

Object Recognition under Various Lighting Conditions

Yasushi Makihara¹, Masao Takizawa², Yoshiaki Shirai¹, and Nobutaka Shimada¹

¹ Osaka University, 2-1, Yamadaoka, Suita, Osaka, 565-0871, Japan,
makihara@cv.mech.eng.osaka-u.ac.jp, WWW home page:
<http://www-cv.mech.eng.osaka-u.ac.jp/~makihara/index.html>
² Minato Medical Science Inc., 137, Kawanishi, Sasayama, Hyogo, 669-2735, Japan

Abstract. This paper describes object recognition under various lighting conditions for a robot which finds a user-specified object and brings it to the user. The system first constructs object models under a known lighting condition. Because the lighting condition in model construction is different from that in object recognition, the system needs to transform the model's color for the current lighting condition. The system estimates color transformation by using only one observed color of a reference object. Last the system tries to recognize a user-specified object based on the estimated lighting condition. When the system fails in recognition, it tries again by using the user's advice. Experiments using real-world refrigerator scenes are shown.

1 Introduction

This paper describes object recognition under various lighting conditions for a service robot which finds a user-specified object and bring it to the user. Nowadays, many service robots using vision are developed such as a robot which brings daily objects [1] and a robot which recognizes apples and books through verbal and gestural interaction [2]. For the vision-based service robots, adaptation to the change of lighting conditions is one of important problems.

There are a variety of researches on color constancy using surface reflectance of objects. Nadimi et al. [5] apply it to shadow detection for moving object based on the fact that shadow regions and illuminated regions have the same surface reflectance.

In addition, problems of color constancy are treated as those of color transformation for different lighting conditions. There are two representative methods for color transformation : gray world and white world. Gray world algorithm [3] assumes that the average reflectance of all the surfaces in a scene is gray. White world algorithm [4] assumes that the brightest pixel corresponds to a scene point with maximum reflectance.

Other methods use color models based on the physics-based light reflection model. Some researchers propose linear color transformation [6] and diagonal color transformation (independent transformation of each RGB channel), which are derived from the physics-based color model. Drew et al. [8] propose a voting scheme with pairs of colors of corresponding features to get candidates of diagonal color transformation.

On the other hand, there are statistics-based approaches to this problem. Miller et al. [9] propose a method of non-linear color transformation using color eigenflows learned from multiple pairs of images of a same scene under different lighting conditions. These

two methods [8], [9] need multiple pairs of reference colors for estimating color transformation. It is however difficult for the robot vision system to get multiple reference colors in unknown lighting conditions.

Our method needs only one reference color for estimating color transformation. The relation between color transformation and the reference color is formulated with the physics-based color model. In learning step, parameters for color transformation are estimated using an observed color of an object and the reference object under various lighting conditions. In object recognition, the system observes a color of the reference object and estimates color transformation with the learned parameters.

The outline of this paper is as follows. In section 2, we describe construction of object models. In section 3, we first introduce a physics-based color model and formulize relation between color transformation and the reference color, and then we describe a learning method of them. In section 4, we briefly describe automatic recognition and interactive recognition. We show experiments of color transformation using objects in a real-life refrigerator in section 5 and give a conclusion of this paper in section 6. Note that manipulation of the objects is described in [10].

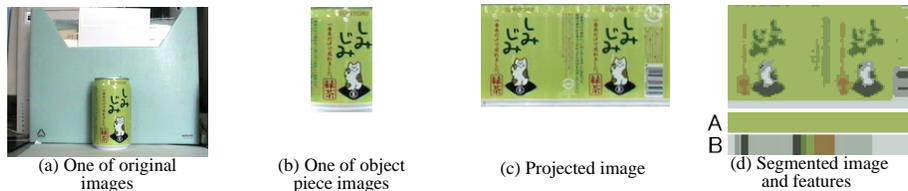
2 Construction of Object Models

2.1 Model Features and Projected Image

Features of our object models consist of the actual size of the object, representative colors, and secondary features such as the color, the position, and the size of uniform color regions other than regions of representative color. Because the object model should contain enough features to recognize it from any directions, we use an image made by projecting the surface texture of the object into a rectangle plane, which we call projected image. The projected image is constructed approximately in the following steps: (1) taking images of the object from each direction (see Fig. 1(a)), (2) extracting each object piece image (see Fig. 1(b)), and (3) merging adjacent object piece images by image mosaicing (see Fig. 1(c)).

