π-SIFT: A Photometric and Scale Invariant Feature Transform

Jae-Han Park, Kyung-Wook Park, Seung-Ho Baeg, and Moon-Hong Baeg
Korea Institute of Industrial Technology (KITECH)
hans1024, kwpark, shbaeg, and mhbaeg@kitech.re.kr

Abstract

For many years, various local descriptors that are insensitive to geometric changes such as viewpoint, rotation, and scale changes, have been attracting attention due to their promising performance. However, most existing local descriptors including the SIFT (Scale Invariant Feature Transform) are based on luminance information rather than color information thereby resulting in instability to photometric variations such as shadows, highlights, and illumination changes. In this paper, we propose a novel local descriptor, π-SIFT, that are invariant to both geometric and photometric variations. In order to achieve photometric invariance, we adopt photometric quasi-invariant features based on the dichromatic reflection model. The performance of the proposed descriptor is evaluated with SIFT.

1. Introduction

In computer vision, the need for a stable local descriptor that is robust to geometric variations such as viewpoint, scaling, and affine transformation has captured the attention of researchers for years. Intensive research efforts have resulted in many robust local descriptors that provide distinctiveness as well as robustness [1, 2]. However, most of the existing local descriptors are based on gray-level images paying little attention to color information.

Color has been investigated for a long time because of its excellent discriminating capability compared to gray-level images and various color models have been introduced. For instance, opponent color space has the characteristic which is invariant to changes in illumination intensity and shadows in addition to isolates the brightness information from RGB color space. Besides, HSV color space is often employed to obtain photometric invariance since the hue is invariant under the orientation of the object with respect to the light source and viewing directions.

In order to get more reliable features, a local descriptor needs to deal with the invariance with respect to imaging conditions including geometric and photometric variations.

In this paper, we propose a novel Photometric and Scale Invariant Feature Transform (PI-SIFT) describing features that are both invariant to geometric and photometric variations. In order to induce photometric quasi-invariant features, we first use the dichromatic reflection model [5] which describes the light reflected at the material surface and the light reflected from the material body. The spatial derivative of this model, which gives the photometric derivative structure of the image, links differential-based features such as edge and corner to the theory of photometric invariance. Next, in order to obtain the features that are invariant to geometric variations such as translation, rotation, and scaling, we build scale-spaces based on the photometric quasi-invariant features. Finally, The similar strategy of SIFT [1] is used to build descriptors.

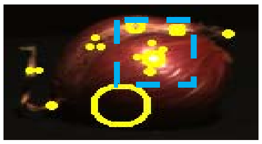
2. Related Work

In recent years, A. E. Abdel-Hakim et al. [3] proposed a novel method (called CSIFT) that aims at not only embedding the color information in the descriptor, but also giving the robustness with respect to both photometric and geometrical changes. Especially, they used color invariance approach, proposed by J. M. Geusebroek et al. [7], to achieve photometric invariance. Even though color invariance method provides a set of photometric invariant derivative filters, the nonlinear transformations for computing photometric invariants have several drawbacks such as instability and loss of discriminative power.

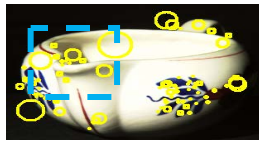
Considering these issues, we focus on describing the robust features without additional dimensions as

978-1-4244-2175-6/08/$25.00 ©2008 IEEE

(a) SIFT features under different illumination directions from left to right.



(b) Highlights



(c) Shadow/shading

Figure 1: Unsuitable SIFT features caused by photometric variations.

building scale-spaces using photometric quasi-invariant features.

3. Photometric Quasi-Invariant Features

We use photometric quasi-invariant features, which are proposed by Joost van de Weijer et al. [4], to detect interest points that are invariant to photometric variations.

