



Contents lists available at ScienceDirect

Computer Vision and Image Understanding

journal homepage: www.elsevier.com/locate/cviu

Note

Illumination robust interest point detection

Murat Gevrekci, Bahadır K. Gunturk*

Department of Electrical and Computer Engineering, Louisiana State University, Baton Rouge, LA 70803, United States

ARTICLE INFO

Article history:

Received 12 February 2008

Accepted 26 November 2008

Available online xxxxx

Keywords:

Corner detection

Interest point detection

Illumination invariance

Harris corner detector

ABSTRACT

Most interest point detection algorithms are highly sensitive to illumination variations. This paper presents a method to find interest points robustly even under large non-uniform photometric changes. The method, which we call illumination robust feature extraction transform (IRFET), determines salient interest points in an image by calculating and analyzing *contrast signatures*. A contrast signature shows the response of an interest point detector with respect to a set of contrast stretching functions. The IRFET is generic and can be used with most interest point detectors. In this paper, we demonstrate that the IRFET improves the repeatability rate of the Harris corner detector significantly (by around 25% on average in the experiments).

© 2008 Elsevier Inc. All rights reserved.

1. Introduction

Interest point detection is a necessary task in a variety of computer vision applications, including object recognition, tracking, stereo imaging, and mosaic construction. Although the term interest point detector is sometimes interchangeably used with the term corner detector; interest points do not necessarily correspond to physical corner points. The performance of an interest point detector is evaluated in terms of its accuracy and consistency. An interest point detector should have good spatial localization and be robust against noise and geometric/photometric changes.

One of the earliest interest point detection algorithms is the Moravec corner detector [1]. In the Moravec corner detection algorithm, a patch around a pixel is taken and compared with the neighboring patches. If the pixel is on a smooth region or an edge, there would be at least one neighboring patch that is very similar. For a corner point, all neighboring patches would be different. Therefore, the corner strength at a pixel is defined as the smallest sum of squared differences between the center patch and its neighboring patches. The problem with this algorithm is that only a finite number of neighboring patches (patches in horizontal, vertical, and diagonal directions) are considered; hence, the algorithm is not isotropic.

Harris and Stephens [2] derives a better corner detector by applying Taylor series expansion to the sum of squared difference between a patch and its arbitrarily shifted version. This expansion yields a 2×2 autocorrelation matrix, whose eigenvalues indicate the magnitudes of two orthogonal gradient vectors. For a corner

point, both magnitudes should be large. Since the calculation of eigenvalues and eigenvectors is costly, a corner strength measure based on the determinant and trace of the autocorrelation matrix is proposed. The local maxima of the corner strength map are chosen as the corner points. This method is known as the Harris or the Plessey corner detector in the literature and it is probably the most popular interest point detector. Shi and Tomasi method [3] is based on the Harris corner detector but uses the minimum of the eigenvalues as the corner strength measure; this turns out to be more robust against affine geometric transformations compared to the corner strength measure of the Harris corner detector.

In [4], phase congruency is used to detect edges and corners. The method is based on the *local energy model*, which postulates that the frequency components are in phase at corners and edges [5]. One advantage of this method is the relative insensitivity against illumination changes. The phase congruency method also has very accurate spatial localization. Mokhtarian and Suomela [6] uses the curvature scale space technique and defines the corner points as the points of maximum absolute curvature. The method has been shown to have very good noise robustness. Another important interest point detector is the SUSAN corner detector [7]. The SUSAN operator finds corners based on the fraction of pixels that are similar to the center pixel within a small circular region. A low value of this fraction indicates a corner. The cornerness response at a pixel is basically obtained by subtracting the number of similar pixels from a fixed geometric threshold. A non-maxima suppression operation determines the corner points. The SUSAN operator is also highly robust to noise. Modifications of the SUSAN operator are used for edge detection and for image denoising. (The well-known Bilateral filter [8] is essentially identical to the SUSAN image denoising filter.)

There are several methods that directly address the problem of extracting illumination robust interest points. For example, [9]

* Corresponding author. Fax: +1 225 578 5200.

E-mail addresses: lgevrek1@lsu.edu (M. Gevrekci), bahadir@ece.lsu.edu (B.K. Gunturk).

applies the Harris corner detection procedure on the logarithm of an image instead of the image itself. Similar to the homomorphic filtering, the underlying idea is to separate the illumination and reflectance components and reduce the effect of the illumination component. The method was later extended by first dividing red and blue channels by the green channel and then taking logarithm [10]. In [11], each color channel is normalized by subtracting mean and then dividing by the square root of the sum of the squares of the pixel values in that channel. This procedure would work well if the illumination changes can be effectively modeled as scale changes. Ref. [12] also models the illumination changes as scale factors, modifies the optical flow equation, and solves for the scale factors in addition to other geometric transformation parameters. When the illumination changes are large, the effect on the pixel intensities would be non-uniform and these methods would likely to perform not as much effectively as in small illumination changes.

