INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH & DEVELOPMENT

ISSN 2278 – 0211 (Online)

A Review on Various Approaches of Feature Extraction for Image Matching

Jayant B. Karanjekar
Assistant Professor, WCEM, Gumgaon, Nagpur, India
Bhavna Naukarkar
Assistant Professor, WCEM, Gumgaon, Nagpur, India

Abstract:
In this paper we present a review on different approaches of feature extraction for subsequent image matching and the literature over the past in the field of feature detection methods. The feature detector and descriptors are needed for the feature extraction process. We take an overview on the various types of feature detectors. The image matching is difficult because the images are captured from different viewpoint, scale, lighting condition, etc. The MSER as a feature detector and SIFT as a feature descriptor obtain good score in most of cases.

Keywords*: Feature extraction, Image matching, Local Features, Affine region detector*

1. Introduction

Image matching is one of the most challenging and important subject of this era. Because the images are captures from the different viewpoint, scale, illumination and so on, it is difficult to find or match the corresponding images from the set of images. When the two images which we want to match are captured from different viewpoint, scale, in different light condition then finding the corresponding regions from both images is the main issue. Occlusion and clutter are also problem in image matching. So it is very important to extract local invariant features for image matching. A local feature [1] is an image pattern of an image which is different from its neighborhood feature. Invariant means the feature which remains constant under various transformations. Many researchers proposed various methods of extracting or detecting invariant features from the image which are invariant to image scale, rotation, occlusion and clutter, viewpoint change, illumination change etc. After the local invariant features detected by the detectors then it is converted into descriptor. The feature descriptor is a code that describe image region, usually given by vector of real number. For the image matching and its applications the feature detectors and feature descriptor is required. The feature detectors which return the same object regions examples are good detectors i.e, detection repeatability; descriptors which match the same object regions between category examples are good descriptors i.e, match count/ratio.

2. Literature Review

Many researchers have done work in the field of feature detection to detect robust feature which can be use in image matching. J. Matas *et al.* [2] introduce the extremal region for the wide baseline image matching and presented a feature detection algorithm for an affine invariant stable subset of extremal region known as maximally stable extremal region [MSER]. The MSERs are the elements of the image which are, closed under affine transformation of image coordinates and invariant to affine transformation of intensity. The extremal means the all pixels inside the MSER region have either higher intensity or lower intensity than that of all pixels at the boundary of the region.

David G. Lowe *et al*. [3] presented a method for extracting distinctive invariant features from images. These features are invariant to image scale and rotation, and can be used in image matching across a substantial range of affine distortion, change in 3D viewpoint, addition of noise, and change in illumination. He also described the SIFT keypoint which are useful due to their distinctiveness, which enables the correct match for a keypoint to be selected from a large database of other key points.

K. Mikolajczyk and C. Schmid *et al.* [4] performed a comparison of the local descriptors. In the comparison the performance of the descriptor is calculated for different type's local region detected by the various feature detectors. The comparison is done between different categories of descriptor like distribution based descriptor, spatial frequency based descriptor which is based on the frequency content of the image, differential descriptor and other descriptor. The comparison includes various descriptors like scale invariant feature transform (SIFT), steerable filters, differential invariants, complex filters, moment invariants etc. The results showed that SIFT obtained best matching score. So their experiment concluded that the SIFT is best descriptor.

Jukka Lankinen *et al*. [5] also presented a comparison of feature detector and descriptor. They took repeatability and matching score as performance parameters but they extended the previous approach of repeatability and matching score [3] to object class level. They performed this comparison against the Caltech-101 dataset. The experiment concluded that the combination of Hessian affine as a feature detector and SIFT as a feature descriptor obtains highest repeatability and matching score.
K. Mikolajczyk *et al*. [6] performed a comparison amongst various affine region detectors. He had done comparison between six different types of affine region detectors which includes Harris-Affine detector, Hessian-Affine detector, MSER (MSER detector), edge-based region (EBR detector), intensity extremal-based region (IBR detector), and salient regions detector. From that comparison it is studied that the performance of the detectors decreases as the viewpoint change increases. Also no detector can perform better than others for all types of scenes and transformation. These detectors have been proposed by number of researchers. The comparison has concluded that the MSER obtains highest score than the other detectors in most cases. The MSER obtains good score on images which have homogeneous regions with distinctive boundaries. The MSER followed by the Hessian – Affine detector followed by the Harris – Affine detector. The Hessian – Affine and Harris – Affine detect more regions in images for image matching where matching regions have occlusion and clutter. Edge Based Region detector is used for the images region which contains intersections of edges. Salient regions detector is suitable for object class recognition. The Salient regions detector obtains lowest score in the comparison. From the comparison it is concluded that the MSER got the highest score in most of the cases.

