A Composite Illumination Invariant Color Feature and Its Application to Partial Image Matching (Draft version)

Masaki KOYASHI†, Nonmember and Keisuke KAMEYAMA††, Member

SUMMARY In camera-based object recognition and classification, surface color is one of the most important characteristics. However, apparent object color may differ significantly according to the illumination and surface conditions. Such a variation can be an obstacle in utilizing color features. Geusebroek et al.’s color invariants can be a powerful tool for characterizing the object color regardless of illumination and surface conditions. In this work, we analyze the estimation process of the color invariants from RGB images, and propose a novel invariant feature of color based on the elementary invariants to meet the circular continuity residing in the mapping between colors and their invariants. Experiments show that the use of the proposed invariant in combination with luminance, contributes to improve the retrieval performances of partial object image matching under varying illumination conditions.

key words: color invariants, photometric invariants, illumination invariant color features, image matching, SIFT, RANSAC, CBIR

1. Introduction

In image matching and search used in applications such as template-based object search and content-based image retrieval (CBIR), the color of the object can be an important descriptor for finding the matching object. However, the apparent color of objects can change drastically according to illumination, surface type and observation condition.

Apparent color differences due to illumination conditions may be compensated to a limited extent by using methods such as local histogram equalization [1]. However, in order to characterize the true color of the object, it is necessary to take into account the physical process of color observation governed by the nature of the light source and the surface condition of the object.

The use of photometric invariants or color invariants has been considered for illumination-invariant object color indexing and characterization. Because the apparent color of an object is greatly determined by the physical process of light scattering at the object surface, it carries various components in addition to the true object color. The idea of photometric invariants is to estimate the components determined only by the object color, canceling out the additional components and variations by taking into account the physical process of ray observation. The causes of such additional components and variations include the differences in object surface conditions and the spectral components of illuminations [2][3][4]. Modeling of perceptual color constancy also aims to achieve a similar goal [5][6].

The photometric invariants have been used in 3D object recognition [7], correspondence, shape and motion estimation [4], and have achieved superior and robust results in comparison with the use of RGB color or grayscale.

This work is based on the color invariants introduced by Geusebroek et al. [8], which is based on Kubelka-Munk physical surface observation theory, and the Gaussian color model. The invariants derived in the work have been successfully used as image descriptors in place of grayscale for solving keypoint correspondence [9]. Among the various color-based image descriptors, it has been reported to have robust results for use in object and scene matching [10][11].

In this work, we start by analyzing how the color invariants by Geusebroek et al. are estimated from RGB colors, and find that a cyclic distance measure on the values of invariants has to be used. The use of this distance measure will reflect the nature that the colors corresponding to the negative and positive extremities of the invariant spectra are actually the same. Further, a novel color invariant feature defined as a composite function of two basic invariants will be defined. Proposal of this invariant intends to improve the selectivity of the corresponding object colors so that identical objects under different illumination will be precisely matched. The distance measure and the invariant will be used in experiments of image keypoint matching and partial image retrieval under varying illumination conditions, in comparison with the other invariants and luminance.

The remainder of the paper is structured as follows. In Sec. 2, the definition and estimation of color invariants introduced by Geusebroek et al. will be reviewed. In Sec. 3, the distance measure and the novel color invariant will be presented. In Sec. 4, the color invariants will be used as image features in partial image matching, and their advantages over conventional
luminance will be investigated.

2. Color Invariants

In this section, we will review the basic color invariants introduced by Geusebroek et al. [8], and point out an issue that may occur when they are used in color invariant-based object matching.

2.1 Color image formation and invariants

When observing the color of an object, the light source, the object, and the observer are concerned. In [8], the spectrum of reflection when the object is illuminated with a light source is modeled as,

\[
E(\lambda, x) = e(\lambda, x)(1 - \rho_f(x))^2R_\infty(\lambda, x) + e(\lambda, x)\rho_f(x),
\]  

with

- \(x\) : position within the image
- \(\lambda\) : wavelength
- \(e(\lambda, x)\) : illumination spectrum of the light
- \(\rho_f(x)\) : Fresnel reflectance at \(x\)
- \(R_\infty(\lambda, x)\) : material reflectivity (color).

