

Finite Element Laplacian Feature Detector

¹Dermot Kerr, ¹Sonya Coleman, ²Bryan Scotney

¹University of Ulster, Magee, BT48 7JL, Northern Ireland

²University of Ulster, Coleraine, BT52 1SA, Northern Ireland

{d.kerr, sa.coleman, bw.scotney}@ulster.ac.uk

Abstract

Recently, interest point detectors and descriptors have become prominent in the field of computer vision and are typically used to determine correspondences between two images of the same scene. We present a design procedure for the Finite Element Laplacian Feature (FELF) Detector which is similar to the multi-scale approach used in the SURF detector and detects blob like features. We illustrate the accuracy of the FELF algorithm in comparison to well known existing techniques and highlight the computational efficiency of the proposed approach.

1. Introduction

Many local features may be extracted from images; for example, corners correspond to points at which the image data have high curvature. Many existing corner detectors find not only real corner points but also other "interesting points" that may not strictly be recognised as corners [1,2]. However, for some particular applications the ability to detect interesting points that are robust to changes within the image is seen as a more desirable characteristic than specifically detecting only real corner points. Blobs are another interesting type of feature prominent in images; generally blobs can be thought of as regions in an image that are brighter or darker than the surrounding regions. With a detected blob we also have an exact point or pixel position relating to the blob maximum or minimum that may be regarded as an interest point. Blob-like features provide complementary information not obtained from corner detectors [3].

The Laplacian of Gaussian [4] is a popular blob detector that detects blob features in an image. However, a limitation of this detector is that features within an image may appear at many different natural scales depending on what they represent, whereas the detector operates at only a single fixed scale. In order to deal with the natural scales at which features may be present, multi-scale detectors have been developed. A multi-scale Laplacian of Gaussian detector may be achieved by appropriately adjusting the size of the Laplacian of Gaussian kernel to obtain a set of kernels that are then applied to the image. Thus, we obtain a set of features detected at multiple scales. However, by applying a detector at multiple scales we increase the difficulty in matching features, as the same feature may be represented at multiple scales [3]. A scale invariant approach seems more appropriate, where the characteristic scale of the feature is identified. This characteristic scale is the scale that best represents the scale of the feature, and it is not related to the resolution of the image, but rather the underlying structure of the de-

tected feature [3]. By using an operator to measure the response of the same interest point at different scales, the scale at which the peak response is obtained can be identified as the characteristic scale.

The Difference of Gaussian detector [2] approximates the Laplacian of Gaussian detector by computing the difference between two Gaussian smoothed images. This approach was used in the SIFT detector [2] to compute an efficient scale space pyramid by sub-sampling images and convolving with differently sized kernels. Maxima and minima are determined by examining the response from the Difference of Gaussian function in the 9-pixel neighbourhood on the same scale level, and then by examining the response at the scale level above and the scale level below. A similar approach is used in the Hessian-Laplace blob detector [5], where second order Gaussian smoothed image derivatives are used to compute the Hessian matrix. This matrix captures the important properties of the image structure. Using a multi-scale approach where kernel sizes are increased, the trace and the determinant of the Hessian matrix are thresholded and blob features detected. The SURF detector [6] uses integral images [7] to provide a multi-scale approach in which simple box filter kernels can be scaled efficiently. An approximate determinant of the Hessian is used to localise features both in space and scale.

Further detectors based on the SURF detector have been proposed including the FESID detector [8] that uses a first order differential operator with combined smoothing to detect corner regions, and a second order differential operator with an approximate determinant of the Hessian to detect the characteristic scale. In this paper we present a finite element based blob detector based on techniques used in the FESID detector [8] and ideas from the SURF detector. In Section 2 we describe the proposed multi-scale detector and in Section 3 performance is evaluated with respect to repeatability and feature matching using the evaluation techniques presented in [5], highlighting improvements when compared with other well known interest point detectors and descriptors. We also provide examples of comparative computational performance evaluation with other well known interest point detectors in Section 4 and provide a summary in Section 5.

2. Finite element detector

The Finite Element Laplacian Feature (FELF) detector uses finite element based second order derivative operators to detect blob-like features that are robust to various transformations. A multi-scale approach is developed similar to that presented in the SURF detector where integral images are embedded for speed and mask sizes

are incremented to produce a multi-scale operator.