Recently, methods to extract scale and affine invariant interest points have been proposed. A review of these methods can be found in [13]. A comprehensive evaluation of the interest point detectors is provided in [14]. Among these methods, the scale invariant feature transform (SIFT) [15] method also presents a region descriptor based on the local histogram of the gradient vectors to achieve illumination invariance in addition to scale invariance. In the illumination invariance experiments of [14], the SIFT method is among the best performers.

In this paper, we present a method, illumination invariant feature extraction transform (IRFET), to improve the robustness of feature detection against illumination changes. The IRFET stretches the histogram of an image around a set of pixel intensities; an interest point detector is applied to each contrast-stretched image to produce a three-dimensional cornerness map. Interest points are determined from this map.

The IRFET is generic and can be combined with a feature detector that creates feature strength map and then applies non-maxima suppression to it. In this paper, we show how the IRFET improves the repeatability rate of the Harris corner detector. Section 2 reminds the standard Harris corner detector. Section 3 explains the idea behind the IRFET method and also presents its application on the Harris corner detector. The robustness of the proposed method over the Harris corner detector is proven in Section 4. Section 5 compares several interest points detectors experimentally; and Section 6 concludes the paper.

2. Harris corner detector

One of the most commonly used interest point detectors in computer vision applications is the Harris corner detector [2]. The Harris corner detector is based on the autocorrelation matrix of the image gradients. The autocorrelation matrix $\mathbf{A}(x, y; I)$ of an image I at a pixel location (x, y) is given as follows:

$$\mathbf{A}(x, y; I) = \begin{pmatrix} \sum_{(m,n) \in N} \left(\frac{\partial}{\partial x} I(m, n) \right)^2 & \sum_{(m,n) \in N} \frac{\partial}{\partial x} I(m, n) \frac{\partial}{\partial y} I(m, n) \\ \sum_{(m,n) \in N} \frac{\partial}{\partial x} I(m, n) \frac{\partial}{\partial y} I(m, n) & \sum_{(m,n) \in N} \left(\frac{\partial}{\partial y} I(m, n) \right)^2 \end{pmatrix}, \quad (1)$$

where $\partial/\partial x$ and $\partial/\partial y$ calculates the gradients in horizontal and vertical directions, respectively; and N is a set of pixels around (x, y) . Usually a weighted sum of the gradients in N is taken using a Gaussian function centered at (x, y) to give more weight to the pixels that are close to (x, y) . The cornerness response function of the Harris corner detector is based on the determinant and trace of the autocorrelation matrix:

$$R_{\text{harris}}(x, y) = \det \mathbf{A}(x, y; I) - k(\text{trace} \mathbf{A}(x, y; I))^2, \quad (2)$$

where k is a small positive constant controlling the cornerness sensitivity of the detector. After calculating the cornerness response for all pixels, non-maxima suppression is used to get the corner points.

3. Illumination robust feature extraction transform

The IRFET is an intuitive method to improve illumination invariance of feature detectors. The underlying idea is to stretch the image contrast as a function of intensity to span the space of possible photometric transformations. By applying a feature detector to a contrast-stretched image, the response of the feature detector under a particular illumination condition is simulated. The collection of the responses under a set of illumination conditions forms a signature for each pixel. The signatures can then be used to characterize a pixel and to find illumination robust interest points.

The contrast stretching function that we use in our experiments is the sigmoid function, which has the following form:

$$f_c(I(x, y)) = \frac{1}{1 + e^{-\gamma(I(x, y) - c)}}, \quad (3)$$

where $I(x, y)$ is the normalized intensity value in the range $[0, 1]$, c is the contrast center around which the contrast is stretched, and γ determines the slope of the sigmoid function. Fig. 1 illustrates the sigmoid function.

By applying the contrast stretching function at different contrast centers, we obtain a set of contrast-stretched images. (Fig. 2 shows a set of contrast-stretched versions of two images at various contrast centers. Notice that these two different images produce similar responses at particular contrast centers. For example, $c = 0.8$ for Image (a) produces an image that is very similar to the output of $c = 0.2$ for Image (b). Likewise, the output of $c = 1.0$ for Image (a) is similar to the output of $c = 0.5$ for Image (a). Also, notice how some details in $c = 0.05$ for Image (b) get apparent as in Image (a).) A feature detector is then applied to each of these contrast-stretched images. In case of the Harris corner detector, first, the autocorrelation matrix at each pixel is found:

$$\mathbf{A}(x, y; I_c) = \begin{pmatrix} \sum_{(m,n) \in N} \left[\frac{\partial}{\partial x} I_c(m, n) \right]^2 & \sum_{(m,n) \in N} \left[\frac{\partial}{\partial x} I_c(m, n) \right] \left[\frac{\partial}{\partial y} I_c(m, n) \right] \\ \sum_{(m,n) \in N} \left[\frac{\partial}{\partial x} I_c(m, n) \right] \left[\frac{\partial}{\partial y} I_c(m, n) \right] & \sum_{(m,n) \in N} \left[\frac{\partial}{\partial y} I_c(m, n) \right]^2 \end{pmatrix}, \quad (4)$$

where $I_c(m, n) \equiv f_c(I(m, n))$ is the contrast-stretched image at contrast center c . And then, the cornerness response is calculated:

$$R(x, y; c) = \det \mathbf{A}(x, y; I_c) - k(\text{trace} \mathbf{A}(x, y; I_c))^2. \quad (5)$$