The reflectivity multiplied by illumination \(e(\lambda, x)\) in Eq. (1) originates from the reflectivity derived in the Kubelka-Munk theory for objects with a layer of colorants [12]. Upon derivation, additional assumptions that, (1) the layer of colorant is sufficiently thick and the backing reflection from the object under the colorant can be ignored, and (2) \(\rho_f R_\infty \approx 0\), have been used [13][14]. By this assumption, the reflectivity in Eq. (1) has also become similar to that used in the Dichromatic Reflection Model by Shafer [15]. Please refer to [13] and [14] for discussions on their equivalence.

Considering a one-dimensional positional space and a white illumination with uniform spectral components for simplicity, \(e(\lambda, x)\) can be replaced by \(i(x)\), resulting in

\[
E(\lambda, x) = i(x)\{\rho_f(x) + (1 - \rho_f(x))^2R_\infty(\lambda, x)\}. \tag{2}
\]

Here,

\[
H = \frac{\partial E}{\partial \lambda} = \frac{E_\lambda}{E_\lambda}, \tag{3}
\]

becomes a property dependent only on the material color (wavelength-dependent reflectivity) \(R_\infty(\lambda, x)\), being independent of viewpoint, surface orientation, illumination direction, illumination intensity and Fresnel reflectance coefficient [8].

The case for matte-surfaced objects under Eq. (2) with ignorable Fresnel reflectance, \(\rho_f(x) \approx 0\) can also be assumed reducing to

\[
E(\lambda, x) = i(x)R_\infty(\lambda, x). \tag{4}
\]

Then,

\[
C = \frac{1}{E(\lambda, x)} \frac{\partial E}{\partial \lambda} = \frac{E_\lambda}{E} \tag{5}
\]

is also independent of viewpoint, surface orientation, illumination direction and illumination intensity. Color invariants, or illumination invariant color features such as \(H\) and \(C\) can be used to characterize the color of the objects under various illumination conditions. In Fig. 1, luminance \(L\) and color invariants \(H\) and \(C\) are compared for the same object under two different illumination conditions. The illumination used in the ALOI dataset [16] from which the images were selected, has a continuous and highly flat spectrum in the visible wavelength (OSRAM tungsten halogen lamp type 64637 with color temperature of 3100K). The difference in the illumination conditions upon obtaining the two images are the positions of the light sources relative to the camera. Note that the variations in \(L\) (Fig. 1(b)) is larger than those in the invariants, most apparent when compared with \(H\) (Fig. 1(c)). The luminance \(L\) was calculated according to \(L = 0.299R + 0.587G + 0.114B\)
using the \(R\)\(GB\) components throughout this work.

Further, a comprehensive set of invariants under different illumination and surface conditions, and those derived using higher-order derivatives of \(E(\lambda, x)\) have been mentioned in [8]. However, many of the invariances hold under rather limited conditions of illumination and object surface. In this work, we will only discuss the use of fundamental invariants \(H\) and \(C\) in color invariant-based matching of objects.

Table 1 summarizes the properties of the invariants \(H\) and \(C\) in comparison with luminance \(L\). Each estimated feature is displayed in grayscale together with its sensitivity to changes of illumination (shadow and intensity) and surface condition (highlight caused by Fresnel reflectance at a glossy surface).

### 2.2 The Gaussian Color Model

Upon estimating the invariants from digital images with \(R\)\(GB\) color components, the Gaussian color model [8] and its conversion to \(R\)\(GB\) basis is used.

Let \(E(\lambda)\) be the true energy of incident light to the camera at wavelength \(\lambda\). Using a probe with a spectral aperture of a Gaussian function \(G(\lambda; \lambda_0, \sigma_\lambda)\) centered at \(\lambda_0\) with scale \(\sigma_\lambda\), the observation of \(E(\lambda)\) at \(\lambda = \lambda_0\) would be,

\[
E_{\lambda_0} = \int E(\lambda)G(\lambda; \lambda_0, \sigma_\lambda) \, \text{d}\lambda.
\]  
(6)

