. The bold values indicate the best achieved results, which is the same in the subsequent table. According to Tab. 1, we find that our (GR/ANR) approach achieves quite promising performance, which is better than all the existing methods. Specifically, the proposed (ANR) algorithm outperforms other methods with an increase of 13.63%, 10.35%, and 8.69% in terms of the mean, median and trimean of angular errors. Moreover, the learning-based methods achieve lower angular error than the static methods in terms of most criteria.

Tab. 1 Performance evaluation for various methods on the ColorChecker database

Method	Mean	Median	Trimean
Grey-world ^[3]	9.49	8.39	8.64
White-patch ^[4]	8.60	6.41	6.96
Grey-edge ^[5]	8.20	5.83	6.70
Grey-world-CCAM ^[6]	8.73	7.25	7.76
Grey-edge-CCAM ^[6]	8.38	5.77	6.72
Weighted grey-edge ^[7]	8.34	5.77	6.61
LSRS ^[8]	7.72	6.53	7.02
DOCC ^[9]	8.15	5.85	6.54
Natural image statistics (NIS) ^[14]	8.15	6.04	6.57
Support vector regression (SVR) ^[11]	7.25	5.21	5.79
PCA correction ^[12]	7.63	6.08	6.56
Thin plate spline (TPS) ^[13]	6.75	5.09	5.46
Moment correction ^[15]	6.53	4.54	4.95
Proposed GR	5.68	4.26	4.59
Proposed ANR	5.64	4.07	4.52

Tab. 2 tabulates the results of various methods on the GreyBall database in terms of the mean, median, and tri-

mean angular errors. It can be found that the proposed (ANR) method still performs better than other methods, achieving an increase of 9.00%, 7.44%, and 7.69% in the criteria of the mean, median and trimean. Also, the learning-based methods are superior to static methods.

Tab. 2 Performance evaluation for various methods on the GreyBall database

Method	Mean	Median	Trimean
Grey-world ^[3]	8.28	7.61	7.75
White-patch ^[4]	7.14	5.89	6.16
Grey-edge ^[5]	7.15	6.15	6.33
Grey-world-CCAM ^[6]	6.82	5.46	5.86
Grey-edge-CCAM ^[6]	8.11	6.84	7.12
Weighted grey-edge ^[7]	7.96	6.74	6.81
LSRS ^[8]	6.03	4.99	5.37
DOCC ^[9]	6.82	5.73	6.05
Natural image statistics (NIS) ^[14]	7.23	5.99	6.27
Support vector regression (SVR) ^[11]	5.17	4.30	4.51
PCA correction ^[12]	5.58	4.45	4.81
Thin plate spline (TPS) ^[13]	5.61	4.52	4.76
Moment correction ^[15]	5.11	3.90	4.16
Proposed GR	4.77	3.85	4.01
Proposed ANR	4.65	3.61	3.84

On both datasets, our GR method exhibits considerable improvement over moment correction, which is attributed to the proposed regularization. Meanwhile, the superiority of ANR over GR verifies the effectiveness of the anchored neighborhood.

Moreover, Fig. 1 illustrates some examples of the ColorChecker dataset corrected by the illuminant estimates of