Note that the cornerness response is a function of the contrast center c . $R(x, y; c)$ for all $(x, y; c)$ is a three-dimensional matrix that can be used to analyze the pixels in terms of the cornerness strength.

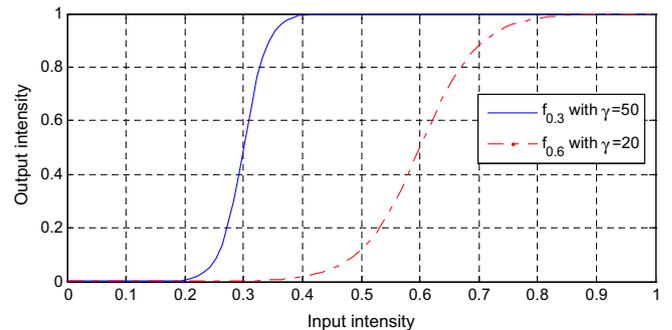


Fig. 1. Sigmoid functions at contrast centers $c = 0.3$ and $c = 0.6$ are plotted. The parameter γ controls the slope of the sigmoid.

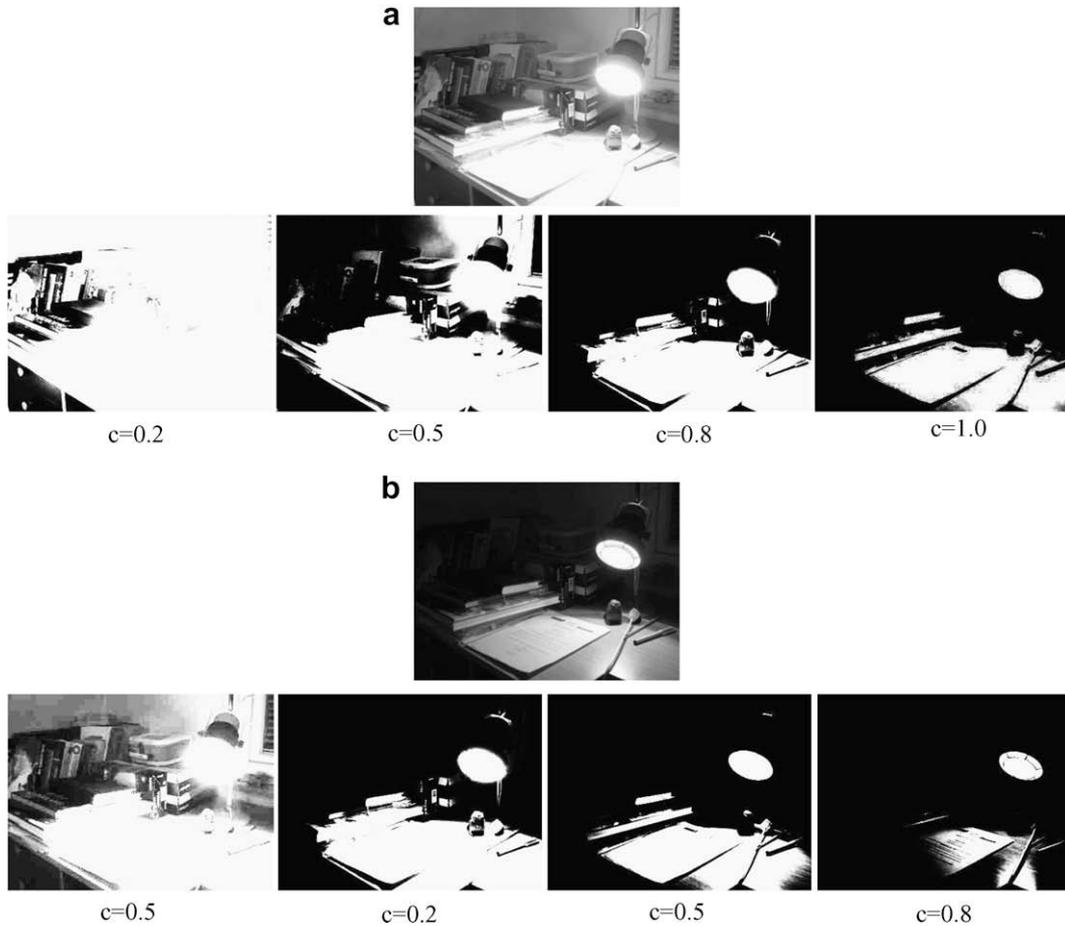


Fig. 2. Contrast stretched versions of two images are displayed at several centers c ($\gamma = 50$ for all images).