Using a Taylor expansion at \(\lambda_0\), the observed function can be approximated as,

\[
\hat{E}_{\lambda_0 + \Delta \lambda, \sigma_\lambda} \approx \hat{E}_{\lambda_0} + \frac{\Delta \lambda}{\sigma_\lambda} \frac{\partial}{\partial \lambda} \hat{E}_{\lambda_0} + \frac{1}{2} \left( \frac{\Delta \lambda}{\sigma_\lambda} \right)^2 \frac{\partial^2}{\partial \lambda^2} \hat{E}_{\lambda_0} + \cdots
\]  
(7)

with

\[
\hat{E}_{\lambda_0} = \int E(\lambda)G_\Lambda(\lambda; \lambda_0, \sigma_\lambda) \, \text{d}\lambda
\]  
(8)

and

\[
\hat{E}_{\lambda_0} = \int E(\lambda)G_{\Lambda\Lambda}(\lambda; \lambda_0, \sigma_\lambda) \, \text{d}\lambda.
\]  
(9)

Here, \(G_\Lambda\) and \(G_{\Lambda\Lambda}\) denote the first and second order derivatives of \(G\) with respect to \(\lambda\).

Truncating Eq. (7) at the second order derivative, an approximation of the spectrum using basis \(
\{ \hat{E}, \hat{E}_\lambda, \hat{E}_{\lambda\lambda} \}\) is made. This is known as the Gaussian color model.

The Gaussian color model approximate the Hering basis[17] corresponding to the chromatic human vision at \(\lambda_0 \approx 520\text{nm}\) and \(\sigma_\lambda \approx 55\text{nm}\). Under this condition, we can approximate \(\hat{E}\), \(\hat{E}_\lambda\), and \(\hat{E}_{\lambda\lambda}\) components from the \(R\)\(GB\) components using the following linear transformation.

\[
\begin{bmatrix}
\hat{E} \\
\hat{E}_\lambda \\
\hat{E}_{\lambda\lambda}
\end{bmatrix} =
\begin{bmatrix}
0.06 & 0.63 & 0.27 \\
0.34 & 0.04 & -0.35 \\
0.34 & -0.6 & 0.17
\end{bmatrix}
\begin{bmatrix}
R \\
G \\
B
\end{bmatrix}
\]  
(10)

Eq. (10) is based on two conversions from the \(R\)\(GB\) space used in [18] to the CIE-1964 \(XYZ\), and from \(XYZ\) to the \(\hat{E}\), \(\hat{E}_\lambda\) and \(\hat{E}_{\lambda\lambda}\) of the Gaussian color model [19]. Therefore, modification of the conversion matrix will be necessary when a different \(R\)\(GB\) space is used, and/or when the \(R\)\(GB\) output of the observing camera requires calibration.

### 2.3 Invariants \(H\) and \(C\) estimated from \(R\)\(GB\) images

By using Eqs. (3) and (10), invariant \(H\) can be estimated from the \(R\)\(GB\) components of the image as,

\[
H = \frac{0.3R + 0.04G - 0.35B}{0.34R - 0.6G + 0.17B}.
\]  
(11)

Also, by using Eqs. (5) and (10), invariant \(C\) can be estimated as

\[
C = \frac{0.3R + 0.04G - 0.35B}{0.06R + 0.63G + 0.27B}.
\]  
(12)

In Fig. 1, \(R\)\(GB\) color, luminance \(L\), invariants \(H\) and \(C\) of an object under two different illumination conditions are shown. Clearly, \(R\)\(GB\) color and luminance \(L\) appears significantly different. On the contrary, the invariants show less change under these conditions.

Among the two, \(H\) which is invariant against shadow, highlight and illumination intensity gives very similar results, maintaining the edges of the red letters at the upper left and the brown diagonal stripes at the upper right. In comparison, measure \(C\) which is defined to be invariant against shadow and illumination intensity shows a larger variation. This is due to
the fact that the surface of the object is mildly glossy. However, the edges of the red characters are preserved.

Also, it is worth noting here that the invariants estimated using the RGB components are blind to metameric changes in the spectral components of light reflected from the objects.