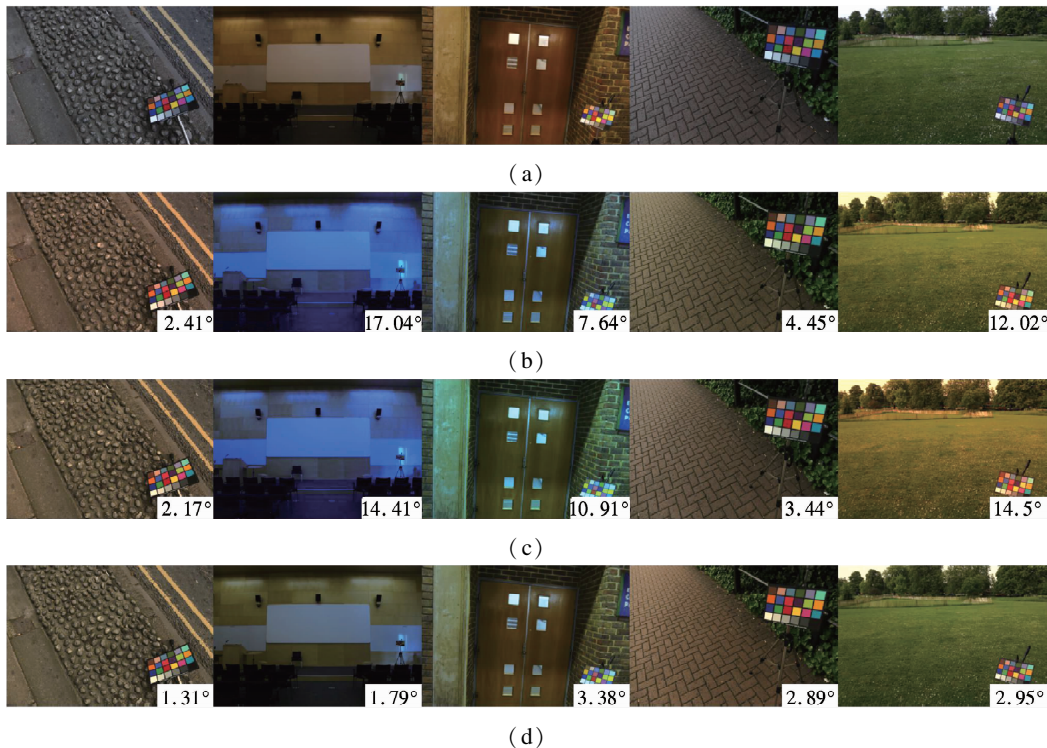


Fig. 1 Corrected images of several algorithms on the ColorChecker dataset. (a) Original images; (b) Images corrected by TPS; (c) Images corrected by moment correction; (d) Images corrected by proposed method

various methods. We can observe that the proposed method achieves visually pleasant results on these images compared with TPS and moment correction.

In addition to the precision, the computational complexity of our approach is the other issue we must consider. For the global variant of our approach, it consumes less than 1 ms in the testing phase, and its computation load comes from the multiplication of a precalculated mapping matrix and a given feature. Compared to our global version, the proposed (ANR) algorithm consumes extra time for searching the anchored point of a test sample. Since a naive idea of pairwise comparison results in $O(m)$ time complexity, this search among m anchored points costs less than 500 ms in our experiments.

3 Conclusion

This paper presents a novel learning-based color constancy method, delivering superior performance over the prevailing methods. First, we present images by statistical edge moments to characterize the intrinsic structures of scene images. Then we apply a square F-norm regularized regression in an anchored neighborhood to learn the mapping between the edge moments and illuminants. Since the mapping matrix associated with the anchored point can be computed offline, the illumination estimation process of a test image turns into the search for the nearest anchored point and multiplication with the stored matrix. The experiments show the superiority of our approach over other methods. In the future, we will explore the effects of adaptive neighborhood and anchor generation, particularly in the real scenario.