By plotting $R(x, y; c)$ as a function of c at a pixel position, we can obtain the “contrast signature” of that pixel. Fig. 3 shows the contrast signatures at several pixels on an image. Specifically, five corner pixels are selected from the image. These pixels are enumerated from 1 to 5. Among these pixels, Pixel 1 is on a corner with the highest contrast. When we look at the contrast signature of that pixel, we see that it has a strong response over a wide range of c . On the other hand, a corner pixel with low contrast returns a large response over a limited range of c . (Although not plotted in the figure, pixels that are on edges and smooth regions return a low response for all values of c .)

Based on the contrast signatures, we can define various cornerness measures. One possible measure is the maximum response

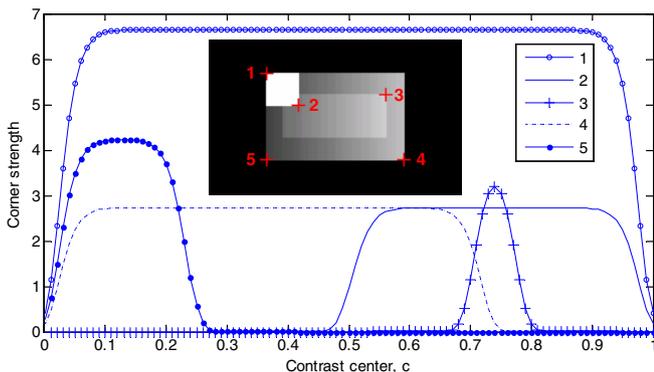


Fig. 3. Sample pixels and the corresponding contrast signatures are shown.

$R_{max}(x, y) = \max_c R(x, y; c)$. Although this makes sense because a corner point, even if it has a low contrast, would return a high $R(x, y; c)$ value at some c , in real applications it turns out to be an unreliable measure because of its sensitivity to noise.

Another possible cornerness measure is the average response over all c , equivalently, the area under the contrast signature:

$$R_{area}(x, y) = \int_0^1 R(x, y; c) dc. \quad (6)$$

The area response $R_{area}(x, y)$ is calculated using the Riemann sum in practice, and it turns out to be a more reliable measure as it favors corners that return high response for a wide range of illumination conditions. In the next section, we will also prove that $R_{area}(x, y)$ is less sensitive to illumination changes compared to $R_{harris}(x, y)$.

Fig. 4 shows these cornerness responses for the image given in Fig. 3. As seen, the standard Harris corner detector misses the corners with low contrast; on the other hand, $R_{area}(x, y)$ and $R_{max}(x, y)$ highlights all the corners well. Among all, $R_{max}(x, y)$ has the best distinguishing response. To test how the cornerness response changes as a function of contrast, we set up another experiment, shown in Fig. 5. The intensity of the white patch in a black image is reduced from 1 to 0; and the responses $R(x, y)$, $R_{area}(x, y)$, and $R_{max}(x, y)$ are measured at the corner of the white patch. It is observed that the standard Harris response drops sharply, while the responses based on the contrast signature are more robust.

These experiments indicate the maximum response $R_{max}(x, y)$ as the best measure; however, as mentioned previously, it turned out that it is not as reliable as $R_{area}(x, y)$ in real applications.

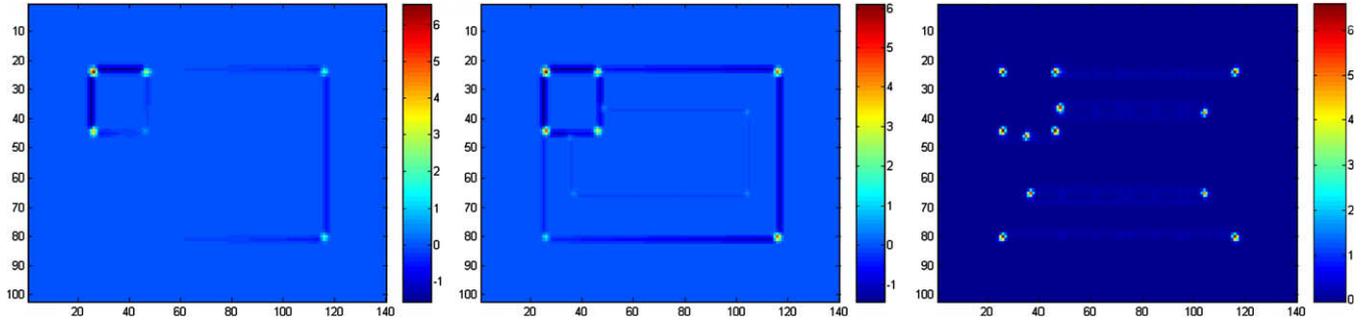


Fig. 4. Left to right: Cornerness responses of the standard Harris corner detector $R_{harris}(x,y)$, IRFET-Harris with $R_{area}(x,y)$, and IRFET-Harris with $R_{max}(x,y)$ are displayed for the image in Fig. 3.