2.4 Analysis

Here, we will look into the nature of the invariants estimated by way of Eqs. (11) and (12). By rewriting these equations in linear combinations of R, G and B, we will have

\[
(0.3 - 0.34H)R + (0.04 + 0.6H)G \\
+ (-0.35 - 0.17H)B = 0 \quad (13)
\]

and

\[
(0.3 - 0.06C)R + (0.04 - 0.63C)G \\
+ (-0.35 - 0.27C)B = 0. \quad (14)
\]

Eqs. (13) and (14) indicate that the set of points (colors) in the RGB space having the same invariant values will always be its subspace, namely a 2-dimensional plane. In the following, we will call them equal-H plane \(P_h(H)\) and equal-C plane \(P_c(C)\), respectively. A term “RGB cube” will be used for the region in the RGB space where there is a corresponding color on a framebuffer. In a 24-bit “full color” framebuffer, the RGB cube is defined as the region where R, G and B components take the value within [0, 255].

Some of the properties of the invariants \(H\) and \(C\) are listed below.

1. Both \(P_h(H)\) and \(P_c(C)\) share the origin \(O\) of the RGB space.
2. Any two colors \((R, G, B)\) and \((kR, kG, kB)\) \((k > 0)\) will have the same invariants.
3. Planes \(P_h(H \approx 0.111111)\) and \(P_c(C \approx -0.010467)\) include line \(R = G = B\) for achromatic colors.
4. \(P_h(0) = P_c(0)\).
5. The planes will rotate \(\pi\) radian as the invariants change from \(-\infty\) to \(\infty\). Therefore, \(P_h(-\infty) = P_h(\infty)\) and \(P_c(-\infty) = P_c(\infty)\).
6. The rotating axes of \(P_h(H)\) and \(P_c(C)\) are \(v_H = (0.619261, 0.518083, 0.590005)\) and \(v_C = (0.736149, -0.324631, 0.593884)\), respectively.
7. Plane \(P_h(H)\) passes through the RGB cube (as it rotates) for any \(H \in (-\infty, \infty)\). In contrast, plane \(P_c(C)\) passes through the RGB cube only for \(C\) in a certain interval. The origin \(O\), however, is always shared by the plane and the cube.

Since the values of the invariants are spanned over \((-\infty, \infty)\), often it is more convenient to use \(\theta_H = \text{arctan}(H) \in (-\frac{\pi}{2}, \frac{\pi}{2})\) and \(\theta_C = \text{arctan}(C) \in (-\frac{\pi}{2}, \frac{\pi}{2})\) as their substitutes of \(H\) and \(C\). This applies especially when the color similarity is evaluated as a function of invarient’s difference. However, it should be noted that angles \(\theta_H\) and \(\theta_C\) are not identical with the rotation angles of \(P_h(H)\) and \(P_c(C)\) from \(P_h(0)\) and \(P_c(0)\), respectively.

2.5 Issues in evaluation of color similarities using invariants

Here, two issues that arise when the invariants are used to evaluate color similarity, will be pointed out.

1. **Discontinuity of the features**

   In Fig. 2, colors corresponding to invariant values \(\theta_H, \theta_C\) are shown for six different radii from the origin for the 24-bit fullcolor RGB space. With these maps, properties 5 and 7 in Sec. 2.4 can be confirmed. From Figs. 2 (c) and (d), it is seen that \(\theta_H\) has a cyclic correspondence to the colors due to the rotation of \(P_h\). The discontinuity of \(H\) at \(H = \pm \infty\) (or \(\theta_H = \pm \pi/2\)) can be an issue when the similarity of two colors at the both sides of the gap needs to be evaluated. Such an example where gradation of colors generate a discontinuous edge of \(\theta_H\) is shown in Fig. 3.

   Invariant \(\theta_C\) also has the same relation with the RGB colors. However, it is not as significant as in \(\theta_H\) because the plane \(P_c\) does not pass through the RGB cube when \(\theta_C\) becomes large.

2. **Detection of opponent color edges using \(\theta_H\)**

   It is observed in Fig. 2(c) that colors with red-yellow hue for \(\theta_C > 0\) and those with blue-green hue \(\theta_C < 0\) share same invariant \(\theta_H\) values. Since yellow and blue are Hering’s opponent colors, it is clear that by using \(\theta_H\) alone, several important color edges cannot be detected.

3. Composite color invariant \(H'\)

   In this section, the main contribution of this paper which addresses the issues pointed out in the previous section will be presented.

