A FAST METHOD TO IMPROVE THE STABILITY OF INTEREST POINT DETECTION UNDER ILLUMINATION CHANGES

Flore Faille

Institute for Real-Time Computer Systems
Technische Universität München, D-80333 Munich, Germany
Flore.Faille@rcs.ei.tum.de

ABSTRACT

This paper presents a method to improve the stability of interest point detection under illumination changes. The response of state–of–the–art detectors depends on image contrast, which leads to unstable detection, especially when the position or orientation of the light source changes. Based on a simple image formation model and on ideas of homomorphic filtering, the Harris corner detector – one of the most widely used detectors – is modified to reduce the illumination influence. The computation effort needed for detection is only minimally increased. The algorithm performances are evaluated on real images using repeatability and false positive rates. The detection stability is enhanced, above all under complex lighting changes.

1. INTRODUCTION

Many computer vision applications like 3D–reconstruction, tracking, registration or recognition tasks rely on the detection of interest points. The interest points and their characterization are the only image information used during further processing. For that reason, a stable detection under varying imaging conditions is important, as misdetections highly influence the performances of higher level tasks.

This work is focused on the problem of illumination changes, which occurs very often between images of a scene taken at distant time instants e.g. due to a movement of the light source(s). As the response of state–of–the–art detectors varies with image contrast, the choice of an appropriate detection threshold becomes almost impossible, particularly if a limited number of points should be detected.

In this paper, a method is proposed to improve the stability of interest point detection. Special care has been given to the algorithm efficiency, in order to keep the computational effort low. Based on a simple image formation model and on ideas of homomorphic filtering, the detection is adapted to the local lighting conditions. The obtained improvements are demonstrated on a very popular interest point detector: the Harris corner detector [1].

Section 2 underlines the sensitivity of interest point detection to lighting changes. In section 3 the principle and implementation details of the proposed method are presented. Experiments and results are described in section 4.

2. INTEREST POINT DETECTION AND ILLUMINATION CHANGES

Many interest point detectors were designed, based for example on local grey value maxima [2], on curvature maxima along contours [3], on the local grey value distribution [4] or on maxima of the local autocorrelation function [1]. Despite their different principles, all detectors are very sensitive to image contrast. The Harris corner detector [1] was chosen in this paper, as it is one of the most widely used.

The local texture around pixel (x, y) is characterized by:

$$\mathbf{C} = G(\sigma) \otimes \begin{bmatrix} f_x^2 & f_x\, f_y \\ f_x\, f_y & f_y^2 \end{bmatrix} [1]. \tag{1}$$

$G(\sigma)$ is a Gaussian with standard deviation σ and $\otimes$ is the convolution operator. f_x and f_y are the first derivatives of the grey value image $f(x, y)$. They are estimated here by convolving f with the derivatives of a Gaussian (with $\sigma_{deriv} = 1.2$) as proposed in [3]. Two high eigenvalues for matrix $\mathbf{C}$ indicate an interest point. For more efficiency, the "cornerness" function R is used:

$$R = \mathrm{Det}(\mathbf{C}) - \alpha\, \mathrm{Tr}^2(\mathbf{C}), \text{ with } 0.04 \leq \alpha \leq 0.06 \text{ [1].} \tag{2}$$

Interest points are then found at local maxima of R above a user–defined threshold T ($T > 0$). Here $\sigma = 3.0$ and $\alpha = 0.06$ are used.

The cornerness function R varies with the local image contrast because it is based on the first derivatives. A detailed analysis of its sensitivity to illumination is given in [5]. Despite the numerous methods to gain constant features from color images (see e.g. [6]), only few interest point detectors that are stable under varying illumination were designed. The following two methods are often used to adapt the detection to the overall lighting conditions: (a)

0-7803-8554-3/04/$20.00 ©2004 IEEE.