References

- [1] Gijsenij A, Gevers T, van de Weijer J. Computational color constancy: Survey and experiments [J]. *IEEE Transactions on Image Processing*, 2011, **20**(9): 2475 – 2489. DOI:10.1109/TIP.2011.2118224.
- [2] Joze H R V, Drew M S. Exemplar-based color constancy and multiple illumination [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014, **36**(5): 860 – 873. DOI:10.1109/tpami.2013.169.
- [3] Buchsbaum G. A spatial processor model for object color perception [J]. *Journal of the Franklin Institute*, 1980, **310**(1): 1 – 26. DOI:10.1016/0016-0032(80)90058-7.
- [4] Land E H. The retinex theory of color vision [J]. *Scientific American*, 1977, **237**(6): 108 – 128. DOI:10.1038/scientificamerican1277-108.
- [5] Van de Weijer J, Gevers T, Gijsenij A. Edge-based color constancy [J]. *IEEE Transactions on Image Processing*, 2007, **16**(9): 2207 – 2214. DOI:10.1109/tip.2007.901808.
- [6] Ying X, Hou L, Hou Y, et al. Canonicalized central absolute moment for edge-based color constancy [C]//*IEEE International Conference on Image Processing*. Melbourne, Australia, 2013: 2260 – 2263. DOI:10.1109/icip.2013.6738466.
- [7] Gijsenij A, Gevers T, van de Weijer J. Improving color constancy by photometric edge weighting [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, **34**(5): 918 – 929. DOI:10.1109/TPAMI.2011.197.
- [8] Gao S, Han W, Yang K, et al. Efficient color constancy with local surface reflectance statistics [C]//*European Conference on Computer Vision*. Zurich, Switzerland, 2014: 158 – 173. DOI:10.1007/978-3-319-10605-2_11.
- [9] Gao S B, Yang K F, Li C Y, et al. Color constancy using double-opponency [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, **37**(10): 1973 – 1985. DOI:10.1109/TPAMI.2015.2396053.
- [10] Gehler P V, Rother C, Blake A, et al. Bayesian color constancy revisited [C]//*IEEE International Conference on Computer Vision and Pattern Recognition*. Anchorage, Alaska, USA, 2008: 4587765-1 – 4587765-8. DOI:10.1109/cvpr.2008.4587765.
- [11] Xiong W, Funt B. Estimating illumination chromaticity via support vector regression [J]. *Journal of Imaging Science and Technology*, 2006, **50**(4): 341 – 348. DOI:10.2352/j.imsimgsci.technol.(2006)50:4(341).
- [12] Cheng D, Prasad D K, Brown M S. Illuminant estimation for color constancy: Why spatial-domain methods work and the role of the color distribution [J]. *Journal of the Optical Society of America A*, 2014, **31**(5): 1049 – 1058. DOI:10.1364/josaa.31.001049.
- [13] Shi L, Xiong W, Funt B. Illuminant estimation via thin-plate spline interpolation [J]. *Journal of the Optical Society of America A*, 2011, **28**(5): 940 – 948. DOI:10.1364/josaa.28.000940.
- [14] Gijsenij A, Gevers T. Color constancy using natural image statistics and scene statistics [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, **33**(4): 687 – 698. DOI:10.1109/TPAMI.2010.93.
- [15] Finlayson G D. Corrected-moment illuminant estimation [C]//*IEEE International Conference on Computer Vision*. Sydney, Australia, 2013: 1904 – 1911. DOI:10.1109/iccv.2013.239.
- [16] Timofte R, de Smet V, van Gool L. Anchored neighborhood regression for fast example-based super-resolution [C]//*IEEE International Conference on Computer Vision*. Sydney, Australia, 2013: 1920 – 1927. DOI:10.1109/iccv.2013.241.
- [17] Mindru F, Tuytelaars T, van Gool L, et al. Moment invariants for recognition under changing viewpoint and illumination [J]. *Computer Vision and Image Understanding*, 2004, **94**(1): 3 – 27. DOI:10.1016/j.cviu.2003.10.011.
- [18] Ciurea F, Funt B. A large image database for color constancy research [C]//*Proceedings of the Eleventh Color Imaging Conference*. Scottsdale, Arizona, USA, 2003: 160 – 164.

基于颜色边缘矩和锚定邻域正则化回归的色彩恒常算法

吴 萌 罗 凯

(西北工业大学航海学院, 西安 710072)

摘要:为提高光照估计的精度并保持相对较快的运行速度,提出了一种新的基于颜色边缘矩和锚定邻域正则化回归的色彩恒常算法. 首先提取不同阶数的颜色边缘矩作为场景图像特征. 然后,在锚定样本的邻域内采用一种迭代的平方 F-范数正则化回归来学习颜色边缘矩特征与光照间的映射矩阵. 最后,对测试样本的光照估计可基于与其最近邻的锚定样本的关联映射矩阵获得,该映射矩阵在训练阶段已被预先计算并存储. 在 2 个标准图像数据集上的实验结果表明,所提算法性能明显优于现有相关算法,其中值角度误差较现有算法至少分别下降了 10.35% 和 7.44%.

关键词:色彩恒常;颜色边缘矩;锚定邻域;最近邻

中图分类号:TP391