3.1. Problem Statement

Figure 1 shows the effects of photometric variations on SIFT descriptor. In Figure 1(a), we can easily see that the number of SIFT features increase or decreases depending on different illumination directions. The blue rectangle areas in Figure 1(b) and 1(c) represent the interest points extracted by highlights and shadow/shading reflected from object's surface. Particularly, since these effects may be continuously changed according to surface geometry variations such as the light source direction and viewing angle, the existing local descriptors may have the poor interest points for the stability and distinctiveness.

3.2. The Dichromatic Reflection Model

In this section, we give a brief description of Shafer's dichromatic reflection model [5].

The dichromatic reflection model decomposes the reflected spectrum from a point in viewing direction, $E_R(\lambda)$, into two additive components (*i.e.*, the light $L_s(\lambda,\mathbf{n},\mathbf{s},\mathbf{v})$ reflected at the material surface (so called *surface reflection component*) and the light $L_b(\lambda,\mathbf{n},\mathbf{s},\mathbf{v})$ reflected from the material body (so called *body reflection component*) for inhomogeneous materials such as papers and plastics (see Figure 2). The parameters $\mathbf{n}, \mathbf{s}$, and $\mathbf{v}$ denote the surface patch normal, the direction of the illumination, and the viewing direction respectively; λ is the wavelength. The *surface reflection component* has approximately the same spectral power distribution as the illumination and appears as highlights on object. On the other hand, *body reflection component* provides the characteristic object color and indicates the properties of object shading.

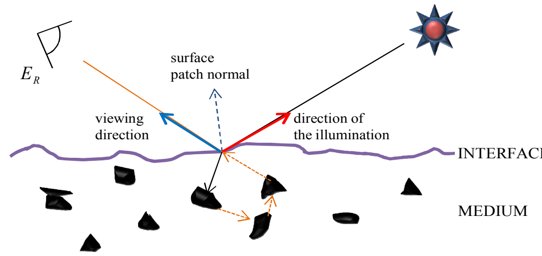


Figure 2: Illustration of the light reflection of inhomogeneous materials

$$E_R(\lambda) = L_b(\lambda,\mathbf{n},\mathbf{s},\mathbf{v}) + L_s(\lambda,\mathbf{n},\mathbf{s},\mathbf{v}) \quad (1)$$

Furthermore, the model separates the spectral reflection properties of L_s and L_b from their geometric reflection properties as follows:

$$E_R(\lambda) = m^s(\mathbf{n},\mathbf{s},\mathbf{v})c^s(\lambda) + m^b(\mathbf{n},\mathbf{s},\mathbf{v})c^b(\lambda) \quad (2)$$

where $c^s(\lambda)$ and $c^b(\lambda)$ are products of spectral power distributions, and geometric terms m^b and m^s model the effect of the illuminant geometry (incident angle, viewing direction, and surface orientation) on the body and surface reflectance, respectively.

3.3. Photometric Quasi-Invariants

In this section, we explain photometric quasi-invariant features and build scale-space using these features.

Consider an image of an infinitesimal surface patch. We assume that the scene consists of inhomogeneous materials such as paper and Fresnel reflectance coefficient has a constant value over the spectrum (*i.e.*, neutral interface reflection model). In addition, for multiple light sources, we assume that the combination can be approximated as a single light source for the local feature. Then, using the spectral sensitivity of *i*-th sensor $s_i(\lambda)$, $i \in \{1,2,3\}$, the measured sensor values at location $\mathbf{x}$ can be given by Shafer's dichromatic reflection model:

$$\begin{aligned} E^i(\mathbf{x}) &= m^b(\mathbf{x})\int_\omega b(\lambda,\mathbf{x})e(\lambda)s_i(\lambda)d\lambda \\ &+ m^s(\mathbf{x})\int_\omega \rho_f(\lambda)e(\lambda)s_i(\lambda)d\lambda \end{aligned} \quad (3)$$

for $E = \{R,G,B\}$ giving the *i*-th sensor response. Further, $b(\lambda,\mathbf{x})$ and $\rho_f(\lambda)$ denote the albedo and Fresnel reflectance respectively. $e(\lambda)$ is the spectral profile of the illuminant; ω denotes the visible spectrum.