   3.1 Feature discrepancy measure considering the cyclic color similarity

   When evaluating the discrepancy of the color invariant
Table 2  Set of colors having equal invariants $C$, $H$ and $H'$ for several values of invariants in $(-\pi/2, \pi/2)$

<table>
<thead>
<tr>
<th>$\theta_C$, $\theta_H$, $\theta_H'$</th>
<th>$P_c(\theta_C)$</th>
<th>$P_h(\theta_H)$</th>
<th>$P_h(\theta_H')$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(\pi/2)$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-1.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-1.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$(-\pi/2)$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

having a cyclic color correspondence, the discrepancy between two values of invariants $\theta_1$ and $\theta_2$ should be,

$$d_{cyclic}(\theta_1, \theta_2) = \min((\theta_1 - \theta_2), \{(\theta_1 + \pi) - \theta_2\}). \quad (15)$$

By introducing this measure, discrepancies in $\theta_H$ and $\theta_C$ will be appropriately evaluated reflecting the similarity of the object color.

3.2 Color invariant feature $H'$

Here, a novel color invariant $H'$ which is calculated by combining the invariants $H$ and $C$ is proposed. In Table 2, colors on equal-$C$ plane $P_c$ and equal-$H$ plane $P_h$ are shown for different values of invariants $\theta_C$ and $\theta_H$. The colored regions show the cross sections of the RGB cube as $P_c$ and $P_h$ rotate. In each map, the origin of the RGB space is shown as $O$. Vectors $v_C$ and $v_H$ are the axes of rotation of $P_h$ and $P_c$, respectively. As previously pointed out, colors of different hues (blue-orange, cyan-red, green-purple, etc) coexist on $P_h$ with a band of achromat color in between. Under an illumination of constant spectrum, it is unlikely that an object of a color in one side of this band to appear to be in the color on the other side (e.g. an orange object observed in blue). From the application point of view, objects or regions having different colors being attributed to same feature value can lead to false matching.

In order to avoid this issue, a new invariant that separates the colors at the achromatic color band is considered. On characterizing the achromatic colors, we chose to use the value of $C$, at which it takes the value of $C = C_0 \approx -0.010467$ or $\theta_C = 00\approx -0.010463$.

The angular expression of the new invariant $H'$ which addresses this issue is defined as,

$$\theta_{H'} = \begin{cases} \frac{\pi}{2} \left( \frac{\theta_C - \theta_{C0}}{\pi} \right) & (\theta_C \geq \theta_{C0}, \theta_H \geq 0) \\ \frac{\theta_H}{\pi} & (\theta_C \geq \theta_{C0}, \theta_H < 0) \\ \frac{\theta_H + \frac{\pi}{2}}{\pi} \left( \frac{\theta_C - \theta_{C0}}{\pi} \right) & (\theta_C < \theta_{C0}, \theta_H \geq 0) \\ \frac{\theta_H + \frac{\pi}{2}}{\pi} & (\theta_C < \theta_{C0}, \theta_H < 0) \end{cases} \quad (16)$$

The new invariant’s correspondence to RGB color is shown in Fig. 4, which is composed by splicing the colors in Fig. 2(c) to follow the conversion defined in Eq. (16). Also, the rightmost column of Table 2 shows the sets of colors that have identical $H'$ values.

Being a function of invariants $H$ and $C$, $H'$ has the following properties.

1. Inherits property 2 of $H$ and $C$ in Sec. 2.4.
2. Invariant to shadow and intensity as identified from Table 1.
3. Addresses issue 2 pointed out in Sec. 2.5.
4. Color correspondence similar to hue.

4. Application to color image matching and retrieval

In this section, the merit of using the color invariants in
object matching and image retrieval will be evaluated.

The color invariants $H$, $C$, $H'$ and their combinations will be used as features for local descriptors in object matching under varying illumination conditions. Matching of objects will be based on finding the correspondence of keypoints and finding the corresponding affine-transformed regions. In the process, the difference in the performances according to the selection of the features and their combinations will be investigated.

4.1 Image data

The images used in the experiments were selected from the Amsterdam Library of Object Images (ALOI) dataset[16]. ALOI is an image collection of 1000 common objects each photographed under various illumination conditions (combinations of 5 lamps arched over the observed object) and camera angles. From the dataset, we selected 50 objects whose material is annotated as “paper”. All images have 384 x 288 pixel dimensionality, with pixels having 24 bit depth of RGB color each.