2.2 Extraction of Model Features

The system first segments the projected image into uniform color regions [11] (see Fig. 1(d)). For each viewing direction, the system reconstructs an object piece image of the uniform color regions and extracts the color of the largest region as the representative color. Then, the system selects the region with the most different color from the representative color as the region of the secondary feature. More detail model construction is described in [12].



In (d), A: representative color, B: color of secondary feature

Fig. 1. Construction of an object model

3 Learning of Lighting Conditions

3.1 Physics-Based Color Model

According to the physics-based reflection model, it is well-known that an observed color $c^{OL} = [c_1^{OL}, c_2^{OL}, c_3^{OL}]^T = [R, G, B]^T$ of an object O under a lighting condition L is expressed as

$$c_i^{OL} = \int F_i(\lambda) S^O(\lambda) E^L(\lambda) d\lambda \quad (1)$$

where λ is wavelength of light, $F_i(\lambda)$ is spectral response of the camera to the i th component of the color, $S^O(\lambda)$ is spectral distribution of a surface reflectance of O and E^L is spectral distribution of a received light under L . In addition, there is a well-known assumption called finite-dimensional linear model [13][14]: these distributions can be written as weighted linear combinations of basis functions. We make the assumption and express these distributions as

$$F_i(\lambda) = \sum_{j=1}^{n_F} f_{ij} F_{ij}(\lambda), S^O(\lambda) = \sum_{k=1}^{n_S} s_k^O S_k(\lambda), E^L(\lambda) = \sum_{l=1}^{n_L} e_l^L E_l(\lambda) \quad (2)$$

where $F_{ij}(\lambda)$, $S_k(\lambda)$, and $E_l(\lambda)$ are basis functions for $F_i(\lambda)$, $S^O(\lambda)$, and $E^L(\lambda)$ respectively, f_{ij} , s_k^O , and e_l^L are weights for each basis function, and n_F , n_S , and n_L are the numbers of basis functions. Substituting eq. (2) to eq. (1), we obtain

$$c_i^{OL} = \sum_{k=1}^{n_S} \sum_{l=1}^{n_E} f_{ikl} s_k^O e_l^L \quad (3)$$

where $f_{ikl} = \int \sum_{j=1}^{n_F} [f_{ij} F_{ij}(\lambda)] S_k(\lambda) E_l(\lambda) d\lambda$. Moreover, rearranging eq. (3) with respect to s_k^O , eq. (3) can be expressed by matrix as

$$c^{OL} = A^L s^O \quad (4)$$

where each i th-row k th-column component of A^L is written as $A_{ik}^L = \sum_{l=1}^{n_E} f_{ikl} e_l^L$.

Next, we deal with color transformation. Let observed colors of the object O under two different lighting conditions L_1 and L_2 be denoted by c^{OL_1} and c^{OL_2} respectively, and then following two equations are obtained from eq. (4).