2.1 Second-order operator

The second order differential operator used in the blob detector is developed through the use of the finite element framework using a virtual mesh with nodes placed at pixel centres. Using a neighbourhood centred on node i , a piecewise linear basis function and a Gaussian test function, the second order directional derivative functional is defined by

$$Z_i^\sigma(U) = \int_{\Omega_i^\sigma} \underline{b}_i \cdot \nabla U \underline{b}_i \cdot \nabla \psi_i^\sigma d\Omega_i \quad (1)$$

where $\underline{b}_i = (b_{i1}, b_{i2})^T$ is a locally constant unit vector, with $b_{i1}^2 + b_{i2}^2 = 1$; b_{i1} and b_{i2} are each chosen to be along the x - or y -directions depending on the operator being constructed. The image function u is approximated by a function $U(x, y) = \sum_{j=1}^N U_j \phi_j(x, y)$, where $\phi_i(x, y)$, $i = 1, \dots, N$ are piecewise linear basis functions defined on a triangular mesh such that

$$\phi_i(x_j, y_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \quad i = 1, \dots, N; j = 1, \dots, N \quad (2)$$

is centred on node i and (x_i, y_i) are the coordinates of the nodal point j . In equation (1) ψ_i^σ is a Gaussian test function restricted to a neighbourhood Ω_i^σ surrounding node i . We construct two operators Dxx_{ij}^σ and Dyy_{ij}^σ representing the second order derivatives in the x - and y -coordinate directions respectively:

$$Dxx_{ij}^\sigma = \int_{\Omega_i^\sigma} \frac{\partial \phi_j}{\partial x} \frac{\partial \psi_i^\sigma}{\partial x} dxdy, \quad i, j = 1, \dots, N \quad (3)$$

$$Dyy_{ij}^\sigma = \int_{\Omega_i^\sigma} \frac{\partial \phi_j}{\partial y} \frac{\partial \psi_i^\sigma}{\partial y} dxdy, \quad i, j = 1, \dots, N \quad (4)$$

The integrals are computed as sums of individual element integrals and are computed only over the neighbourhood Ω_i^σ , rather than the entire image domain Ω as ψ_i^σ has support restricted to Ω_i^σ .

2.2 Image representation

In order to provide an efficient image representation, the finite element blob detectors have incorporated the use of integral images introduced by Viola and Jones [7]; more recently integral images have been a key aspect of the SURF detector and have previously been successfully used with the FESID detector [8], as integral images provide a means of fast computation using small convolution filters.

If an intensity image is represented by an array of $n \times n$ samples of a continuous function $u(x, y)$ of image intensity on a domain Ω , then the integral image value $I_\Sigma(\mathbf{x})$ at a pixel location $\mathbf{x} = (x, y)$ is the sum of all pixel values in the original image I within a rectangular area formed by the origin of the image and location \mathbf{x} , and can be described as,

$$I_\Sigma(\mathbf{x}) = \sum_{s=0}^{i \leq x} \sum_{t=0}^{j \leq y} I(s, t) \quad (5)$$

The time and number of operations required to compute the image sum over any rectangular area of the integral image is independent of the size of that region, as four memory reads and three additions are required to compute this region, or indeed any rectangular region regardless of its size.

2.3 Finite element Laplacian detector

An integral image is first constructed using the method outlined in Section 2.2 and the blob detection stage is then performed. Using the same multi-scale approach as the SURF detector we select the first filter size of a 9×9 pixel region. In our approach we partition the 9×9 pixel region slightly differently as illustrated in Figure 1.

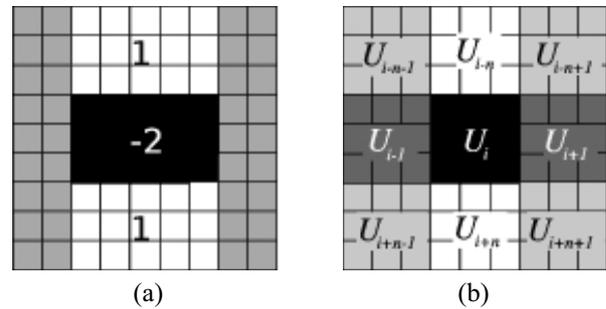


Figure 1. 9×9 filter partitioning for (a) SURF and (b) FELF detector

This approach differs from the SURF detector in that we need to compute 9-regions for each operator, rather than the 3 or 4 regions that are computed with the SURF detector. The filter partitioning allows the operator values to be simply mapped to the appropriate 3×3 region on the 9×9 filter. The sum of the Dxx_{ij}^σ and Dyy_{ij}^σ operator masks is computed, resulting in a Laplacian operator, and these values are appropriately mapped to each of the 9×9 filter regions. The interest point blob strength is then represented by convolution of this Laplacian filter with the intensity values for the filter regions, normalised to take account of the size of the filter used, i.e. in this case, 9×9 . This normalised response indicates the blob response at that particular spatial location with the absolute Laplacian value indicating the blob strength and with the local neighbourhood maxima representing the blob centre or interest point location.