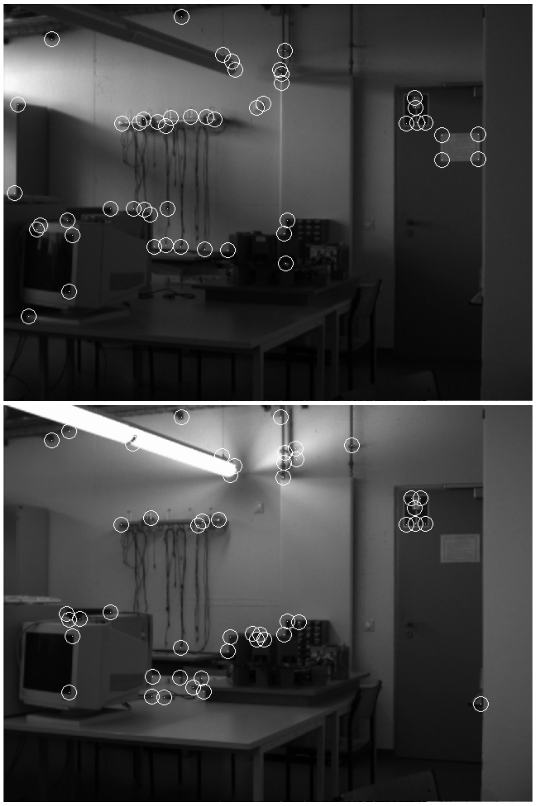

Fig. 1. Interest point detection on images of a scene under two different illuminations ($N = 50$). Only 28.0% of the points are redetected.

the threshold is set proportional to the maximum cornerness value [3], or (b) the N interest points with the highest cornerness values are selected [7]. As method (b) reaches the best stability under lighting changes [5], it will be used here for detection instead of the user–defined threshold. However this method still is insufficient to reach a good stability under complex lighting variations, as shown in [5] or in Fig. 1, where only 28% of the interest points are redetected after the lamps have been turned on.

To improve stability, we introduced and compared three algorithms in [5], which locally adapt detection to lighting conditions: derivative normalization using local image energy, threshold adaptation using local mean and variance of R and local clustering–based threshold computation. The algorithm proposed in this paper enhances the local derivative normalization method, reducing noise sensitivity as well as computation time. Its results are comparable to those of the best method (the local threshold adaptation) despite the simpler image model and the 16 % lower computation time. While the local threshold adaptation method achieves a more uniform interest point distribution in the image, the algorithm presented here performs better when saturation occurs.

3. PROPOSED IMPROVEMENT

3.1. Principle

Images are generated by the light reflected from the observed scene. For Lambertian surfaces, the grey value image $f(x, y)$ can be modelled as [8]:

$$f(x, y) = i(x, y)r(x, y), \tag{3}$$

where $i(x, y)$ and $r(x, y)$ represent the illumination and the reflectance components. Most of the time, the illumination can be assumed to vary slowly in space. This is the basis of homomorphic image filtering. By taking the logarithm of $f(x, y)$ (which is positive), the multiplicative relation can be transformed to an additive relation:

$$\ln f(x, y) = \ln i(x, y) + \ln r(x, y). \tag{4}$$

The logarithm of the illumination $\ln i(x, y)$ still can be assumed to vary slowly in space, so that illumination independent image information can be gained e.g. by applying a high–pass filter to $\ln f(x, y)$.

As most interest point detectors are based on some kind of high–pass filtering, the idea of homomorphic filtering can be applied to improve their stability under lighting variations. For the Harris detector, the derivatives in (1) need only be calculated on the logarithm of the image. The illumination varies slowly in space ($i_x \approx i_y \approx 0$), so:

$$\frac{\partial \ln f}{\partial x} = \frac{f_x}{f} = \frac{i_x r + i r_x}{ir} \approx \frac{r_x}{r}\,, \; \frac{\partial \ln f}{\partial y} \approx \frac{r_y}{r}. \tag{5}$$

i_x, i_y, r_x and r_y are the derivatives of i and r. The illumination influence on $\mathbf{C}$ and R disappears. Consequently a stable detection can be achieved with a fixed threshold T ($T > 0$). This new detector is equivalent to a detection based on normalized image derivatives $f_x/\bar{f}$ and $f_y/\bar{f}$, where $\bar{f}$ is the mean grey value over the neighborhood used to estimate f_x and f_y. The detection is locally adapted to the lighting conditions, yielding more stable results than obtained with the usual Harris detector. The execution time is only minimally increased. Because of the simple image formation model, instabilities can be expected close to specularities, to depth or surface normal discontinuities, and when sharp light or shadow patterns appear.

3.2. Implementation

If implemented in a straightforward manner, the proposed method encounters problems in very dark image regions. First, the noise influence is strongly amplified, as the divisor in (5) takes very low values. Additionally, pixels having a grey value equal to 0 cannot be handled.

