FESID: Finite Element Scale Invariant Detector

Dermot Kerr[1], Sonya Coleman[1], and Bryan Scotney[2]

[1] School of Computing and Intelligent Systems, University of Ulster,
Magee, BT48 7JL, Northern Ireland
[2] School of Computing and Information Engineering, University of Ulster, Coleraine,
BT52 1SA, Northern Ireland

Abstract. Recently, finite element based methods have been used to develop gradient operators for edge detection that have improved angular accuracy over standard techniques. A more prominent issue in the field of image processing has become the use of interest point detectors and to this end we expand upon this research developing a finite element scale invariant interest point detector that is based on the same multi-scale approach used in the SURF detector. The operator differs in that the autocorrelation matrix is used to select the interest point location and the derivative and smoothing operations are combined into one operator developed through the use of the finite element framework.

1 Introduction

Local features such as corners or interest points have been used to find point correspondences between images as they are well suited to matching and robust to occlusion, clutter and content change [1,2]. The process of finding correspondences can be separated into three distinct stages: selection of interest points, description of interest points and finally, matching of interest points between different images. Here we focus on the first stage only, interest point detection.

Standard corner detectors, such as the Harris and Stephens corner detector [3], find points in an image at one particular scale, but points within an image may occur at many natural scales depending on what they represent. In order to deal with the natural scales at which features may be present recent corner detectors have been developed to work on multiple scales, thereby having the ability to detect all points no matter what scale they are represented at [1,2,4,5,6]. However, by representing the same feature at many scales we increase the difficulty of matching the interest points due to the increased numbers of points. A scale invariant approach seems more appropriate where the *characteristic scale* of the underlying 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, rather the underlying structure of the detected feature [6]. 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.

Many interest point detectors appear in the literature, and indeed many detectors do not even detect interest points, rather interesting regions corresponding

to features such as blobs, edges, or interest points. One example is the Harris-Laplace [6] that expands on the Harris and Stephens corner detector [3]. This detector uses a multi-scale approach to find Harris and Stephens corner points, which are then refined using a Laplacian scale selection to select the characteristic scale. The Hessian-Laplace detector [7] uses a similar multi-scale approach, although uses second-order derivatives for interest point selection leading to detection of blob and ridge type features. One of the most popular scale invariant detectors in the literature is the SIFT detector [5]. This implementation is very efficient where the Laplacian of Gaussian is approximated using a Difference of Gaussians. More recently the SURF detector [1] has been developed combining techniques from the Hessian-Laplace detector, the SIFT detector and using an novel image representation technique enabling improved computational performance, even over the already efficient SIFT detector.

In this paper we propose a scale invariant feature detector using the Linear^2-Gaussian corner detector [8,9,10] for interest point selection. This corner detector, developed through the use of the finite element framework, combines the separate derivative calculation and smoothing operations common to corner detectors such as the Harris and Stephens corner detector. We incorporate techniques from the Harris-Laplace detector and the SURF detector, allowing detection of point-like features rather than blob-like features, ensuring accuracy and efficient computation. In Section 2 we detail the proposed scale invariant detector, including construction of the point selection, and scale selection operators developed using the finite element framework. In Section 3 we perform comparative evaluation and provide a summary in Section 4.

2   A Finite Element Based Detector

The Harris-Laplace detector [2] uses a scale adapted Harris and Stephens corner detector to select the spatial location of interest points, and the scale adapted Laplacian of Gaussian to determine the peak response scale. Detected interest points are then refined using finer grained scales to obtain the interest point location where the Laplacian of Gaussian is maximum in scale. The SURF detector uses integral images with the Hessian matrix to determine interest point location. This approach allows for the detection of blob structures at a location where the determinant of the Hessian matrix is maximum. The finite element scale invariant detector (FESID) combines techniques from the Harris-Laplace detector and the SURF Fast Hessian detector. FESID uses the Linear^2-Gaussian corner detector [8,9,10] to detect interest points using a similar multi-scale approach to that used in the SURF detector. Selected interest points are then refined by rejecting points where the point scale (measured using a second-order finite element Linear-Gaussian operator) are not maximum. The integral image approach [11] used in the SURF detector is incorporated to increase computational efficiency.

