Joint Use of Luminance and Color Invariants in Partial Image Retrieval

Wataru Matsumoto  
Graduate School of Systems and Information Engineering,  
University of Tsukuba  
1-1-1, Tennodai, Tsukuba, Ibaraki 305-8573, Japan  
Email: matsumoto@adapt.cs.tsukuba.ac.jp

Keisuke Kameyama  
Faculty of Engineering, Information and Systems,  
University of Tsukuba  
1-1-1, Tennodai, Tsukuba, Ibaraki 305-8573, Japan  
Email: Keisuke.Kameyama@cs.tsukuba.ac.jp

Abstract—Object color plays an important role in object image matching. Color invariants are robust features under varying illumination conditions. In this work, we analyze the estimation process of the color invariants by Geusebroek et al. from RGB images, and propose a novel invariant feature \( H' \) based on the basic invariants to meet the circular continuity residing in the mapping between colors and their invariants and improve the color selectivity. The proposed invariant \( H' \) in combination with luminance, contributes to enhance the retrieval performances of partial image matching under varying illumination conditions.

I. INTRODUCTION

In image matching, color is an important descriptor for representing the features of the object and finding the matching object. However, the apparent color of objects can change drastically according to light source, object surface and observation condition.

Geusebroek et al. introduced color invariants that are dependent of the object surface color, and these invariants can be estimated robustly regardless of the illumination conditions [1]. The invariants are based on Kubelka-Munk theory and the Gaussian color model for the measurement of the reflected light spectrum. The invariants were applied as a image descriptor in the place of luminance for finding keypoints and describing the object features [2]. Among a number of color-based image descriptors for comparisons, they are reported that CSIFT using color invariants exhibited robust results for application in object and scene image matching [3].

In this work, we analyze the property of the invariants’ estimation process from RGB colors, and propose a novel color invarient feature \( H' \) defined as a composite function of two basic invariants \( H \) and \( C \). The novel invariant intends to improve the color selectivity so that it enables a more sensitive object identification under different illumination conditions.

II. COLOR INVARIANTS

Geusebroek et al. introduced color invariants as illumination invariant features derived from a surface reflection model based on Kubelka-Munk theory [4]. Assumptions of this model is similar to those used in the Dichromatic Reflection Model by Shafer [5]. Among the color invariants, we will focus on two invariants named \( H \) and \( C \) in [1].
\[ \sigma_\lambda \approx 55\text{nm}. \] Based on this, we can calculate \( \hat{E}, \hat{E}_\lambda, \) and \( \hat{E}_{\lambda\lambda} \) approximately using the following linear transformation.

\[
\begin{pmatrix}
\hat{E} \\
\hat{E}_\lambda \\
\hat{E}_{\lambda\lambda}
\end{pmatrix} =
\begin{pmatrix}
0.06 & 0.63 & 0.27 \\
0.3 & 0.04 & -0.35 \\
0.34 & -0.6 & 0.17
\end{pmatrix}
\begin{pmatrix}
R \\
G \\
B
\end{pmatrix}
\tag{6}
\]

With the use of these approximate values, color invariances \( H \) and \( C \) can be estimated as,

\[
H = \frac{\partial E}{\partial \lambda} \approx \frac{0.3R + 0.04G - 0.35B}{0.34R - 0.6G + 0.17B},
\tag{7}
\]

and

\[
C = \frac{1}{E(\lambda)} \frac{\partial E}{\partial \lambda} \approx \frac{0.3R + 0.04G - 0.35B}{0.06R + 0.63G + 0.27B}.
\tag{8}
\]

### B. Properties of \( H \) and \( C \)

A comprehensive set of invariants under different illumination and surface conditions, and those derived using higher-order derivatives of \( E(\lambda) \) have been mentioned in [1]. In this work, however, we will only discuss the use of fundamental invariants \( H \) and \( C \) in color invariant-based matching of objects.

Invariants \( H \) and \( C \) have the following properties [7].