The use of $\ln(1 + f(x, y))$ instead of $\ln f(x, y)$ proved experimentally to be a good workaround against both problems. In bright regions, adding 1 to the grey values has a

Fig. 2. Suppression of the illumination influence on the derivatives. Upper left: grey value image. Upper right: gradient $|f_x| + |f_y|$. Lower left: gradient of $\ln f$. Lower right: gradient of $\ln(1 + f)$ with pre–processing. All gradient images were scaled to the maximum grey value.

negligible effect on the derivatives. In dark regions, it actually helps to reduce noise influence. Furthermore the following pre–processing is applied to minimize noise–induced false detections in very dark areas: all pixels with a grey value smaller than V are replaced by the mean value in their 3×3 neighborhood. V was chosen experimentally to be 3.

Figure 2 illustrates how the illumination influence on the derivatives is suppressed and how noise effects in dark regions can be reduced by the presented implementation.

4. EXPERIMENTAL RESULTS

The results of the proposed detector on the images of Fig. 1 are given in Fig. 3. The threshold ($T = 0.0002$) was chosen to obtain the same number of interest points in both top images ($N = 50$). The detection stability is enhanced: 56.6% of the points are now redetected. The needed execution time is approximately 9.5% higher than for the Harris detector using a fixed threshold. It is in most cases shorter than for the Harris detector selecting the N best points.

To evaluate the proposed method, it was tested on two image series presenting a scene under varying illumination conditions. The performances are evaluated with the repeatability rate, a criterion introduced in [3] to measure the stability of interest point detection. This rate is the ratio between the number of redetected points and the number of points in the reference image. A point is considered redetected if it is in the 3×3 neighborhood of a reference point. The false positive rate – the number of non–redetected points in the current image divided by the number of points in the current image – must be considered too, as the number of detections is not constant between images.

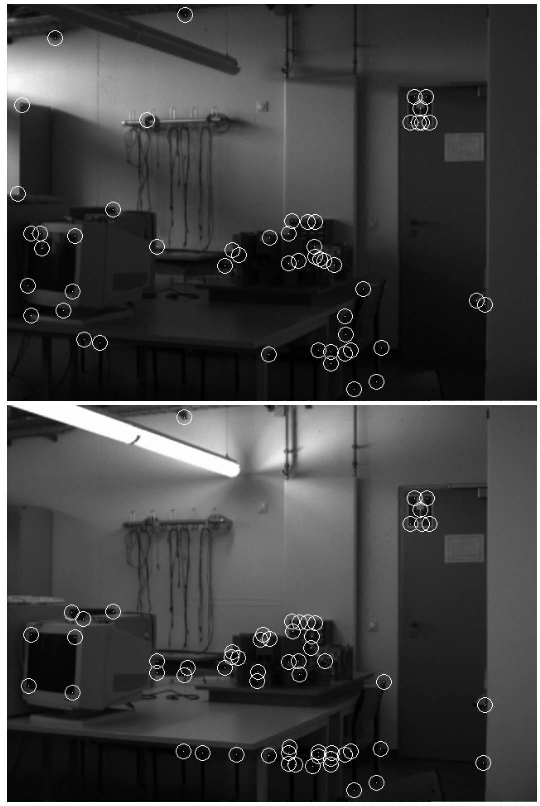

Fig. 3. Interest point detection with the proposed method on the images of Fig. 1. 56.6% of the points are redetected.

The first considered illumination variation is a change of the lighting intensity, which was simulated by varying the camera shutter time. This global illumination change can be compensated by selecting the N best interest points. The image with the medium shutter time was chosen as reference. The repeatability and false positive rates for the proposed detector and for the Harris detector with selection of the N best points are given in Fig. 4. The scene was chosen to show the differences between the detectors. The first 5 images ($m < 0.45$ in Fig. 4) are relatively difficult for the here proposed method, because of the large dark areas. More than 25% of the image must be pre–processed, which results in lower repeatability. However, the false positive rate stays low, showing the strength of the implementation. The last 5 images ($m > 1.4$) present increasingly saturated areas, which can be better overcome with the proposed detector. For the medium images, the performances of both methods are equivalent, as could be expected.

In the second image series (Fig. 5), complex illumination changes were generated by varying the number, type (sunlight, neon and normal lamps) and position of the light sources. The first image of the series is the reference image. To better visualize the results, a complexity measure

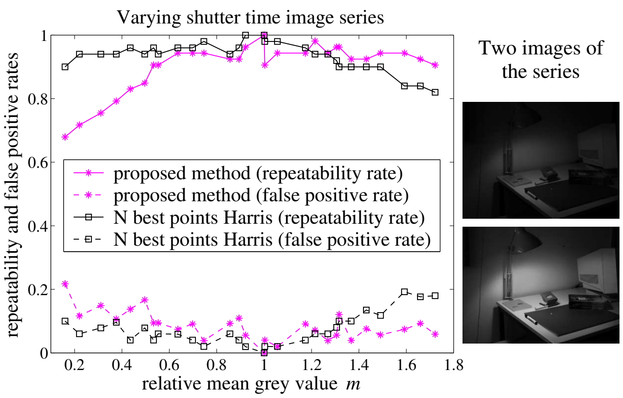

Fig. 4. Detection stability for an image series with varying lighting intensity.