2.1   First Order Linear^2-Gaussian Operator

The Linear^2-Gaussian operator [8] is developed through the use of the finite element framework and in the FESID detector is used to detect the location of the interest points. In [12], near circular Gaussian-based image derivative operators have be developed via the use of a virtual mesh and are proven to reduce angular errors when detecting edges over a range or orientations and in [9] the Linear^2-Gaussian operator was shown to have comparable performance with the Harris and Stephens corner detector. The Linear^2-Gaussian operator is based on the first directional derivative $\partial u/\partial b \equiv \underline{b} \cdot \underline{\nabla} u$ which measures a product of the directional derivative is approximated by the functional

$$H_i^\sigma (U) = \int_{\Omega_i^\sigma} (\underline{b}_i \cdot \underline{\nabla} U)^2 \, \psi_i^\sigma \, d\Omega_i. \tag{1}$$

We consider a locally constant unit vector $\underline{b}_i = (b_{i1}, b_{i2})^T$, where $b_{i1}^2 + b_{i2}^2 = 1$; $\underline{b}_{i1}$ and $\underline{b}_{i2}$ 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 $\phi_i(x,y)$, $i = 1, \ldots, N$, are piecewise linear basis functions defined on a triangular mesh such that

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

is centred on node i, and (x_i, y_i) are the coordinates of the nodal point j. In equation (1) $\psi_i^\sigma \in T_\sigma^h$ is a Gaussian test function confined to a neighbourhood Ω_i^σ surrounding node i that enables Gaussian smoothing to be built into the neighbourhood operator. We refer to the use of combination of two piecewise linear basis functions and one Gaussian test function as the Linear^2-Gaussian operator.

We use the Harris and Stephens [3] cornerness measure to measure the strength of the interest point thus it is required to compute the smoothed x^2, y^2, and xy-derivative values. As the Linear^2-Gaussian operator combines two piecewise linear basis functions and one Gaussian test function it is possible to directly compute a smoothed product of the directional derivatives. For example, in the case of H_{xy}^σ we choose $\underline{b}_{i1}$ and $\underline{b}_{i2}$ to be along the $x-$ and $y-$coordinate directions respectively; and for $H_{x^2}^\sigma$ and $H_{y^2}^\sigma$ we choose $\underline{b}_{i1} = \underline{b}_{i2}$ along the $x-$ and $y-$coordinate directions respectively. Thus, we obtain directly measures of the products of the directional derivatives rather than from the product of measures of the directional derivatives. These can be defined by

$$K_{ij}^\sigma = \int_{\Omega_i^\sigma} \frac{\partial \phi_j}{\partial x} \frac{\partial \phi_j}{\partial x} \psi_i^\sigma \, dx dy, \; i,j = 1, \ldots, N \tag{3}$$

and

$$L_{ij}^\sigma = \int_{\Omega_i^\sigma} \frac{\partial \phi_j}{\partial y} \frac{\partial \phi_j}{\partial y} \psi_i^\sigma \, dx dy, \; i,j = 1, \ldots, N \tag{4}$$

and

$$M_{ij}^{\sigma} = \int_{\Omega_i^{\sigma}} \frac{\partial \phi_j}{\partial x} \frac{\partial \phi_j}{\partial y} \psi_i^{\sigma} \, dxdy, \ i,j = 1, \ldots, N. \tag{5}$$

The integrals are sums of the 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}. Thus, by using two piecewise linear basis functions along with a Gaussian basis function, we generate the non-linear operators $H_{x^2}^{\sigma}$, $H_{y^2}^{\sigma}$, and H_{xy}^{σ}.

2.2   Second Order Linear-Gaussian Operator

A finite element based approach is again used to define and construct the second-order derivative Linear-Gaussian operators, again using a neighbourhood centred on node i and comprised of a set S_i^{σ} of elements, and supporting a Gaussian test function ψ_i^{σ} centred on the central node i, similar to the approach presented in Section 2.1 only using one piecewise linear basis function and one Gaussian test function. To construct the second order Linear-Gaussian operator, a second order directional derivative functional similar to that defined in equation (1) is used

$$Z_i^{\sigma}(U) = \int_{\Omega_i^{\sigma}} \underline{b}_i \cdot \underline{\nabla} U \cdot \underline{\nabla} \psi_i^{\sigma} \, d\Omega_i \tag{6}$$