1. Set of points (colors) in the RGB space having the same invariant values will be a 2-dimensional subspace (plane), namely the equal-\( H \) plane \( P_h(H) \) and equal-\( C \) plane \( P_c(C) \), respectively.
2. Planes \( P_h(0.1111) \) and \( P_c(-0.0104) \) include line \( R = G = B \) for achromatic colors.
3. \( P_h(0) = P_c(0) \).
4. The planes will rotate \( \pi \) radian as the invariants change from \(-\infty\) to \( \infty \). Therefore, \( P_h(-\infty) = P_h(\infty) \) and \( P_c(-\infty) = P_c(\infty) \).
5. The rotating axes of \( P_h(H) \) and \( P_c(C) \) are \( \psi_H = (0.6193, 0.5181, 0.5900) \) and \( \psi_C = (0.7361, -0.3246, 0.5939) \), respectively.
6. Plane \( P_h(H) \) passes through the RGB cube (where colors are assigned in a framebuffer) for any \( H \in (-\infty, \infty) \). In contrast, plane \( P_c(C) \) passes through the RGB cube only for \( C \) in a certain interval.

Since the invariant values are spanned over \((-\infty, \infty)\), often it is more convenient to use \( \theta_H = \arctan(H) \in (-\frac{\pi}{2}, \frac{\pi}{2}) \) and \( \theta_C = \arctan(C) \in (-\frac{\pi}{2}, \frac{\pi}{2}) \) as their substitutes of \( H \) and \( C \).

### III. Composite Color Invariant \( H' \)

In Table I, 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 the planes rotate. In each map, the origin of the RGB space is shown as \( O \). Vectors \( \psi_c \) and \( \psi_h \) are the axes of rotation of \( P_h \) and \( P_c \), respectively. Colors of different hues (blue-orange, cyan-red, green-purple, etc) coexist on \( P_h \) with a band of achromatic color in between. From an application point of view, object 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 introduced. On characterizing the achromatic colors, we chose to use the value of \( C \), at which it takes the value of \( C = C_0 \approx -0.0105 \).

The angular expression of the new invariant \( H' \) which

<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_c(\theta_H) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>((-\pi/2))</td>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
</tr>
<tr>
<td>(-1.2)</td>
<td><img src="image4.png" alt="Image" /></td>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
</tr>
<tr>
<td>(-0.9)</td>
<td><img src="image7.png" alt="Image" /></td>
<td><img src="image8.png" alt="Image" /></td>
<td><img src="image9.png" alt="Image" /></td>
</tr>
<tr>
<td>(-0.6)</td>
<td><img src="image10.png" alt="Image" /></td>
<td><img src="image11.png" alt="Image" /></td>
<td><img src="image12.png" alt="Image" /></td>
</tr>
<tr>
<td>(-0.3)</td>
<td><img src="image13.png" alt="Image" /></td>
<td><img src="image14.png" alt="Image" /></td>
<td><img src="image15.png" alt="Image" /></td>
</tr>
<tr>
<td>0</td>
<td><img src="image16.png" alt="Image" /></td>
<td><img src="image17.png" alt="Image" /></td>
<td><img src="image18.png" alt="Image" /></td>
</tr>
<tr>
<td>0.3</td>
<td><img src="image19.png" alt="Image" /></td>
<td><img src="image20.png" alt="Image" /></td>
<td><img src="image21.png" alt="Image" /></td>
</tr>
<tr>
<td>0.6</td>
<td><img src="image22.png" alt="Image" /></td>
<td><img src="image23.png" alt="Image" /></td>
<td><img src="image24.png" alt="Image" /></td>
</tr>
<tr>
<td>0.9</td>
<td><img src="image25.png" alt="Image" /></td>
<td><img src="image26.png" alt="Image" /></td>
<td><img src="image27.png" alt="Image" /></td>
</tr>
<tr>
<td>1.2</td>
<td><img src="image28.png" alt="Image" /></td>
<td><img src="image29.png" alt="Image" /></td>
<td><img src="image30.png" alt="Image" /></td>
</tr>
<tr>
<td>((\pi/2))</td>
<td><img src="image31.png" alt="Image" /></td>
<td><img src="image32.png" alt="Image" /></td>
<td><img src="image33.png" alt="Image" /></td>
</tr>
</tbody>
</table>
addresses this issue is defined as,