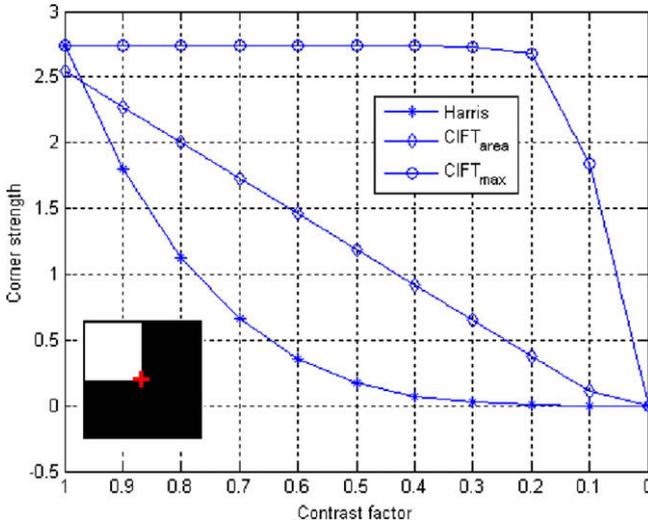


Fig. 5. The intensity of the white patch is reduced from 1 to 0. The cornerness responses of the standard Harris corner detector $R_{harris}(x,y)$, IRFET-Harris with $R_{area}(x,y)$, and IRFET-Harris with $R_{max}(x,y)$ as a function of this intensity are plotted for the marked corner pixel.

4. Proof of illumination robustness

In this section, we will prove that $R_{area}(x,y)$ is more robust against illumination changes than the Harris response $R_{harris}(x,y)$ is. First, we incorporate the camera response function, which is the main source of illumination sensitivity under large photometric changes, into the picture. Let $z(m,n)$ be the irradiance of a scene and $g_x(\cdot)$ be the camera response function that converts the irradiance into pixel intensity:

$$I(m,n) = g_x(z(m,n)). \quad (7)$$

The parameter α is the parameter associated with the illumination change. For example, in the classical formulation of the camera response function [16], $g_x(z(m,n)) = a_1 + a_2(\alpha z(m,n))^{a_3}$, where (a_1, a_2, a_3) are the parameters specific to the camera, and α represents gain and/or exposure time of the sensor. In the more general polynomial formulation [17], $g_x(z(m,n)) = \sum_{k=0}^{K-1} a_k(\alpha z(m,n))^k$.

We will now drop the spatial positions from our notations for simplicity purposes, define $q(I) \equiv \det \mathbf{A}(I) - k(\text{trace} \mathbf{A}(I))^2$, and rewrite the Harris corner detector response R_{harris} and the area response R_{area} as a function of the irradiance as follows:

$$R_{harris} = q(I) = q(g_x(z)) \quad (8)$$

and

$$R_{area} = \int_0^1 q(f_c(I)) dc = \int_0^1 q(f_c(g_x(z))) dc. \quad (9)$$

In order to prove that R_{area} is less sensitive to illumination changes than R_{harris} is, all we need to do is to show that $(\partial R_{area} / \partial \alpha) \leq (\partial R_{harris} / \partial \alpha)$. Using the chain rule, the derivative of R_{harris} can be written as

$$\frac{\partial R_{harris}}{\partial \alpha} = \frac{\partial q}{\partial I} \frac{\partial I}{\partial \alpha} = q' \frac{\partial g_x(z)}{\partial \alpha}. \quad (10)$$

Note that (10) requires both q and g_x to be differentiable. While we are not interested in the exact formulas of $\partial q(x,y) / \partial I(m,n)$ or $\partial I(m,n) / \partial \alpha$, these derivatives exist and can be calculated by using the definitions of \mathbf{A} , q and g_x . Similarly, we can write the derivative of R_{area} as follows:

$$\frac{\partial R_{area}}{\partial \alpha} = \frac{\partial}{\partial \alpha} \int_0^1 q(f_c(g_x(z))) dc = \int_0^1 q' f_c' \frac{\partial g_x(z)}{\partial \alpha} dc. \quad (11)$$

Combining the last two equations, we get $\partial R_{area} / \partial \alpha = \partial R_{harris} / \partial \alpha \int_0^1 f_c' dc$; therefore, to complete the proof it is sufficient to show that $\int_0^1 f_c' dc \leq 1$. Using the definition of the sigmoid, we obtain

$$\int_0^1 f_c' dc = \int_0^1 \frac{\partial}{\partial s} \left(\frac{1}{1 + e^{-\gamma(s-c)}} \right) dc = \frac{1}{1 + e^{-\gamma s}} - \frac{1}{1 + e^{-\gamma(s-1)}}, \quad (12)$$

where $s \in [0,1]$. The maximum value of this function is reached when $s = 0.5$, where its value is 1. Therefore, $\int_0^1 f_c' dc \leq 1$ for any value of s regardless of the value of γ .