Again we choose $\underline{b}_{i1}$ and $\underline{b}_{i2}$ to be along the $x-$ and $y-$coordinate directions depending on the the operator being constructed. Here we construct the operators N_{ij}^{σ}, O_{ij}^{σ}, and P_{ij}^{σ} representing the second order x-, y-, and xy-coordinate directions respectively.

$$N_{ij}^{\sigma} = \int_{\Omega_i^{\sigma}} \frac{\partial \phi_j}{\partial x} \frac{\partial \psi_i^{\sigma}}{\partial x} \, dxdy, \ i,j = 1, \ldots, N \tag{7}$$

and

$$O_{ij}^{\sigma} = \int_{\Omega_i^{\sigma}} \frac{\partial \phi_j}{\partial y} \frac{\partial \psi_i^{\sigma}}{\partial y} \, dxdy, \ i,j = 1, \ldots, N \tag{8}$$

and

$$P_{ij}^{\sigma} = \int_{\Omega_i^{\sigma}} \frac{\partial \phi_j}{\partial x} \frac{\partial \psi_i^{\sigma}}{\partial y} \, dxdy, \ i,j = 1, \ldots, N. \tag{9}$$

2.3   Image Representation

In order to provide an efficient image representation we have adopted the approach of integral images introduced by Viola and Jones [11]; more recently integral images have been a key aspect of the Speeded-Up Robust Features (SURF) detector [13]. Using integral images provides a means of fast computation when 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,

$$I_\Sigma(\mathbf{x}) = \sum_{i=0}^{i \le x} \sum_{j=0}^{j \le y} I(i, j). \tag{10}$$

The time and number of operations required to compute 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.

2.4   Finite Element Scale Invariant Detector Operation

An integral image is constructed using the method outlined in Section 2.3 and the the interest point detection stage is then performed. Using the same multiscale approach as the SURF detector we first select the first filter size of 9×9. The 9×9 filter is partitioned slightly different from the approach used in the SURF detector as illustrated for the case of the y-direction filter in Figure 1.
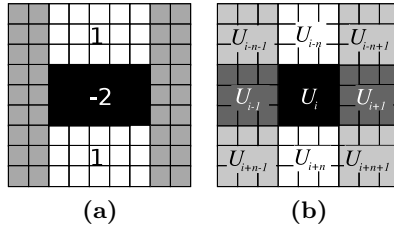
Fig. 1. 9×9 filter partitioning for (a) SURF and (b) FESID

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 operator values, mapped from the Linear²-Gaussian operator values obtained in Section 2.1, are then convolved with the sum of the pixel intensities from each of the areas U_{i-n-1}, U_{i-n}, U_{i-n+1}, U_{i-1}, U_i, U_{i+1}, U_{i+n-1}, U_{i+n}, and U_{i+n+1}, illustrated in Figure 1(b). The filter outputs are then used to compute the interest point strength using the function

$$R = \left(H_{U_{x^2}} H_{U_{y^2}} - \left(H_{U_{xy}} \right)^2 \right) - k \left(H_{U_{x^2}} + H_{U_{y^2}} \right)^2, \tag{11}$$

over the complete integral image using the 9×9 box filter. This is repeated for all the scales using the same approach as the SURF detector, where for example filter sizes of 9×9, 15×15, 21×21, 27×27 are used within the first octave, creating a set of interest point strengths over the full range of scales.

Using the second-order operators from Section 2.2, the interest point scale strength is computed using a similar approach. Where an interest point is deemed to be present the scale strength is determined using the same method as the SURF detector, by constructing the Hessian matrix and computing the normalised determinant,

$$\mathbf{H} = \begin{bmatrix} H_{xx} & H_{xy} \\ H_{xy} & H_{yy} \end{bmatrix}. \tag{12}$$

$$\mathrm{Det}\,(H_{approx}) = H_{xx}H_{yy} - (0.9 \times H_{xy})^2 \tag{13}$$

where the value 0.9 in equation 13 is used to balance the Hessian matrix determinant [13]. Again, this is repeated for each of the octaves creating a set of interest point scale responses over the full range of scales. Interest points that are not local maxima in a $3 \times 3 \times 3$ neighbourhood or above pre-defined thresholds for both the interest point strength and the interest point scale response are rejected. Interest points are interpolated in 3D to localise the interest point's location, where rather than fitting the selected interest points with a 3D

Fig. 2. Example detected FESID interest points using the graffiti image [7]

quadratic function [5,13] to provide sub-pixel maxima, we simply interpolate in each direction $x-$, $y-$, and scale separately.