\[
\theta_{H'} = \begin{cases} 
\theta_H - \frac{\pi}{2} & (\theta_C \geq \theta_{C_{0}}, \theta_H \geq 0) \\
\theta_H & (\theta_C \geq \theta_{C_{0}}, \theta_H < 0) \\
\theta_H + \frac{\pi}{2} & (\theta_C < \theta_{C_{0}}, \theta_H \geq 0) \\
\theta_H & (\theta_C < \theta_{C_{0}}, \theta_H < 0)
\end{cases}
\] (9)

The rightmost column of Table I shows the sets of colors that have identical \(H'\) values. The \(H'\) being a function of \(H\) and \(C\), it is invariant to shadows and changes of illumination intensity. Most importantly, it has a color correspondence similar to hue and solves the issue of \(H\) and \(C\) pointed out above.

Here, for comparative purpose, hue is introduced and defined as

\[
\theta_{Hue} = \arctan \left( \frac{\sqrt{3}(G - B)}{2R - G - B} \right). \quad (10)
\]

Hue is invariant to shadows, highlights and changes of illumination intensity[8]. The rotating axis of equal-\(Hue\) plane is \(v_{Hue} = \left( \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right)\). Fig.1 shows the images of the hue circle converted to color invariant features \(H, H'\) and \(Hue\). There is a similarity that each value of feature are changed by hue color transition, because these values cycle around the rotation axes. The position of the rotation axes of hue is slightly different from those of \(H\) and \(H'\).

IV. IMAGE MATCHING AND RETRIEVAL EXPERIMENT

The Luminance \(L\), \(Hue\), color invariants \(H, C, H'\) and their combinations will be used as features for local descriptors in object matching under varying illumination conditions. Here, the difference in the keypoints correspondence and the retrieval performance according to the employed image feature set were evaluated.

A. Experimental setup

1) Image dataset: In all experiments, we used the Amsterdam Library of Object Images (ALOI) dataset[9], which is a collection of 1000 common objects images having various color and texture observed from different angles under various camera and illumination positions. It also includes images observed under types of different illuminations. However, only images illuminated by a tungsten halogen lamp with near-flat continuous spectrum have been used in this work. All images have \(384 \times 288\) pixel dimensionality, with pixels having 24 bit depth of \(RGB\) color.

2) Local descriptor: Scale-Invariant Feature Transform (SIFT) [10] using illumination invariant features known as Colored SIFT : (CSIFT) [2] was used as local descriptors.

3) Distance measure: In evaluating the feature discrepancy between two values \(\theta_1\) and \(\theta_2\) \((-\pi/2 \leq \theta_2 \leq \theta_1 < \pi/2)\) of an invariant or hue, a cyclic distance measure

\[
d_{cyclic}(\theta_1, \theta_2) = \min((\theta_1 - \theta_2), \{\theta_2 + \pi - \theta_1\}) \quad (11)
\]

was used.

4) Correspondence of features and regions: Matching of CSIFT descriptor was based on the nearest neighbor search in the feature space. A union of corresponding point pairs were used in resolving the matching regions in the two images. The pairs were further interpreted as matched regions by finding the homography between the regions by RANdom SAMple Consensus (RANSAC)[11]. The goodness of match for the two regions were evaluated using the normalized cross-correlation (NCC) of the corresponding regions.

5) Evaluation: For evaluating the correspondences of two images, the F-measure defined as

\[
J_F = \frac{2N_{TP}}{N_{TP} + N_{FP} + (N_{TP} + N_{FN})} \quad (12)
\]

is used. Here, \(N_{TP}, N_{FP}\) and \(N_{FN}\) denote the numbers of image matches judged as true positive (TP), false positive (FP) and false negative (FN), respectively.