4.2 Local descriptor

It is important that image structures are modeled with scale-rotation invariance and stability against noise. Therefore, we employ Scale-Invariant Feature Transform (SIFT) [20] as local descriptors of the variation of illumination invariant features. SIFT using illumination invariant features is known as Colored SIFT : (CSIFT) [9]. In CSIFT, local features can utilize color information with robustness to the illumination variation. Upon evaluating the feature similarity, the distance measure $d_{cyclic}$ in Eq. (15) was used for $H$, $C$ and $H'$.

4.3 Correspondence of local features

Let $\mathcal{P}_q = \{\hat{p}_q^1, \hat{p}_q^2, \cdots, \hat{p}_q^m\}$ be $m$ local features in the query image $I_q$, and $\mathcal{P}_d = \{\hat{p}_d^1, \hat{p}_d^2, \cdots, \hat{p}_d^n\}$ be $n$ local features in the database image $I_d$. Each feature is represented as a vector of SIFT descriptor, and the local feature $\hat{p}_d^i$ corresponding to $\hat{p}_q^j$ in query $I_q$ should meet the following nearest neighbor condition

$$\hat{p}_d^i = \arg \min_{\hat{p}_d^i \in \mathcal{P}_d} ||\hat{p}_q^j - \hat{p}_d^i||.$$  

(17)

In the following, correspondences will be evaluated with each image feature used in the experiment, and a union of corresponding point pairs will be utilized in resolving the matching regions in the two images.

4.4 Correspondence of regions

Assume that $k$ correspondences $\mathcal{C} = \{c_1, c_2, \cdots, c_k\}$ of local features $c_j = (\hat{p}^{c_j}_q, \hat{p}^{c_j}_d)$ are obtained. Here $\hat{p}^{c_j}_q, \hat{p}^{c_j}_d$ are the homogeneous coordinate representations of features in both images. The most likely affine transform (homography) between the regions including the corresponding keypoints will be estimated by RANSAC (Random Sample Consensus)[21] method.

4.5 Correlation between matching region candidates

The goodness of match for the two regions in $I_q$ and $I_d$ will be evaluated using the normalized cross-correlation (NCC) of the corresponding regions. Let $I_q', I_d$ be the transformed image of $I_q$ by an affine transformation estimated by the keypoint correspondences. The NCC between $I_q'$ and $I_d$ is calculated as

$$R_N(I_q', I_d) = \frac{\sum_D(I_q'(x) - \bar{I}_q')(I_d(x) - \bar{I}_d)}{\sqrt{\sum_D(I_q'(x) - \bar{I}_q')^2 \times \sum_D(I_d(x) - \bar{I}_d)^2}}.$$  

(18)

Here, $D$ denotes the corresponding region, and $\bar{I}_q$ and $\bar{I}_d$ are the averages of pixel values in $D$.

The conclusive correlation is defined as the average NCC of $R$, $G$ and $B$ channels.

4.6 Evaluation scores

An example of keypoint correspondence between two images is illustrated in Fig. 5. The keypoints in one of the images ($I_1$) will be labeled according to the correspondences at the other image ($I_2$). Keypoints matched
4.7.2 Procedure and evaluation

As Set 1, shown in Fig. 6. This set of images will be referred to as Set 1. An example of image variation for a single object is nation source positions and 3 different camera angles.

4.7.1 Image set

In the first experiment, the performance of solving the local correspondence of object keypoints in two images was compared.

In this experiment, the local matching using illumina-

tion condition (named "ℓ₅" in the ALOI dataset [16]) with less shade casted on the object surface was used as the query image. Seven other images under different illumination conditions ("ℓ₁" - "ℓ₇") were used as database images. These images were compared in the query-database image pairs were checked. Since the query and database images had the same object appearances except for the lighting, the correspondence was judged to be "correct" when the coordinates of the point pairs were within a radius margin of 5 pixels in both images. Upon evaluation of the discrepancy of feature values, δcycle in Eq. (15) was used. Further, the F-measure in Eq. (19) was used to evaluate the goodness of match in each image pair.