3. Feature Detection Methods

In this section we have described the various popular feature detection methods.

3.1. Harris – Affine Detector

The Harris – Affine is an affine normalization which detects features in scale space and then determine elliptical region for each feature. This detector is based on the Hessian Matrix. The scale-selection is based on the Laplacian, and the shape of the elliptical region is determined with the second moment matrix of the intensity gradient proposed by T. Lindeberg and J. Garding *et al.* [7, 8]. This matrix describes the gradient distribution in a local neighborhood of a point:

$$M = \sigma_D^2 g(\sigma_I) * \begin{bmatrix} I_x^2(x,\sigma_D) & I_x I_y(x,\sigma_D) \\ I_x I_y(x,\sigma_D) & I_y^2(x,\sigma_D) \end{bmatrix} \quad (1)$$

The eigenvalues of this matrix represent two principal signal changes in a neighborhood of the point. This detector finds point for which both curvatures are significant. The features are stable in arbitrary lightning conditions.

3.2. Hessian – Affine Detector

The Hessian – Affine is similar to the Harris – Affine detector. This detector is based on the Hessian matrix:

$$H = H(x,\sigma_D) = \begin{bmatrix} I_{xx}(x,\sigma_D) & I_{xy}(x,\sigma_D) \\ I_{xy}(x,\sigma_D) & I_{yy}(x,\sigma_D) \end{bmatrix} \quad (2)$$

The matrix uses the second derivative which gives strong response to the image which contains blobs and ridges. This detector finds the regions with local maximum of the determinant.

3.3. Edge Based Region Detector

The edge based region detects regions in an image by using the edges present in the image. The points at which image brightness changes sharply are typically organized into a set of curved line segments termed edges. Moreover, by using the edges, the dimensionality of the problem can be significantly reduced. The 6D search problem over all possible affinities can be reduced to a one-dimensional problem by exploiting the nearby edges geometry. In this detector we have to start with Harris corner point p proposed by Harris and Stephans *et al.*[9] and neighborhood edges extracted with canny edge detector proposed by Canny *et al.*[10]. The regions are extracted at various scales to increase the robustness.

3.4. Intensity Based Region Detector

This detector starts from an intensity extrema and studies the intensity pattern along rays emanating from this point and forms arbitrary region which then replaced by elliptical region. The following function is evaluated along each ray:

$$f(t) = \frac{abs(I(t) - I_0)}{\max\left(\frac{\int_0^t abs(I(t) - I_0)}{t}, d\right)} \quad (3)$$

With t an arbitrary parameter along the ray, $I(t)$ the intensity at position t, I_0 the intensity value at the extremum, and d a small number which has been added to prevent a division by zero. At point where function reaches extremum is invariant under affine geometric and linear photometric transformations. Typically, a maximum is reached at positions where the intensity suddenly increases or decreases. All points corresponding to maxima of $f(t)$ along rays originating from the same local extremum are linked to form an affine invariant region.

3.5. Maximally Stable Extremal Region detector

A Maximally Stable Extremal Region (MSER) is a connected region of an appropriately thresholded image. 'Extremal' means all pixels inside the MSER have either higher intensity (Bright extremal regions) or lower (Dark extremal regions) intensity than all the pixels on its outer boundary. Detection of MSER is related to thresholding. The concept of MSER [1] can be explained as follow:

- Threshold the image I by gray-levels. We will consider the pixels below a threshold as 'black' and above or equal as white.
- If we see all the thresholded images, we will see first image a white image and at last we will see the black image at end.
- The set of all connected components of all frames is the set of all maximal regions.
- The minimal regions could be obtained by inverting the intensity of I and running the same process.