Diffuse light that occurs in outdoor/indoor scene (*e.g.*, diffuse light coming from the sky or causing by reflectance from walls) cannot be modeled by Eq. (3). Shafer expands Eq. (3) by introducing the diffuse light, a, by third term.

$$E^i(\mathbf{x}) = m^b(\mathbf{x})\int_\omega b(\lambda,\mathbf{x})e(\lambda)s_i(\lambda)d\lambda + m^s(\mathbf{x})\int_\omega \rho_f(\lambda)e(\lambda)s_i(\lambda)d\lambda + \int_\omega a(\lambda)s_i(\lambda)d\lambda \quad (4)$$

If the sensors $s_i(\lambda)$ are narrowband with spectral response and are approximated by delta functions $s_i(\lambda) = \delta(\lambda - \lambda_i)$, then the reflection function can be simplified to

$$E^i(\mathbf{x}) = e^i\left(m^b(\mathbf{x})b^i(\mathbf{x}) + m^s(\mathbf{x})\rho_f{}^i\right) + a^i \quad (5)$$

Here, the photometric derivative structure of the image can be computed by calculating the spatial derivative of Eq. (5):

$$\begin{aligned} E^i_x(\mathbf{x}) &= e^i(\underbrace{m^b_x(\mathbf{x})b^i(\mathbf{x})}_{\text{shading-shadow}} + \underbrace{m^b(\mathbf{x})b^i_x(\mathbf{x})}_{\text{body reflectance}} + \underbrace{m^s_x(\mathbf{x})}_{\text{specular}}) \\ &= G^n\left(\mathbf{x};\sigma^2\right) * E^i(\mathbf{x}) \end{aligned} \quad (6)$$

where the subscript, x, indicates spatial differentiation and spatial differential quotients are be obtained by convolution $E^i(\mathbf{x})$ with n-order Gaussian derivative filters $G^n\left(\mathbf{x};\sigma^2\right)$ at any scale σ. Since we assume neutral interface reflection model, $\rho_f{}^i$ has a constant value and is independent of $\mathbf{x}$. Note that the spatial derivative is composed of shading-shadow, body reflectance, and specular change.

For a given image $f(x,y)$, its linear scale-space is a family of derived signals $L\left(x,y;\sigma^2\right)$ as follows:

$$L\left(x,y;\sigma^2\right) = G\left(x,y;\sigma^2\right) * f(x,y) \quad (7)$$

At any scale in scale-space, it is possible to apply local derivative operators to the scale-space. That is,

$$\begin{aligned} L_{x^m y^n}\left(x,y;\sigma^2\right) &= \partial_{x^m y^n}\left(L\left(x,y;\sigma^2\right)\right) \\ &= \left(\partial_{x^m y^n} G\left(x,y;\sigma^2\right)\right) * f(x,y) \end{aligned} \quad (8)$$

Here, such scale-space derivatives can be computed by convolving $f(x,y)$ with Gaussian derivative operators due to the commutative property between the derivative operator and the Gaussian smoothing operator. Therefore, the photometric derivative structure, $E^i_x(\mathbf{x})$, is built in scale-space at any scale. And then, in order to accomplish photometric invariance, the photometric derivatives need to be transformed to the underlying color space which is uncorrelated with respect to photometric variations. For this purpose, we use the opponent color space and hue (for more details see [4]).