$$c^{OL_1} = A^{L_1} s^O, c^{OL_2} = A^{L_2} s^O \quad (5)$$

Supposing $n_S = 3$, we obtain linear color transformation from eqs. (5) as

$$c^{OL_2} = (A^{L_2})^{-1} A^{L_1} c^{OL_1} = A^{L_2 L_1} c^{OL_1}. \quad (6)$$

3.2 Color Estimation under Unknown Lighting Condition

In order to estimate the color c^{OL} of a target object O under an unknown lighting condition L in object recognition, we use a color of a reference object O_r (the refrigerator's door or frame in this case). Fig. 2 shows an outline of this problem setting. First, given the colors $c^{O_r L_r}$ of object O_r and $c^{O L_r}$ of O under a known lighting condition L_r in model construction, each surface reflectance s of them is expressed with eq. (4) as

$$s^{O_r} = (A^{L_r})^{-1} c^{O_r L_r} \quad (7)$$

$$s^O = (A^{L_r})^{-1} c^{O L_r}. \quad (8)$$

Next, given the color $c^{O_r L}$ of O_r under the unknown lighting condition L , it is written with eq. (3) as

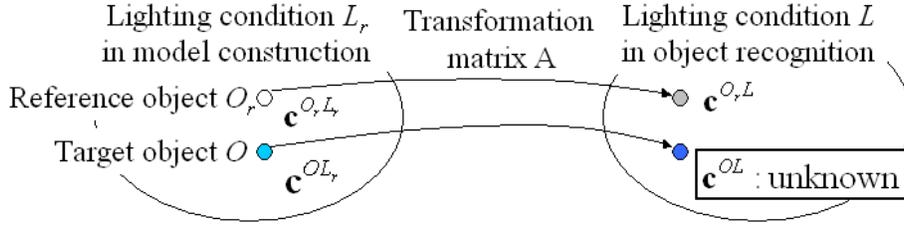


Fig. 2. Problem setting of color estimation

$$c_i^{O_r L} = \sum_k \sum_l f_{ikl} s_k^{O_r} e_l^L \quad (9)$$

Substituting s^{O_r} obtained from eq. (8) to the above equation and rearranging it, we obtain the unknown illumination e^L as

$$e^L = (B^{O_r})^{-1} c^{O_r L} \quad (10)$$

where each i th-row l th-column component of B^{O_r} is written as $B_{il}^{O_r} = \sum_k \sum_j f_{ikl} (A^{L_r})_{kj}^{-1} c_j^{O_r L_r}$. At last, we obtain $c^{O L}$ from eqs. (3), (8), and (10) as

$$c_i^{O L} = \sum_j \sum_m \alpha_{imj} c_m^{O_r L} c_j^{O L_r} \quad (11)$$

where each α_{imj} is a parameter for color transformation estimation and is equal to $\sum_k \sum_l f_{ikl} (A^{L_r})_{kj}^{-1} (B^{O_r})_{lm}^{-1}$. For simplicity, we define $c^{O L_r}$, $c^{O L}$, and $c^{O_r L}$ as c^m , c^o , and c^r respectively and express eq. (11) by the following two equations.

$$c^o = A(c^r; \alpha) c^m \quad (12)$$

$$A(c^r; \alpha)_{ij} = \sum_k \alpha_{ijk} c_k^r \quad (13)$$

The former means the color of target object is transformed linearly with transformation matrix $A(c^r; \alpha)$ and the latter means each component of the transformation matrix $A(c^r; \alpha)$ is written as a linear combination of the observed color c^r of the reference object.

Actually, the light $e^{O L}$ illuminating target objects is different from $e^{O_r L}$ illuminating the reference object in object recognition. While the reference object is illuminated only by external light, the target object is illuminated by both external light and constant internal light such as a lamp in the refrigerator. With eq. (10), the light $e^{O L}$ and $e^{O_r L}$ are expressed as follows.

$$\begin{aligned} e^{O_r L} &= e^{L_{ext}} &= (B^{O_r})^{-1} c^{O_r L} \\ e^{O L} &= e^{L_{ext}} + e^{L_{int}} &= (B^{O_r})^{-1} c^{O_r L} + e^{L_{int}} \end{aligned} \quad (14)$$

where $e^{L_{ext}}$ is the external light and $e^{L_{int}}$ is the constant internal light. Substituting $e^{O L}$ for e^L in eq. (3), the color transformation includes the constant terms. In addition, the observed colors are modified by gamma compensation and contrast compensation. Considering them, we add constant terms to the eqs. (12) and (13) for better approximation. We redefine c as $c = [c_1, c_2, c_3, 1]^T = [R, G, B, 1]^T$ and let the number of basis functions be 4. As a result, color transformation matrix A becomes a 4 by 4 matrix and includes 16 components. However, 4th-row components of A are fixed as $[0, 0, 0, 1]$ and actual components of A and α are 12 and $12 \times 4 = 48$ respectively.