A circular region with a diameter that is $3\times$ the detected scale is then used to define the region around the interest point for computing descriptors, in a manner similar to the Harris-Laplace and the SURF detectors. Figure 2 illustrates an example image with detected interest points; in this image the threshold has been set higher than usual to limit the number of displayed points for visual clarity. It is interesting to note the Linear^2-Gaussian operator's susceptibility to detecting points around edge features as shown on the image.

3   Experimental Results

Evaluation of FESID 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[1]. In the evaluation, the detectors used for comparison with FESID are limited to those that are most similar in terms of operation. A full evaluation of the different detectors using the same software and images has been carried out in [1,7], and the reader is refered to this work for full details.

Using the repeatability metric, first introduced in [14], we explicitly compare the geometrical stability of detected interest points between different images of a scene under different viewing conditions. In the FESID detector we describe a circular region with a diameter that is $3\times$ the detected scale of the interest point, similar to the approach in [7,13]. 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[13], the interest points are deemed to correspond. For more information on how the detected regions are measured the reader is refered to [7].

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 and rotation and image compression. We have performed comparative evaluation with the SURF detector, Harris-Laplace detector, and Hessian-Laplace detector using the complete image dataset with different geometric and photometric transformations.

Figure 3 illustrates the repeatability of corresponding regions detected in each of the image sets. In all cases the thresholds used are those provided by the authors of the original detectors in the testing software, and in the case of the finite element scale invariant detector (FESID), a fixed threshold of 75 is used for the cornerness response, and a threshold of 50 is used for the determinant

---

[1] The images and testing software can be downloaded from:
   http://www.robots.ox.ac.uk/~vgg/research/affine/
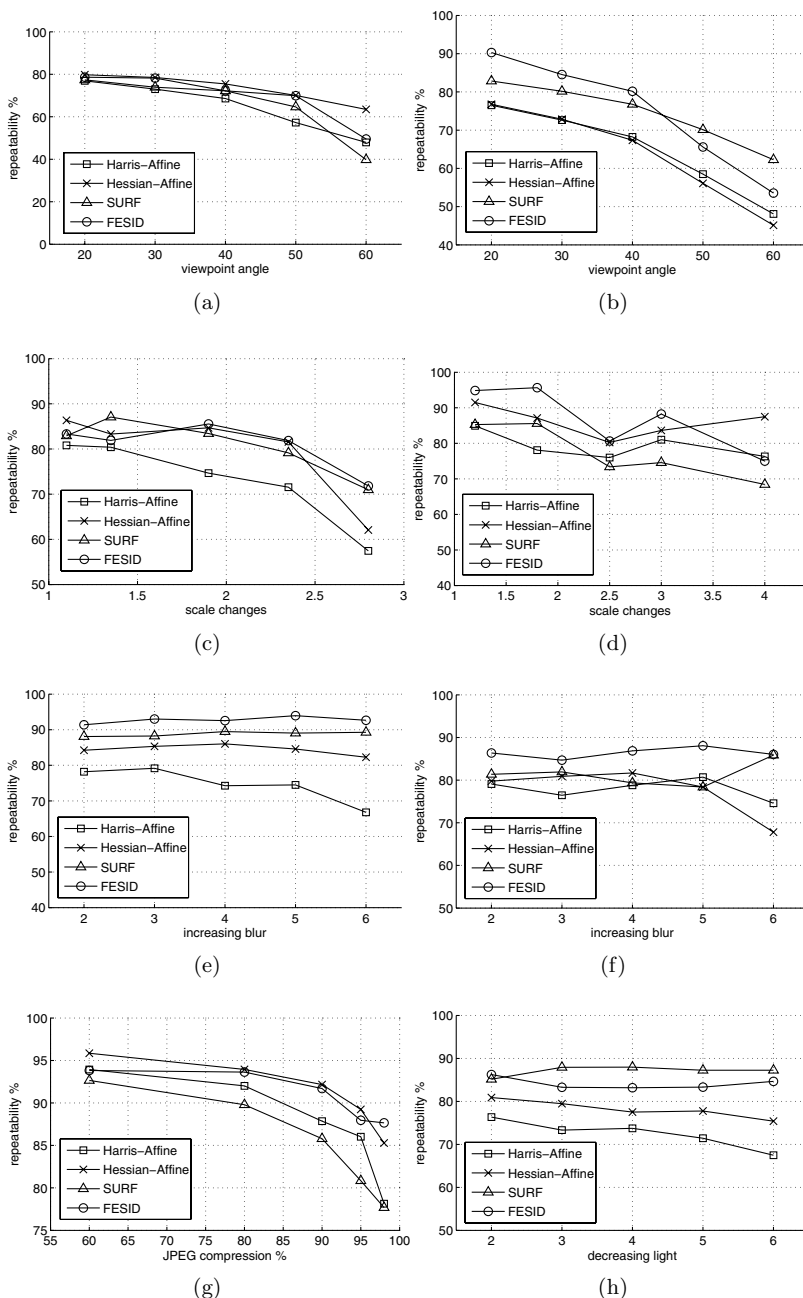
Fig. 3. Repeatability score for image sequences: (a) Graffiti - viewpoint change, (b) Wall - viewpoint change, (c) Boat - zoom and rotation, (d) Bark - zoom and rotation, (e) Bikes - image blur, (f) Trees - image blur, (g) UBC - JPEG compression, and (h) Leuven - illumination change