5. Experiments

In this section, we report the experiments demonstrating how the IRFET improves the repeatability rate of the Harris corner detector under large photometric changes and also compare with several state-of-the-art corner detectors under the same conditions.

There are three data sets, which are shown in Figs. 6–8. There are non-uniform illumination changes and severe saturation in these images. In each set, one image is chosen as the reference image, and the repeatability rate is computed for each of the remaining images.

There are different measures of repeatability rate for interest point detectors; we use the following measure. Suppose N_1 is the number of interest points in Image 1; N_2 is the number of interest points in Image 2; and N is the number of common interest points. Then, the repeatability rate is defined as $N / \min(N_1, N_2)$. An interest point in an image is repeated if there is an interest point within a 3×3 neighborhood of the other image.

In the experiments, we used the following parameters. For the Harris corner detector, the value of k was 0.04. In calculating the autocorrelation matrix, we used a Gaussian function with standard deviation of 1 to get the weighted sum of the gradients within a

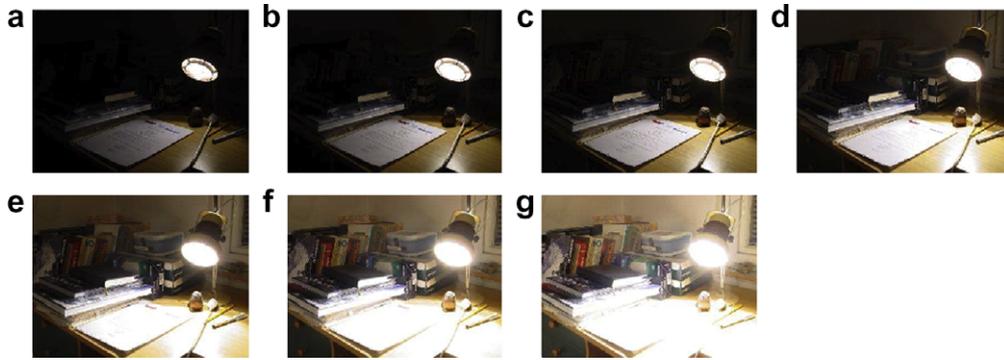


Fig. 6. Data set 1. Image in (d) is chosen as reference. Repeatability rates between this image and the other images are computed. The images in (a), (b), (c), (e), (f), and (g) are labeled as 1–6 in Fig. 9. The image size is 300×400 .



Fig. 7. Data set 2. Image in (a) is chosen as the reference image. Repeatability rates between this image and the other images are computed. The images in (b), (c), and (d) are labeled as 1, 2, and 3 in Fig. 9. The image size is 480×640 .

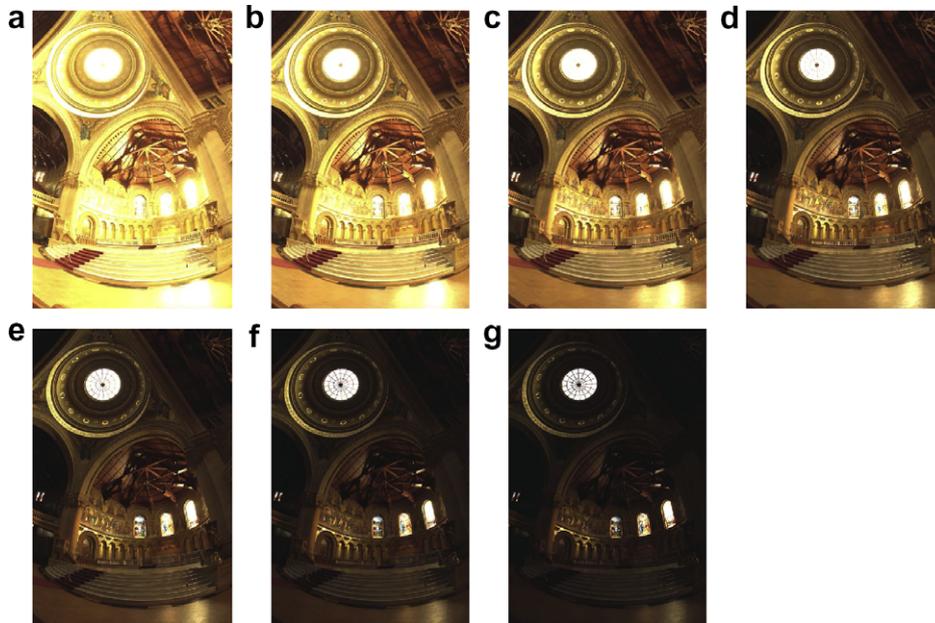


Fig. 8. Data set 3. Image in (d) is chosen as reference. Repeatability rates between this image and the other images are computed. The images in (a), (b), (c), (e), (f), and (g) are labeled as 1–6 in Fig. 9. The image size is 768×512 .