4.7.3 Results

Average F-measure of the image pairs, for cases using features L, H, C and H′, respectively, are shown in Table 3.

The results in Table 3 show that the performance using luminance L as in conventional SIFT was inferior to those when the invariants C and H′ were used. This shows the advantage of using color invariant features in matching object regions observed in color under varying illumination conditions.

However, the F-measure was relatively low when invariant H was used. This is mainly due to the nature of H that opponent colors are included in the same equal-H plane fₜ₈(θH). Due to this nature, the number of false positive and false negative keypoints increased, resulting in low performance.

Similar errors can also occur when using C, where it is apparent in Pₖ(0) of Table 2. However, illumination changes at colors of red and blue hues are matched suitably as seen in Pₖ(−0.9) and Pₖ(0.6) of the same table. Some change in the apparent hue under different illumination condition may also have been admitted as in the color set of Pₖ(−0.6).

The proposed invariant H′ significantly improved the matching F-measure by H to a level comparable to those by C. It is clear that it was beneficial to separate the opponent colors in Pₖ(H) to define the invariant H′.

4.8 Experiment 2 : Partial image matching under illumination variation

In this experiment, the local matching using illumina-

---

**Table 3** F-values of keypoint correspondences using features L, H, C and H′.

<table>
<thead>
<tr>
<th>Invariant</th>
<th>L</th>
<th>H</th>
<th>C</th>
<th>H'</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-measure</td>
<td>0.47</td>
<td>0.38</td>
<td>0.58</td>
<td>0.54</td>
</tr>
</tbody>
</table>

---

**Fig. 6** Images of an object taken under eight different illumination conditions, included in Set 1 and Set 2-Illum-Angle.

**Fig. 7** Images from Set 2-Rotation. An object taken at eight different rotation angles.
tion invariant features was further applied to partial image retrieval. The database included images of objects taken under different illumination conditions and observation angles. The query images are partial images of the objects also with variations in scaling, rotation and illumination conditions. A situation of image retrieval using a partial query image taken under an uncontrolled condition to be matched against dictionary images (in product catalogs or web pages), obtained via a controlled condition was assumed. Here, the difference in the retrieval performance according to the employed image feature set was evaluated.

4.8.1 Image set

The images of 50 objects from ALOI were used as database images. They were compiled into two sets focusing on different observation conditions.

- **Set 2-Illum-Angle**
  Images taken under different illumination positions and camera angles. Same as Set 1 with less objects. Includes 50 (objects) × 8 (illuminations) × 3 (angles) = 1200 images.

- **Set 2-Rotation**
  Images of objects rotated horizontally −45°, −30°, −15°, 0°, 15°, 30° and 45° against the camera. Includes 50 (objects) × 7 (rotation angles) = 350 images. Examples of the images for an identical object is shown in Fig. 7.

4.8.2 Procedure and evaluation

The partial query images were cut out from several database images. For each object, three query images were prepared as follows.

- **Query A**: Partial image of an object enlarged by 200%: 26 images (Fig. 8(b)).
- **Query B**: Partial image in Query A rotated 45°, in its original size: 26 images (Fig. 8(c)).
- **Query C**: Same portion as Query A from a different image which is significantly darker: 21 images (Fig. 8(d)). Five cases that the portion have become too dark to be conceived even to a human eye, were omitted.

Altogether, a total of 73 query images were used. The 26 Query A images are shown in Fig. 9.

The image feature used for keypoint matching in CSIFT and subsequent region matching were luminance $L$, invariants $H, C, H'$, luminance-invariant pairs $\{L, H\}, \{L, C\}$ and $\{L, H'\}$. Since the database includes multiple images of same objects, there are multiple correct retrievals for a query. Upon retrieval, all matched regions in the database images having positive NCC ($R_N > 0$) to the query were included in the retrievals.

For each query image, a set of images meeting either of the following criteria were selected as the correct retrieval.

- The object in the database image was identical with the one from which the query image was selected (having the same object ID).
- The image included the query image regardless of the object ID (e.g. having the same logo).

However, matching to objects of different colors were judged to be false even if the patterns or texts were identical. See the upper row of Fig. 9 for such possibilities with identical product names on packages of different colors.