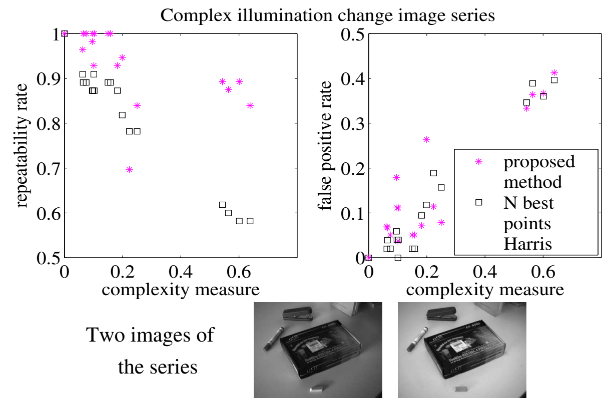

Fig. 5. Detection stability for an image series with complex lighting changes.

is defined to characterize the lighting variation between two images f_1 and f_2:

$$\mathrm{CM}(f_1, f_2) = \mathrm{CM}(f_2, f_1) = \sigma\left(\frac{f_1 - \mu_1}{\sigma_1} - \frac{f_2 - \mu_2}{\sigma_2}\right). \quad (6)$$

μ_i and σ_i are the mean and standard deviation of the grey values of f_i. CM is the standard deviation of the difference between the two zero–normalized images. CM $= 0$ if the two images are related by $f_1 = af_2 + b$, i.e. if an adaptation to the overall lighting condition is enough for a stable detection. Consequently the higher the complexity measure, the worse the stability of the Harris detector with selection of the N best points. For the varying shutter time image series of Fig. 4, CM takes values of about 0.05 due to noise. The repeatability and false positive rates for the second image series and the two detectors are given in Fig. 5 in dependence on the complexity measure. The proposed method allows a good improvement of the stability, especially for complexer lighting changes. In addition to these results, the stability of the presented detector is reduced when unmodelled phenomena such as e.g. specularities occur.

5. CONCLUSION

In this paper a method is presented to improve the stability of the Harris detector under illumination changes. For this, the ideas of homomorphic filtering are used to reduce the influence of illumination on the filter response. The computation time is only minimally increased in comparison to the original detector. The performance of the proposed method was tested on image series implying different illumination variations. Compared to the Harris detector with selection of the N best points, the detection stability stays good for global lighting changes (e.g. a lighting intensity change) and is improved for complex illumination changes (e.g. a movement of the light source). Due to the simple image formation model used, the detector is not robust when phenomena like specularities or sharp light/shadow patterns appear. This will be object of further research, for example by taking into account color information.

6. REFERENCES

[1] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proc. of the 4th Alvey Vision Conference*, 1988.

[2] David G. Lowe, "Object recognition from local scale–invariant features," in *Proc. of the International Conference on Computer Vision ICCV*, 1999, pp. 1150–1157.

[3] Cordelia Schmid, Roger Mohr, and Christian Bauckhage, "Evaluation of interest point detectors," *International Journal of Computer Vision*, vol. 37, no. 2, pp. 151–172, 2000.

[4] S. M. Smith and J. M. Brady, "SUSAN – a new approach to low level image processing," *International Journal of Computer Vision*, vol. 23, no. 1, pp. 45–78, May 1997.

[5] Flore Faille, "Adapting interest point detection to illumination conditions," in *Digital Image Computing: Techniques and Applications (DICTA 2003)*, Dec. 2003, vol. 1, pp. 499–508.

[6] Theo Gevers and Arnold W. M. Smeulders, "Color–based object recognition," *Pattern Recognition*, vol. 32, pp. 453–464, 1999.

[7] Tinne Tuytelaars, *Local, Invariant Features for Registration and Recognition*, Ph.D. thesis, Katholieke Universiteit Leuven, 2000.

[8] Rafael C. Gonzales and Richard E. Woods, *Digital Image Processing*, Addison-Wesley, 1992.