Similarly, blob responses are computed over further scales by increasing the overall size of the filter, but maintaining the 9 regions. For example, within the first octave filter sizes of 9×9 , 15×15 , 21×21 , and 27×27 are used, and each of these filters has 9 individual regions of size 3×3 , 5×5 , 7×7 , and 9×9 respectively. The blob response is computed over a total of 4 octaves that each contains 4 scale ranges. Blob responses that are not maxima or minima in the immediate neighbourhood of the selected blob are rejected by examining a $3 \times 3 \times 3$ neighbourhood (two spatial dimensions and a scale dimension) around the selected blob. Due to the large filters used to compute the blob locations any remaining blobs that are above a threshold are interpolated in 3D to accurately identify the exact blob spatial location.

3. Experimental Results

Evaluation of FELF was performed using the set of test images and testing software provided from the collaborative work between Katholieke Universiteit Leuven, Inria Rhone-Alpes, Visual Geometry Group and the Center for Machine Perception and available for download [9]. In the evaluation, the detectors used for comparison with FELF are limited to those that are most similar in terms of operation. A full evaluation, using the same software and images, of the different detectors used here for comparison has been carried out in [5], and the reader is referred to this work for full details.

Two important parameters characterise the performance of a feature detector, the average number of corresponding features detected in images under different geometric and photometric transformations, both in absolute (corresponding features) and relative terms (feature repeatability i.e. percentage-wise), and the accuracy and localisation of those detected features. The repeatability metric, first introduced in [3], explicitly compares the geometrical stability of detected interest points between different images of a scene under different viewing conditions. The test image set consists of real structured and textured images of various scenes, with different geometric and photometric transformations such as viewpoint change, image blur, illumination change, scale, rotation and image compression. We have performed comparative evaluation with the SURF detector and Harris-Laplace detector using the complete image dataset with different geometric and photometric transformations (for example see Figure 2, or [5] for full details).

For the detectors presented here we describe a circular region with a diameter that is $3\times$ the detected scale of the interest point, similar to the approach in [3, 5]. The overlap of the circular regions corresponding to an interest point pair in a set of images is measured based on the ratio of intersection and union of the circular regions. Thus, where the error in pixel location is less than 1.5 pixels, and the overlap error is below 60%, similar to the evaluation of the SURF detector [6], the interest points are deemed to correspond. For more information on how the detected regions are measured the reader is referred to [7].



Figure 2. Example scale and rotation change sequence

In Figure 3(a) and 3(b) we present comparative evaluation of the detectors on the textured viewpoint change scene. The FELF detector shows improved repeatability over the other detectors and shows a similar number of corresponding regions to the SURF detector. Evaluation with the blur change on a structured scene is presented in Figure 3(c) and Figure 3(d) where the FELF detector performs best in terms of corresponding regions although not as well in the repeatability measure. In Figure 3(e) and Figure 3(f) the repeatability and number of corresponding

regions are compared for the illumination change sequence, where the FELF detector performs best in the number of corresponding regions and shows similar performance for the repeatability score. Finally, Figure 3(g) and Figure 3(h) show the repeatability and number of corresponding regions for the illumination change scene; again FELF performs similarly to SURF for repeatability

4. Computational Comparison

We have compared the times to process a single image (800×640 pixels) using the Harris-Laplace, SURF, and FELF detectors. In the case of Harris-Laplace the detector used is the authors' own implementation. In the case of the SURF detector the version used is that supplied in the OpenCV library [10]. In the case of the FELF detector the results are computed using optimised code. The computation times are presented in Table 1, with the times computed by averaging 5 runs of each detector on an Intel Core 2 Duo 3.00GHz CPU with 2Gb RAM.