Moreover, the observed color is dependent on inter reflection of objects, the shadows of objects, and the position in the refrigerator. we regard however these local effects as small and neglect them.

3.3 Learning of Parameters for color transformation estimation

In order to learn the parameters α , more than 16 pairs of c^m, c^o , and c^r are necessary . Because this color transformation includes errors in fact, the parameters should be calculated by the least square method. The error δ under the lighting conditions in object recognition and an objective function S_o to minimize are as follows.

$$\delta_{ij} = A_{ij}(c_{ij}^r; \alpha) c_{ij}^m - c_{ij}^o \quad (15)$$

$$S_o = \sum_{i=1}^N \sum_{j=1}^{n_i} \delta_{ij}^T \delta_{ij} \quad (16)$$

where N is the number of objects, n_i is the number of observation values for each object i . However, these have a trouble because weights for the errors under darker conditions are less than those under lighter conditions and color transformation to the darker conditions does not work well. In order to deal with the errors under various lighting conditions equivalently, we use the error under the lighting condition in model construction and express it as

$$\Delta = A(c^r; \alpha)^{-1} \delta = c^m - A(c^r; \alpha)^{-1} c^o. \quad (17)$$

In addition, because errors of low-reflectance objects are less than those of high-reflectance objects, we use Mahalanobis distance for normalization of the errors. At last, we obtain the parameters α by minimizing the following objective function.

$$S = \sum_{i=1}^N \sum_{j=1}^{n_i} \Delta_{ij}^T \Sigma_i^{-1} \Delta_{ij} \quad (18)$$

where Σ_i is a covariance matrix of color transformation error for each i , which is calculated as $\Sigma_i = (1/n_i) \sum_{j=1}^{n_i} \Delta_{ij} \Delta_{ij}^T$. Though eq. (18) is minimized by solving $\partial S / \partial \alpha = 0$, it is nonlinear and therefore we minimize eq. (18) with the conjugate gradient method. The initial α is given by minimization of eq. (16), which is easily solved by a linear equation $\partial S_o / \partial \alpha = 0$.

Fig. 3 shows color transformation of an object model. The original model M is transformed to the model M_1 in the darker condition L_1 and the model M_2 in the lighter condition L_2 and colors of these models are almost the same as those of the corresponding object in the refrigerator.



(black box : reference object, white circle : target objects in the refrigerator)

Fig. 3. Color transformation result

4 Object Recognition

4.1 Automatic object recognition

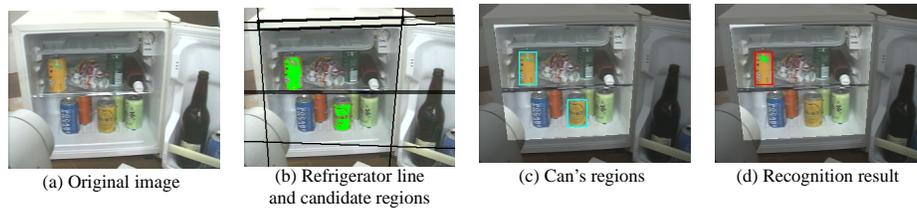
A procedure of automatic recognition of a user-specified object contains 5 steps: (1) extraction of the refrigerator, (2) color transformation of models, (3) extraction of candidate regions for objects, (4) determination of object types (can, bottle, or PET bottle), and (5) matching of the extracted regions to the model.