response over the complete image set. The threshold values where selected to achieve a similar number of interest points as the SURF detector.

4   Summary

In this paper we have introduced a finite element scale invariant interest point detector, presenting the finite element method for constructing first- and second-order derivative operators that are used for interest point selection and scale selection. We describe the finite element scale invariant interest point detector algorithm and show that the Harris and Stephens cornerness measure is used for spatial localisation of the interest point, and the determinant of the Hessian matrix is used for scale selection.

Use of the Linear^2-Gaussian corner detector with its built-in smoothing provides advantages over the standard Harris and Stephens detector. Comparative evaluation of FESID has been performed with other well known interest point detectors using the set of images and testing software developed in [7]. The evaluation illustrated that in most cases the finite element detector had improved or comparable repeatability rates and detected a similar or greater number of corresponding regions compared with other well known interest point detectors.

References

1. Bay, H., Ess, A., Tuytelaars, T., Gool, L.V.: Speeded-Up Robust Features (SURF). CVIU 110(3), 346–359 (2008)
2. Mikolajczyk, K., Schmid, C.: Scale & Affine Invariant Interest Point Detectors. IJCV 60(1), 63–86 (2004)
3. Harris, C., Stephens, M.: A combined corner and edge detector. In: Proceedings of Alvey Vision Conference, vol. 15, pp. 147–151 (1988)
4. Dufournaud, Y., Schmid, C., Horaud, R.: Matching images with different resolutions. In: Proceedings of IEEE CVPR, vol. 1, pp. 612–618 (2000)
5. Lowe, D.G.: Distinctive Image Features from Scale-Invariant Keypoints. IJCV 60(2), 91–110 (2004)
6. Mikolajczyk, K., Schmid, C.: Indexing based on scale invariant interest points. In: Proceedings of ICCV, vol. 1, pp. 525–531 (2001)
7. 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(1), 43–72 (2005)
8. Coleman, S., Kerr, D., Scotney, B.: Concurrent Edge and Corner Detection. In: Proceedings of IEEE ICIP, pp. 273–276 (2007)
9. Kerr, D., Coleman, S., Scotney, B.: Near-Circular Corner and Edge Detection Operators. In: Proceedings of IEEE IMVIP, pp. 7–14 (2007)
10. Coleman, S., Scotney, B., Kerr, D.: Integrated edge and corner detection. In: Proceedings of ICIAP (2007)
11. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. CVPR 1, 511–518 (2001)

12. Scotney, B., Coleman, S.: Improving angular error via systematically designed near-circular Gaussian-based feature extraction operators. Pattern Recognition 40(5), 1451–1465 (2007)
13. Bay, H., Tuytelaars, T., Van Gool, L.: Surf: Speeded up robust features. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006. LNCS, vol. 3951, pp. 404–417. Springer, Heidelberg (2006)
14. Schmid, C., Mohr, R., Bauckhage, C.: Comparing and Evaluating Interest Points. In: Proceedings of ICCV, pp. 230–235. IEEE Computer Society Press, Los Alamitos (1998)