7×7 local neighborhood. For the non-maxima suppression, the pixels with cornerness response less than a threshold were eliminated, and the local maxima within 3×3 regions were found for the remaining pixels. The threshold value, below which the corresponding pixels are eliminated, was set to 2% of the maximum value of the cornerness response. For the IRFET-Harris algorithm, the signatures were obtained for $c \in [0, 1]$ with step size of 0.05 and with $\gamma = 50$, and R_{area} values were calculated using the Riemann sum. The values of the step size and the γ parameter were decided after some trial-and-error to have a good overall performance. In addition to the standard Harris and the IRFET-Harris, we included

results of the Phase Congruency [4], SIFT [15], and SUSAN [7] methods. The software for these methods are provided by the original authors. For these methods, the parameters were chosen as the default values given in these softwares.

The results are given in Fig. 9. In the figure, both the repeatability rate and the number of repeated corners are plotted. As seen in the results, the IRFET improves the repeatability rate of the standard Harris corner detector by around 25% on average. The improvement is larger for images with large photometric differences. When the number of repeated corners is examined, it can be seen that the number of repeated corners is significantly larger

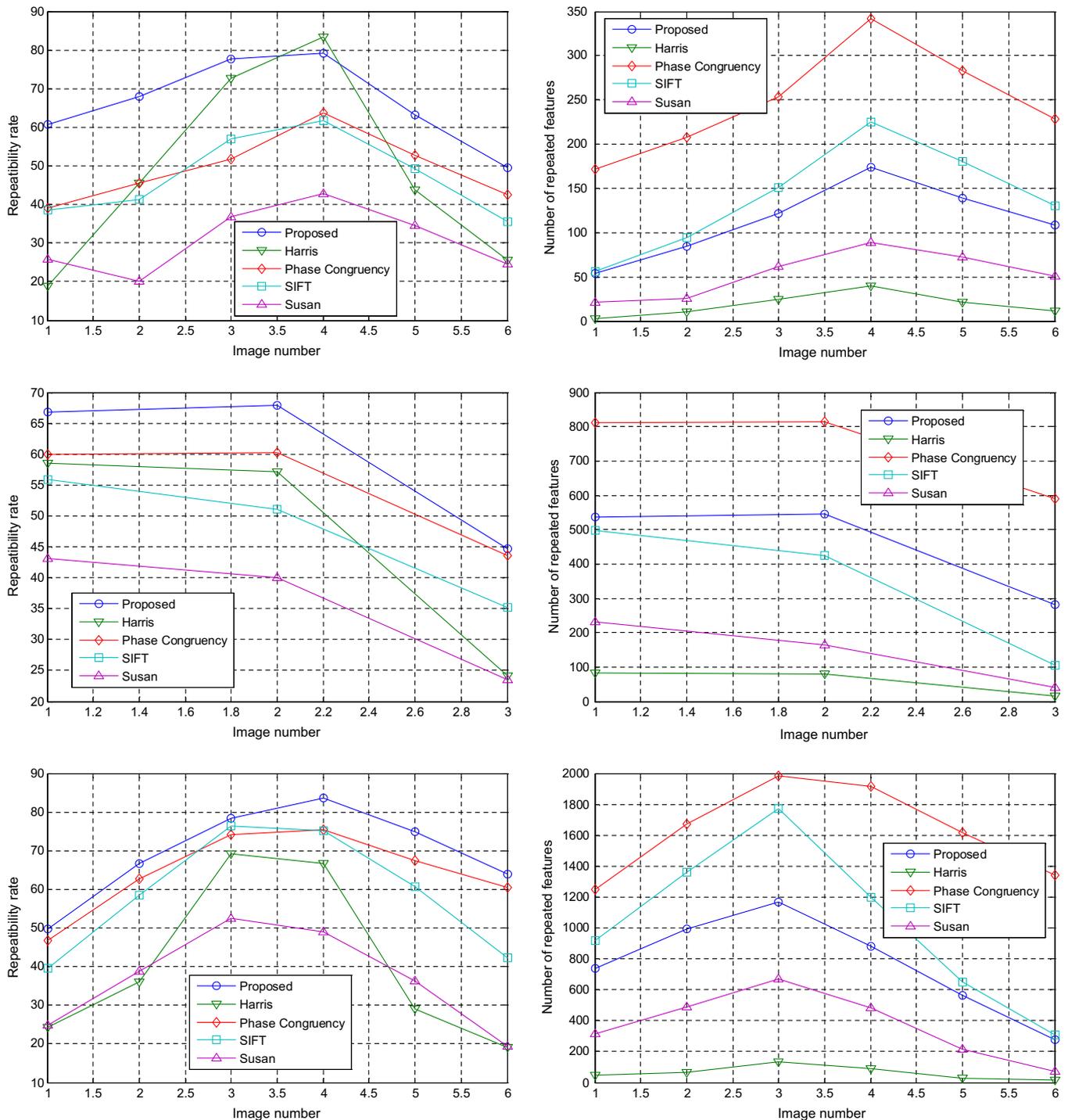


Fig. 9. Top to bottom: Experimental results (the repeatability rates and the number of repeated corners) for data set 1, data set 2, and data set 3, respectively.