In the case of a white illuminant, the opponent color space, $OC = \{O^1, O^2, O^3\}$, is known as the orthonormal transformation invariant with respect to specularities, m^s. The photometric derivatives are transformed to the opponent color space as follows:

$$\begin{aligned} O^1_x &= \frac{R_x - G_x}{\sqrt{2}} = \frac{e\left(m^b_x(\vec{x})\left(b^R(\vec{x}) - b^G(\vec{x})\right) + m^b(\vec{x})\left(b^R_x - b^G_x\right)\right)}{\sqrt{2}} \\ O^2_x &= \frac{(R_x + G_x - 2B_x)}{\sqrt{6}} \end{aligned} \quad (9)$$

Now the opponent color space becomes invariant to m^s. However, since these transformations are still variant to lighting geometry, m^b, these also need to be transformed to hue for obtaining invariance to both the specularities and lighting geometry as follows:

$$\begin{aligned} hue_x &= \tan^{-1}\left(\frac{O^1_x}{O^2_x}\right) \\ &= \tan^{-1}\left(\frac{\sqrt{3}\left(\left(b^R(\vec{x}) - b^G(\vec{x})\right) + \left(b^R_x(\vec{x}) - b^G_x(\vec{x})\right)\right)}{\left(\left(b^R(\vec{x}) + b^G(\vec{x}) - 2b^B(\vec{x})\right) + \left(b^R_x + b^G_x - 2b^B_x\right)\right)}\right) \end{aligned} \quad (10)$$

3.4. PI-SIFT Descriptor

As mentioned above, in order to detect the interest points that are robust to both changes, we substitute photometric quasi-invariant features into scale spaces.

First, we build the pyramid similar to the DoG pyramid of SIFT using these features. Also, we construct Gaussian pyramid based on Hue color space instead of it based on gray-scale image to select the more insensitive interesting points to illumination change. We follow the similar strategy of the SIFT in building our PI-SIFT descriptor. That is, we detect candidate points at the extrema of scale-spaces and the maximum geometrical stability of the stable candidate points is achieved by interpolation after eliminating unstable candidate points (for more details see [1]).

4. Experimental Results

We evaluate the PI-SIFT descriptor with respect to photometric variations such as illumination direction and light source changes along with the SIFT.

We use the nearest neighbor-based ratio matching [6]. In this strategy, two regions are matched if the $\|D_P - D_Q\| / \|D_R - D_Q\| < t$, where D_P is the first nearest neighbor to D_Q, D_R the second nearest neighbor to D_Q and t is a threshold. The results are presented in terms of recall vs. 1-precision as described in [6]. The recall-precision curves are formed as the value of t varies. The matching thresholds vary from 0.1 to 0.99 inclusively. We set initial σ to 1.2 for the fair comparison.