Table 1. Computation times

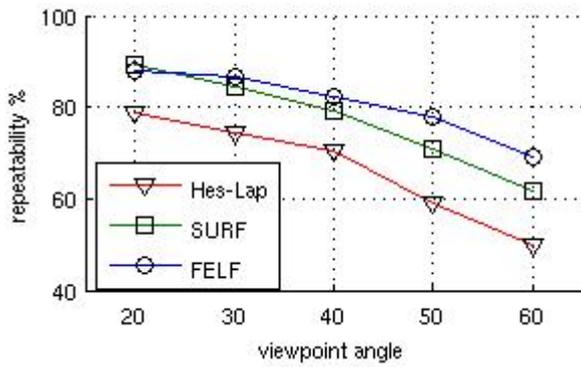
Method	Time (s)	FPS
Hessian-Laplace	0.871	1.148
SURF (inc SURF-E descriptors)	0.636	1.572
FELF (inc SURF-E descriptors)	0.541	1.848

5. Conclusions and Future Work

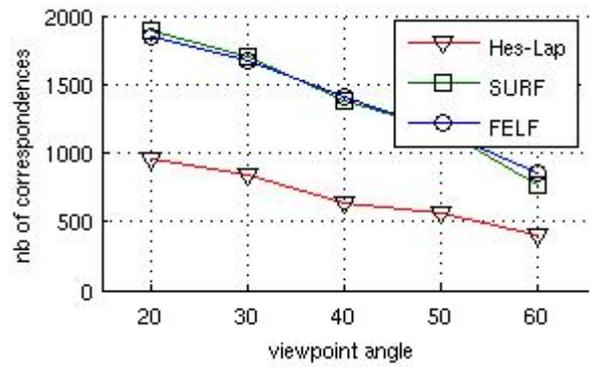
The results indicate that the FELF detector has improved performance in terms of computation time and is generally similar to SURF in repeatability. In some sequences such as the structured blur change the SURF detector performs better than the FELF detector in terms of the repeatability measure. This is likely to be due to the fact that the SURF detector uses three derivative approximations rather than the one derivative approximation used in the FELF detector.

References

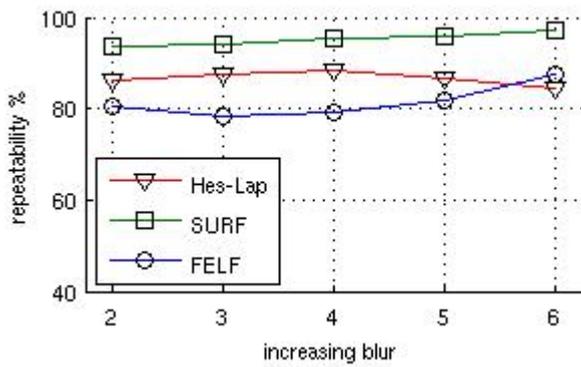
- [1] Tuytelaars, T., Mikolajczyk, K., "Local invariant feature detectors: a survey", *Foundations and Trends in Computer Graphics and Vision*, 177-280, 3, 3, 2008.
- [2] Lowe, D.G., "Distinctive Image Features from Scale-Invariant Keypoints", *IJCV*, 91-110, 60, 2, 2004.
- [3] Mikolajczyk, K. and Schmid, C., "Scale & Affine Invariant Interest Point Detectors", *IJCV*, 60, 63-86, 2004.
- [4] Marr, D., Hildreth, E., "Theory of Edge Detection", *Proceedings of the Royal Society of London*, 215-217, 1980.
- [5] Mikolajczyk, K., Tuytelaars, T., Schmid, C., Zisserman, A., Matas, J., Schaffalitzky, F., Kadir, T., Van Gool, L., "A Comparison of Affine Region Detectors", *IJCV*, 65, 43-72, 2005.
- [6] Bay, H., Ess, A., Tuytelaars, T., Van Gool, L., "Speeded-Up Robust Features" *CVIU*, 110, 346-359, 2008.
- [7] Viola, P., Jones, M., "Rapid object detection using a boosted cascade of simple features", *CVPR*, 1, 511-518, 2001.
- [8] Kerr, D., Coleman, S. and Scotney, B., "FESID: Finite Element Scale Invariant Detector", *Proc. of ICIAAP, LNCS 5716*, pp.72-81, 2009.
- [9] <http://www.robots.ox.ac.uk/~vgg/research/affine/>
- [10] Bradski, G., "The OpenCV Library", *Dr. Dobb's Journal of software Tools*, 2000.