for the IRFET-Harris compared to the standard Harris. The number of repeated corners might be important in some applications, for example, when registering images with relatively small overlapping areas. One might also try to improve the repeatability rate by increasing the threshold of the non-maxima suppression or by ranking the pixels according to the cornerness strength and then taking the top, say, M of them. For these scenarios, the IRFET-Harris has larger space for improvement than the standard Harris.

While improving the standard Harris, the IRFET-Harris outperforms the other interest point detectors as well. In all experiments,

the IRFET-Harris has higher repeatability rate than the Phase Congruency, SIFT, and SUSAN methods.

6. Conclusions

In this paper, we proposed a method, IRFET, to detect interest points robustly under varying illumination conditions and demonstrated how it improves the repeatability rate of the standard Harris corner detector. The IRFET-Harris also performs better than several other state-of-the-art corner detection algorithms. An

important drawback of the IRFET method is the computational complexity. In the current implementation, if the contrast center c is sampled at m points, then the computational complexity of the IRFET is approximately m times the computational complexity of the method it is applied. Of course, for applications where performance is more important than computational complexity, the IRFET would be very beneficial. Also note that the method is parallel in nature: the response for each contrast center can be computed independently. This allows parallel implementation of the method. A possible future work is to combine the IRFET method with affine invariant interest point detectors to obtain affine and photometric invariant feature detectors.

Acknowledgments

This work was supported in part by the National Science Foundation under Grant No. 0528785. The authors would like to thank Martin Cadik for providing the first data set. The third data set is obtained from the website of Dr. Paul Debevec (<http://www.debevec.org/>).

References

- [1] H. Moravec, Obstacle avoidance and navigation in the real world by a seeing robot rover, Doctoral Dissertation Technical Report CMU-RI-TR-80-03, Robotics Institute, Carnegie Mellon University, September 1980.
- [2] C. Harris, M. Stephens, A combined corner and edge detector, in: Proceedings of the 4th Alvey Vision Conference, 1988, pp. 147–152.
- [3] J. Shi, C. Tomasi, Good features to track, in: 1994 IEEE Conference on Computer Vision and Pattern Recognition (CVPR'94), 1994, pp. 593–600.
- [4] P. Kovesi, Image features from phase congruency, *Videre* 1 (3) (1999) 1–26.
- [5] M. Morrone, J. Ross, D. Burr, R. Owens, Mach bands are phase dependent, *Nature* 324 (3) (1986) 250–253.
- [6] F. Mokhtarian, R. Suomela, Robust image corner detection through curvature scale space, *IEEE Transactions on Pattern Analysis & Machine Intelligence* 12 (20) (1998) 1376–1381.
- [7] S. Smith, J. Brady, Susan: a new approach to low-level image-processing, *International Journal of Computer Vision* 23 (1) (1997) 45–78.
- [8] C. Tomasi, R. Manduchi, Bilateral filtering for gray and color images, in: Proc. of the Int. Conf. Computer Vision, 1998, pp. 839–846.
- [9] F. Faillie, A fast method to improve the stability of interest point detection under illumination changes, in: Proc. of the IEEE Int. Conf. Image Processing, 2004, pp. 2673–2676.
- [10] F. Faillie, Stable interest point detection under illumination changes using colour invariants, in: Proc. of the British Machine Vision Conference, 2005.
- [11] M.S. Drew, J. Wei, Z.-N. Li, Illumination-invariant image retrieval and video segmentation, *Pattern Recognition* 32 (1999) 1369–1388.
- [12] B. Georgescu, P. Meer, Point matching under large image deformations and illumination changes, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26 (6) (2004) 674–688.
- [13] K. Mikolajczyk, C. Schmid, Scale and affine invariant interest point detectors, *International Journal of Computer Vision* 60 (1) (2004) 63–86.
- [14] K. Mikolajczyk, C. Schmid, A performance evaluation of local descriptors, *IEEE Transactions on Pattern Analysis & Machine Intelligence* 27 (10) (2005) 1615–1630.
- [15] D. Lowe, Distinctive image features from scale-invariant keypoints, *International Journal of Computer Vision* 60 (2) (2004) 91–110.
- [16] S. Mann, Comparametric equations with practical applications in quantigraphic image processing, *IEEE Transactions on Image Processing* 9 (8) (2000) 1389–1406.
- [17] T. Mitsunaga, S. Nayar, Radiometric self calibration, in: Proc. of the IEEE Int. Conf. Computer Vision and Pattern Recognition, vol. 1, 1999, pp. 374–380.