For each set of retrieved images, the F-measure was calculated according to Eq. (19).
4.8.3 Results and discussions

**General tendency**

In Figs. 10 and 11, the average F-measure for the two datasets using different combinations of features are shown. The plots in the two graphs have similar tendencies, although there were some increase in mismatch of keypoints for rotated images in Set 2-Rotation. This shows that the selection of the feature combination played a major role in determining the performance, and the ranks of the performances reflected those found in Experiment 1.

**Difference by query**

For the scaled query (Query A), luminance $L$ achieved better correspondence between images of different scales, contrary to the cases where the invariants were used. It should be mentioned that when $L$ was used, objects with different colors were falsely matched.

When the rotated queries (Query B) were used, the advantage of using $L$ decreased, however, it also shows that the luminance serves well when there are less differences in the illumination between the query and database images, because the majority of the corresponding portions of the database images are illuminated similarly to the Queries A and B.

When a significantly dark query (Query C) was used, the advantage of using $L$ with $C$ or $H'$ becomes obvious. In contrast, the performance dropped significantly for $L$ and $H$. For $L$, this drop was expected as the feature is variant against the illumination change. Improvement from $H$ to $H'$ shows that confining the equal-$H'$ color set to similar hue colors was a plus.

**Difference by feature**

Joint use of $L$ and other invariants contributed to improve performance for Queries A and B. However, for Query C, using $L$ with $C$ or $H'$ seemed to have negative contribution in Set-2-Illum-Angle. Among the single features, the proposed $H'$ recorded near-constant high F-measures. Among the combined features, $L$ and $H'$ performed best.

Consequently, it became clear that the use of invariants, especially $C$ and the newly proposed $H'$ for partial image retrieval under changing illumination conditions, gave much robust results when compared with luminance $L$. As a whole, joint use of $L$ and $C$ or $H'$ gave the most stable performances for different query images. As a descriptor for color image matching, comparison with conventional features such as hue having a similar cyclic nature is due. Also, evaluations using outdoor images for real-world applications are essential.

5. Conclusion

In this work, the natures of basic illumination invariant color features $H$ and $C$ estimated from color digital images introduced by Geusebroek et al. were investigated. In order to use the features in illumination invariant object matching, a new distance measure that considers the cyclic similarity of the colors were introduced. Also, a novel invariant $H'$ as a function of $C$ and $H$ was proposed to improve the selectivity of the hue of the object color under various lighting conditions. The invariant features $H$, $C$ and $H'$ together with luminance $L$ were used in matching of partial object images by way of keypoint matching using CSIFT and region matching based on RANSAC. The experimental results showed the superiority of using the invariants in keypoint matching, and the same advantage was observed in image matching as well.

References


Masaki KOBAYASHI received the B.E. degree from the College of Information Science, University of Tsukuba, Japan in 2008. In 2010, he received the M.E. degree from the Graduate School of Systems and Information Engineering, University of Tsukuba. Since 2010, he has been working at Canon Inc. where he currently is a research staff. His interests include image processing, pattern recognition and content-based image retrieval.

Keisuke KAMEYAMA received the B.E., M.E., and Dr. Eng. degrees in electrical engineering (1989), precision machinery engineering (1991), and computer science (1999), respectively, all from Tokyo Institute of Technology, Tokyo, Japan. From 1992 to 2000 he worked as a Research Associate at Tokyo Institute of Technology. In 2000, he joined the Tsukuba Advanced Research Alliance (TARA), University of Tsukuba. Since 2007, he works as an Associate Professor at the Graduate School of Systems and Information Engineering, University of Tsukuba. His research interests include pattern recognition, signal processing, and adaptive information processing methods. Dr. Kameyama is a member of IEICE, JNNS and IEEE.
Fig. 2 Colors in RGB cube mapped to H − C space for several radii $r$ from the origin in RGB space. Each RGB component ranges from 0 to 255, and the farthest point (white) is at $r = 255\sqrt{3} \approx 441.7$. (a) $r = 10$, (b) $r = 100$, (c) $r = 255$, (d) $r = 300$, (e) $r = 350$, (f) $r = 420$. Points in H − C space having no corresponding RGB color (outside of the RGB cube) are shown either in white ((a)-(d)) or gray ((e),(f)).