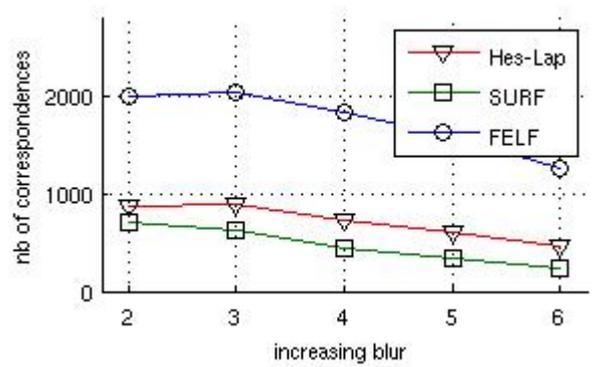
(a) Repeatability for viewpoint change (textured)



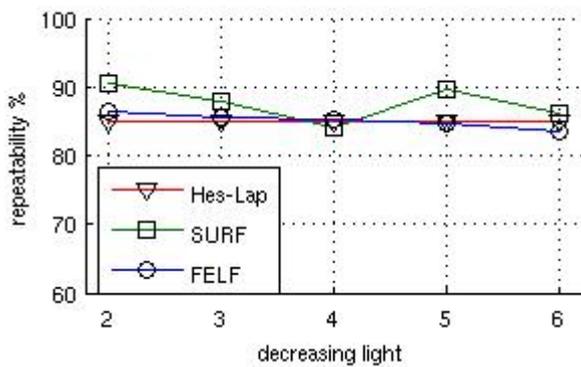
(b) Correspondences for viewpoint change (textured)



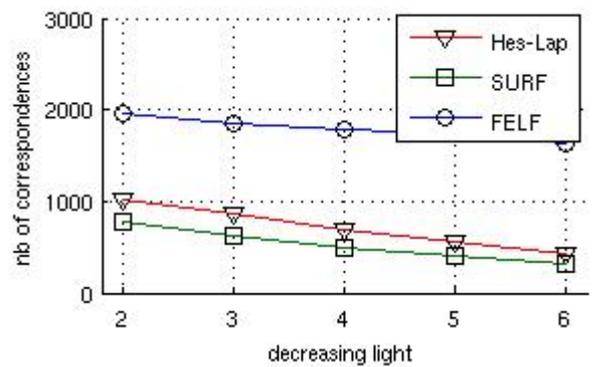
(c) Repeatability for blur change (structured)



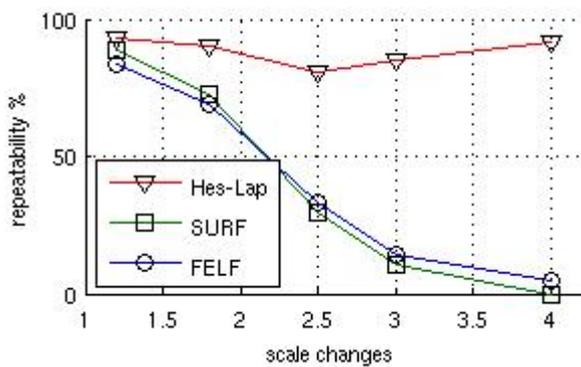
(d) Correspondences for blur change (structured)



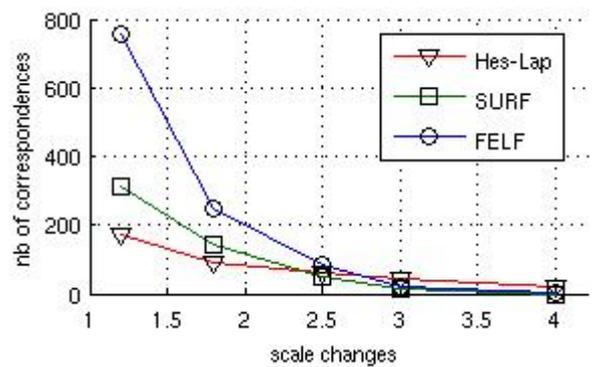
(e) Repeatability for illumination change



(f) Correspondences for illumination change



(g) Repeatability for scale change



(h) Correspondences for scale change

Figure 3. Repeatability score and number of corresponding regions for Wall sequence (a) & (b), Bikes sequence (c) & (d), Leuven sequence (e) & (f), and Bark sequence (g) & (h).