In step (1), the system extracts the edge lines of the refrigerator region by Hough transformation (see Fig. 4(b)). The system also calculates the average color c_r of the extracted refrigerator's door (see Fig.4(b)) and uses it for color transformation estimation in eq. (13).

Fig. 4 shows an example of object recognition from step (3) to (5). The user specifies a left-top yellow can (grapefruit juice) in this example. The system first extracts candidate regions with the representative color (see Fig. 4(b)) and determines can's regions by checking the size of regions and edges around them (see Fig. 4(c)). Though a right-bottom yellow can (lemon tea) is also extracted at this stage, the system recognizes the left-top can as the specified object by checking secondary features (see Fig. 4(d)). More detail recognition methods and results are described in [12].

4.2 Interactive object recognition

When the system fails in automatic recognition, it uses dialogues with the user via speech to obtain additional information. For example, when the color of the target object is shifted due to the change of the lighting condition, the system may not extract candidate regions (see Fig. 5(b)). Then, the system asks the user the approximate position of the object. Next, the system extends the range of the representative color by considering possible changes of lighting conditions and tries again to extract candidate regions in the given position (see Fig. 5(c)). Finally, the system recognizes the target object (see Fig. 5(d)). More detail dialogue cases and interactive recognition methods are described in [15].



(In (b), line : extracted refrigerator, bold box : reference object region, white region : candidate regions for a target object)

Fig. 4. An example of automatic object recognition

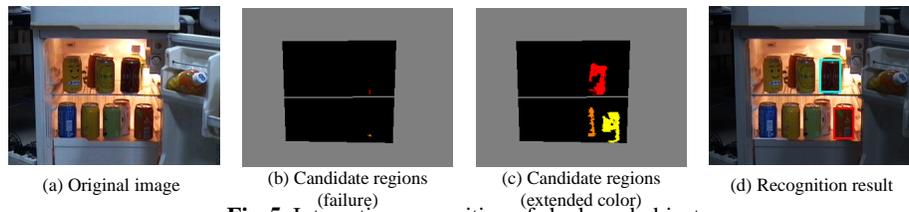


Fig. 5. Interactive recognition of shadowed objects

5 Experiments

We made an experiment on color transformation with 3 kinds of refrigerators. We use both pocket scenes and inside scenes as shown Fig. 6, so we use total 6 types of scenes in this experiment. First, we take images under various lighting conditions. Next, we randomly choose objects in the images as a learning set and assign the others as a test set. Second, for the learning set, we extract pairs of observed color c_o of a target object and c_r of the reference object (refrigerator's door or frame) and store those pairs together with the color c_m of a corresponding original object model. Next the system learns the parameters α for color transformation estimation with those color pairs. Last the system tries to extract object regions from images of the test set.

We compared our method with the diagonal transformation [7]. In the diagonal transformation, each component of the transformation matrix A is calculated only with a pair of the color c^{rm} of the reference object in model construction and its observed color c^r in object recognition as $A_{ii} = c_i^r / c_i^{rm}$.

Table 1 shows the results of the experiments. In both methods, most of the objects are extracted by automatic recognition or interactive recognition with extension of color range (see 4.2). However, the success rates of automatic recognition in our method are higher than that in the diagonal transformation. Moreover, our method works well for both white reference objects and colored reference objects such as a light green door (a) or a light yellow frame (d).

Table 1. Success rate of extracting candidate regions for each method

frige type	method	Learning sets				test sets			
		#pair	success1	success2	failure	#pair	success1	success2	failure
(a) large size	diag	320	48.7	43.5	8.8	320	45.9	40.8	13.3
	ours		94.7	5.3	0.0		92.3	7.7	0.0
(b) large size pocket	diag	219	83.1	16.9	0.0	219	82.0	18.0	0.0
	ours		95.9	4.1	0.0		93.2	6.8	0.0
(c) small size	diag	256	46.9	47.7	5.4	256	46.9	47.5	5.6
	ours		94.1	5.9	0.0		92.0	8.0	0.0
(d) small size pocket	diag	160	32.5	36.2	31.3	160	30.9	39.7	29.4
	ours		97.5	2.5	0.0		95.3	4.4	0.3
(e) middle size	diag	456	65.4	31.4	3.2	456	64.4	32.2	3.4
	ours		94.1	5.9	0.0		94.1	6.0	0.0
(f) middle size pocket	diag	140	65.0	34.3	3.2	140	69.6	30.0	0.4
	ours		95.0	5.0	0.0		90.7	8.9	0.4