B. Experiment 1: Robustness of keypoint correspondences

Here, we evaluated the robustness of keypoint correspondences by using color invariant features as image descriptors under different illumination conditions.

1) Image set: The images used in this experiment were 156 object images selected from the ALOI dataset whose material is "paper". Each object has 24 images of 8 illumination directions and 3 camera angles. Therefore, there are \(156 \times 24 = 3744\) images in the dataset. In Fig. 2, there are examples of images from 8 different illumination conditions.

2) Evaluation method and procedure: We selected one image of particular illumination direction \(l_8\) in Fig. 2, and evaluated the keypoint correspondences to the other 7 images. First, the images were converted into values of invariant features. Second, keypoints were extracted using CSIFT. Then, corresponding keypoints were searched employing a nearest neighbor search. Here, if an error in the coordinates of the corresponding keypoint is less than 5 pixels, then it is treated as a correct match.
Hue invariants of keypoints between well illuminated (it appears that there is a significant difference in number
Fig. 3. Differences in the number of same object’s keypoints under various illumination positions. F-measure is the average of 156 keypoint estimation applying color invariants under different illuminations.

C. Experiment 2: Partial image matching

3) Results: Table II shows the result of corresponding keypoint estimation applying color invariants under different illumination positions. F-measure is the average of 156 "paper" objects. Under such a condition of various illuminations, invariants C and H' are superior in stability than L. Owing to the improvement of color selectivity, H' shows higher performance than H. Hue is inferior to H' in stability in spite of the similar properties. Because this phenomenon was unexpected, a further research is needed.

Fig. 3 shows differences in the number of same object’s keypoints under 8 illumination conditions. These numbers are averages of 156 "paper" object images. In luminance L, it appears that there is a significant difference in number of keypoints between well illuminated (l3, l4, l5, l7, l8) and faintly illuminated (l1, l2, l6) images. Therefore, luminance is unstable under such illuminant conditions. On the other hand, although there are slight differences, other features (H, C, H', and Hue) are relatively stable. In C, numbers of keypoints are few. This is believed to be due to the effect of limited range in θC because of colors that do not exist in the RGB color space (see Table I). Numbers of keypoints in Hue are relatively small when compared with H'. This could be the reason for the low performance of matching.

<table>
<thead>
<tr>
<th>Feature</th>
<th>L</th>
<th>H</th>
<th>C</th>
<th>H'</th>
<th>Hue</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-measure</td>
<td>0.470</td>
<td>0.391</td>
<td>0.585</td>
<td>0.541</td>
<td>0.394</td>
</tr>
</tbody>
</table>

Fig. 4. (a) Examples of Rotation images. Query images: (b) Original object image. (c) Query A from the framed area of (b), (d) Query B (rotated Query A), and (e) Query C (darker Query A)

Table II. F-MEASURE OF KEYPOINT CORRESPONDENCES USING FEATURES L, H, C, H' AND Hue

C. Experiment 2: Partial image matching

In experiment 2, we apply the invariant features H, C and H' in partial image retrieval. Matching will be based on finding the correspondence of keypoints in the two images and finding the corresponding affine-transformed regions.

1) Image set: The images used in the experiments were 50 object images selected from 156 objects whose material is "paper" in the ALOI dataset. The images were compiled into two sets focusing on different observation conditions.

- **Set Illum-Angle**: Images taken under different illumination positions and camera angles. Includes 50 (objects) × 8 (illuminations) × 3 (angles) = 1200 images.

- **Set 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. 4(a).

2) Query image: The partial query images were cut out from several database images. Altogether, a total of 73 query images were used. For each object, three query images were prepared as follows.