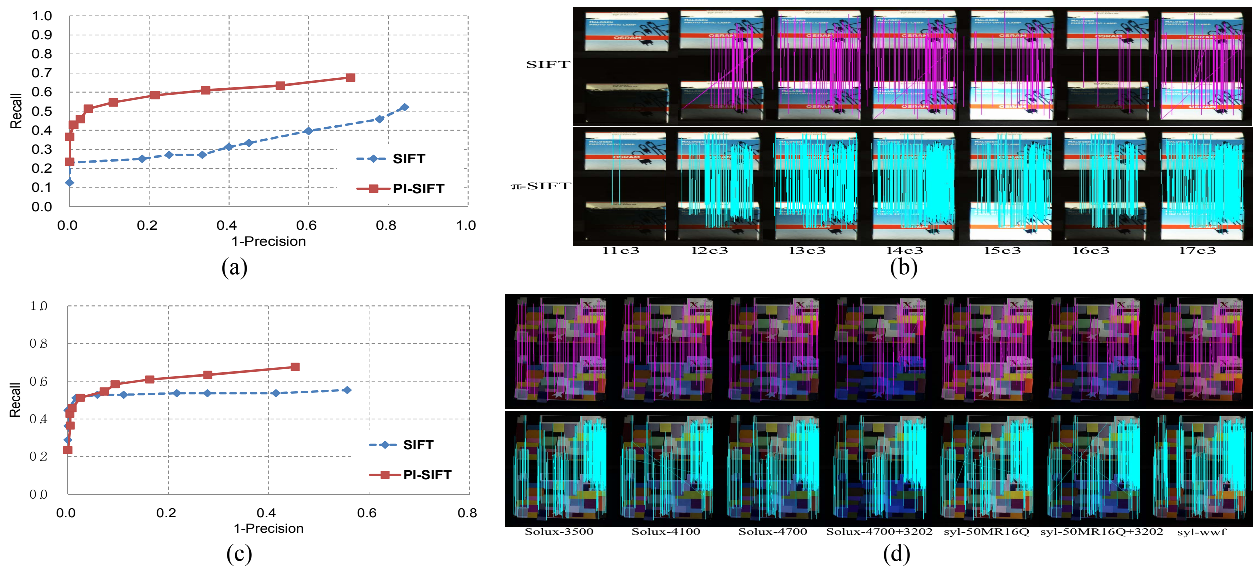


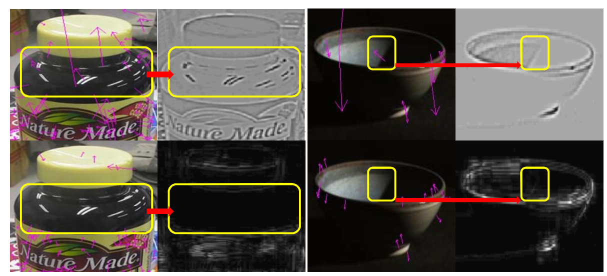
Figure 3: (a) and (b) represent the results for varying illumination directions (*i.e.*, form l1c3 to l7c3). (d) and (e) show the results under different illuminant.

Under the illuminant direction changes, PI-SIFT represents better performance than SIFT as shown in Figure 3(a). Specially, in Figure 3(b), the l6c3 denotes clearly the advantage of PI-SIFT. On the other hand, in Figure 3(c) and (d), although PI-SIFT has many features than SIFT, we can show that PI-SIFT has the result similar to SIFT since SIFT also normalizes the feature vectors to reduce the effects of illumination change.

Figure 4 shows the local derivatives and the photometric quasi-invariant features at $\sigma = 1.2$ in scale-space and the interest points detected by SIFT and PI-SIFT respectively. As shown in Figure 4, PI-SIFT is more robust than SIFT for photometric changes due to the photometric quasi-invariant features.

5. Conclusion

In this paper, we presented a novel local descriptor, PI-SIFT, which is insensitive to photometric variations in addition to geometric invariance. The dichromatic reflection model is used for extracting photometric quasi-invariant features and the similar approach to the SIFT is used for obtaining geometric invariance. Experiments show that our method gives similar performance or outperforms other descriptors. This method may be applicable to image retrieval, object detection and recognition.



(a) Highlights (b) Shadow/shading

Figure 4: above (SIFT) and bottom (PI-SIFT) illustrates depict the features extracted by SIFT and PI-SIFT under photometric variations, respectively.

References

[1] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *Int. J. Computer Vision*, 60(2):91-110, 2004.

[2] H. Bay, T. Tuytelaars, and L. V. Gool. SURF: Speeded up robust features. In *Proc. ECCV* 2006.

[3] A. E. Abdel-Hakim and A. A. Farag. CSIFT: A SIFT descriptor with color invariant characteristics. In *Proc. CVPR*, pages 1978-1983, 2006.

[4] J. van de Weijer, T. Gevers, and J. M. Geusebroek. Edge and corner detection by photometric quasi-invariants, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 27(4):625-630, April 2005

[5] S. A. Shafer. Using color to separate reflection components. *Color Res. Appl.*, 10(4):210-218, 1985.

[6] K. Mikolajczyk and C. Schmid. A performance evaluation of local descriptors. In *Proceedings of CVPR*, pages 257-263, June 2003.

[7] J. M. Geusebroek, R. van den Boomgaard, A. W. M. Smeulders, and H. Geerts. Color invariance. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 23(12):1338-1350, 2001.