3.6. Salient Region Detector

The Salient Region Detector was proposed by Kadir and Bardy *et al.* [11] is based on information theory. This detector is based on the pdf of intensity values computed over an elliptical region. Detection proceeds in two steps first, at each pixel x the entropy H of the pdf p(I) is calculated over various scales s.

$$H = -\sum_{I} p(I) \log p(I) \qquad (4)$$

The probability distribution p(I) is based on the intensity distribution in a circular neighborhood of radius *s* around x. Local maxima of the entropy are recorded. These are candidate salient regions. Second, for each of the candidate salient regions the magnitude of the derivative of p(I) with respect to scale *s* is computed as

$$W = \frac{s2}{2s\ -\ 1} \sum_{I} \left| \frac{\partial p(I; s)}{\partial s} \right| \qquad (5)$$

The saliency Y is then computed as

Y = W * M (6)

The candidate salient regions over the entire image are ranked by their saliency *Y*, and the top *P* ranked regions are retained.

4. Performance Evaluation

The most important measure that is used for comparing the detectors is the repeatability rate. The number of correspondences in both images is compared, and the smaller of the numbers is used as minimum when calculating the repeatability. The repeatability R can be calculated as:

$$R = \frac{Corresponding\ regions}{Detected\ regions} * 100\ \% \qquad (8)$$

The repeatability measurement tends to give better results if the number of correspondences is higher. The table shows the repeatability of detectors in most of the cases.

Detector	Repeatability
Harris Affine Detector	15%
Hessian Affine Detector	28%
EBR	15%
IBR	30%
MSER	45%
Salient Region Detector	10%

Table 1

5. Conclusion and Future Work

In this paper we present a review on various approaches of feature detection. It is found that the performance of the detector degrades as the angle of viewpoint increases. So no detector performs better than the other. In most of the cases the MSER obtains highest score as a feature detector and the SIFT obtains highest score as a feature descriptor. So combination of the MSER and SIFT will get good score in feature detection for subsequent image matching. The development of algorithm containing integration of MSER and SIFT for the feature detection will be the part of future area of work.

6. References

1. T. Tuytelaars and K. Mikolajczyk, “Local Invariant Feature Detectors: A Survey” Foundations and Trends in Computer Graphics and Vision Vol. 3, No. 3 (2007) 177–280_c 2008.
2. J.Matas, O. Chum,M. Urban, and T. Pajdla, “Robust wide-baseline stereo from maximally stable extremal regions,” in Proc. Brit. Mach. Vis. Conf., Cardiff, U.K., 2002, pp. 384–393.
3. D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” Int. J. Comput. Vis., vol. 60, no. 2, pp. 91–110, Nov. 2004.
4. K. Mikolajczyk and C. Schmid, “A performance evaluation of local descriptors,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 27, no. 10,pp. 1615–1630, Oct. 2005.
5. Jukka Lankinen, Ville Kangas, and Joni-Kristian Kamarainen “A comparison of local feature detectors and descriptors for visual object categorization by intra class repeatability and matching,” 21th Int. Conf. on Pattern Recognition, 2012.
6. K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, and J. Matas, “A comparison of affine region detectors,” Int. J. Comput. Vis., vol. 65, no. 1/2, pp. 43–72, Nov. 2005.
7. T. Lindeberg, “Direct estimation of affine image deformation using visual front-end operations with automatic scale selection,” in Proceedings of the International Conference on Computer Vision, pp. 134–141, 1995.
8. T. Lindeberg and J. Garding, “Shape-adapted smoothing in estimation of 3-D shape cues from affine deformations of local 2-D brightness structure,” Image and Vision Computing, vol. 15, no. 6, pp. 415–434, 1997.
9. C. Harris and M. Stephens, “A combined corner and edge detector,” in Alvey Vision Conference, pp. 147–151, 1988.
10. J. Canny, “A computational approach to edge detection,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 8, no. 6, pp. 679–698, 1986.
11. T. Kadir and M. Brady, “Scale, saliency and image description,” International Journal of Computer Vision, vol. 45, no. 2, pp. 83–105, 2001