- **Query A**: Partial image of an object enlarged by 200% : 26 images (Fig. 4(c)).
- **Query B**: Partial image in Query A rotated 45°, in its original size : 26 images (Fig. 4(d)).
- **Query C**: Same portion as Query A from a different image which is significantly darker : 21 images (Fig. 4(e)).

The set of Query A images is shown in Fig. 5.
3) **Image feature**: Features used for keypoint matching in CSIFT and region matching were luminance \( L \), \( Hue \), invariants \( H \), \( C \), \( H' \), pairs \( \{ L, H \} \), \( \{ L, C \} \), \( \{ L, H' \} \) and \( \{ L, Hue \} \). Upon retrieval, all matched regions in the database images having positive NCC \( (R_N > 0) \) to the query were included in the retrievals.

4) **A correct retrieval**: 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.

5) **Results**: F-measure for different types of queries and combinations of features are shown in Fig. 6.

In the scaled (Query A) and rotated (Query B) queries, luminance \( L \) achieved better correspondence than the invariants, probably because there was sufficient illumination to obtain a number of keypoints enough to determine the correspondence between the query and database images. When dark queries (Query C) were used, the advantage of using invariants \( C \) and \( H' \) became obvious. In contrast, the performance dropped significantly for \( L \) and \( H \). Improvement from \( H \) to \( H' \) shows that confining the equal-\( H' \) color set to similar hue colors was a plus.

Except for Query C in Illum-Angle where joint use of \( L \) with \( C \) or \( H' \) had a slightly negative contribution, joint use of \( L \) and other invariants contributed to improve performances.

Performances when \( Hue \) was used as the descriptor was generally low, which was unexpected. Although further investigation is necessary, it can be considered to be due to its high variance against different illumination conditions in the dataset.

The performances showed similar tendencies for both datasets of Illum-Angle and Rotation. However, some increase in mismatch of keypoints were observed for rotated database images in Set Rotation.

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. The use of invariant \( H' \) is recommended in dark lighting conditions especially.

V. Discussion

The reason for the superior performances of color invariants \( C \) and \( H' \) may be explained by the actual invariance of the descriptors under varying illumination conditions. In Fig. 7, the variances of \( L \), \( H \), \( C \), \( H' \) and \( Hue \) are compared through the 8 different lighting conditions. The variances are avarges of all keypoints in selected 7 different color objects images included in the image set Illum-Angle and calculated for each descriptor after normalizing their ranges to unity.

It is clear that the stability of the descriptor directly contributes to the image matching performances under varying illumination conditions in Fig. 6. The \( Hue \) is supposed to be invariant under an ideal white illumination, however, results in Fig. 7 show that this is not the case for the images used in the experiment.

It is confirmed that most cases using combination of \( L \) and \( H' \) improved retrieval performance in Fig. 6. The results of a particular image in Fig. 8 shows comparison when \( L \), \( H' \) and \( L + H' \) features were used in two retrievals.
Fig. 7. Variances of $L$, $H$, $C$, $H'$ and $\text{Hue}$ under different illumination angles.

Fig. 8. Comparison of F-measures when $L$, $H'$, and $L+H'$ features were used for two retrievals. (a) Search result when $H'$ is used to supplement $L$. (b) Search result when $L$ works to supplement $H'$. The images are examples of "correct" retrieval image and query, respectively.

It appears that there is a good matching respectively (Fig. 8 (a), (b)), so it is believed to be related to good performance when $L+H'$ combination was used. Investigation and tuning of invariant measures for specific observation conditions should be beneficial for real-world implementation of illumination invariant object recognition.

VI. Conclusion

In this work, we proposed a newly color invariant $H'$ based on color invariants $C$ and $H$ to enhance the selectivity of object color under varying illumination conditions. Luminance, hue and different color invariants were compared as image descriptors in the stage of keypoints correspondences using CSIFT and region matching of partial object images based on RANSAC. Experimental results showed the advantage of joint use of luminance and $C$ or $H'$ in partial image matching under varying illumination conditions.

REFERENCES