Success1 and success2 indicate success rates (%) of extraction of the candidate regions in automatic and interactive object recognition respectively.



((a),(c),(e): inside scenes, (b),(d),(f): pocket scenes)

Fig. 6. Scene examples for each refrigerator (black box : reference object region)

6 Conclusion

We proposed a method for estimating color transformation in object recognition under various lighting conditions only with one observed color of the reference object based on the physics-based color model. The parameters for color transformation estimation are calculated by the minimization of sum of Mahalanobis distance for color transformation errors under the standard lighting condition in model construction so that color transformation errors of all objects under various lighting condition may be treated equivalently. We show the effectiveness of our method with experiments under various lighting conditions with some refrigerators.

Future works are as follow.

- Recognition of specular objects and transparent objects
- Implementation of our system on the mobile robot
- Experiments with the robot

References

1. Y. Takahashi et al.: Development of the mobile robot system to aid the daily life for physically handicapped. Proc. of Int. Conf. on Machine Automation 2000. (2000) 549-554
2. T. Takahashi et al.: Human-Robot Interface by verbal and Nonverbal Communication. Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems. (1998) 924-924
3. G. Buchsbaum, "A spatial processor model for object color perception", J. Franklin Inst. 310. (1980)
4. J. J. McCann, J. A. Hall, and E. H. Land, "Color mondrian experiments: The study of average spectral distributions", J. Opt. Soc. Amer., Vol. 67, No. A. (1977)
5. S. Nadimi et al.: Moving Shadow Detection Using a Physics-based Approach. Proc. of 16th Int. Conf. on Pattern Recognition. Vol. 2. (2002) 701-706
6. D. H. Marimont and B. A. Wandell, "Linear models of surface and illuminant spectra", J. Opt. Soc. Amer., Vol. 11. (1992)
7. Mark S. Drew, J. Wei, and Ze-Nian Li, "Illumination-invariant image retrieval and video segmentation", Pattern Recognition, Vol. 32. (1999) 1369-1388
8. Mark S. Drew, Ze-Nian Li, and Zinovi Tauber, "Illumination color covariant locale-based visual object retrieval", Patter Recognition, Vol. 35, No. 8. (2002) 1687-1704
9. E. G. Miller et al.: Color Eigenflows: Statistical Modeling of Joint Color Changes. Proc. of 8th Int. Conf. on Computer Vision. Vol. 1. (2001) 607-614
10. Y. Makihara et al: A Service Robot Acting by occasional Dialog -Object Recognition Using Dialog with User and Sensor-Based Manipulation-. J. of Robotics and Mechatronics, Vol. 14, No. 2. (2002) 124-132
11. Y. Shirai: Three-Dimensional Computer Vision. Springer-Verlag. (1987) 62-68
12. Y. Makihara et al: Object Recognition Supported by User Interaction for Service Robots. Proc. of 16th Int. Conf. on Pattern Recognition. Vol. 3. (2002) 561-564
13. D. B. Judd et al.: Spectral distribution of typical daylight as a function of correlated color temperature. J. Opt. Soc. Am. Vol. 54. (1964) 1031-1040
14. J. Cohen, "Dependency of the spectral reflectance curves of the Munsell color chips", Psychonomical Science, Vol. 1. (1964) 369-370
15. Y. Makihara et al: Object Recognition Supported by User Interaction for Service Robots. Proc. of 5th Asian Conf. on Computer Vision. Vol. 2. (2002) 719-724