Towards Invariant Interest Point Detection of an Object

Md. Saiful Islam[1], Andrzej Sluzek[2], Zhu Lin[1], Meng Joo Er[2]
School of Computer Engineering[1], Intelligent Systems Centre[2]
Nanyang Technological University
Singapore 639798.
Email: saiful@pmail.ntu.edu.sg , assluzek@ntu.edu.sg , danielin@pmail.ntu.edu.sg, emjer@ntu.edu.sg

ABSTRACT

Detection of some interest points on an object is useful for many applications, such as local shape description of the object, recognition of the object in clutter environment etc. The same object present in different images can have some geometric and photometric transformations with respect to one another. The detection method should be robust to all these transformations. We describe relative scale Harris method for interest point detection. This method is robust to linear geometric transformations. A threshold selection method is also described for invariance to intensity change, partial occlusion and clattered environment. Unlike multi-scale methods our method is fast enough to be suitable for real-time applications.

Keywords: interest point, relative scale, invariance, repeatability

1. INTRODUCTION

A reference object and a test image are given. The reference object present in the test image can have some geometric and photometric transformation, with respect to the given reference object. Moreover, the surrounding of the object of interest could be cluttered with other objects and object itself could be partially occluded. In all the conditions, we like to detect similar sets of interest points on the reference object and on the object of interest in the test image. Harris detector [Har88] is a classical work on interest point detection. Later, it was improved by Schmid et al [Sch00] for better repeatability rate of interest points, in the presence of relative rotation. The detectors fail if there is a large scale change for the object of interest present in the image. To alleviate the problem, Mikolajczyk et al [Mik01], [Mik04] investigated scale-space interest point detectors, using the Harris-Laplacian method. But this method could be computationally quite expensive for real time applications.

In this paper, we introduce relative scale Harris method for scale invariant interest point detection. It is assumed that the relative scales of the same object presents in different images are known *a prior*. For intensity invariant interest points, a heuristic based method of threshold selection is proposed. So, our detection method is robust for different kinds of geometric (translation, rotation and scale change) and photometric transformations (intensity scaling and intensity shift). This method is also robust for cluttered environment and partial occlusion. Unlike multi-scale interest point detection methods [Mik01], [Mik04], our method is computationally quite fast. The proposed method of interest point detection could be quite useful for many applications such as local shape description, object recognition in cluttered environment etc.

Our detection method is described in Section 2. Section 3 presents some experimental results to show the effectiveness of the method. Section 4 concludes the paper with future works.

WSCG SHORT papers proceedings, ISBN 80-903100-9-5
WSCG'2005, January 31-February 4, 2005
Plzen, Czech Republic.

2. METHOD

In this section, our goal is to detect two similar sets of points on the same object presents in two images, irrespective of different kinds of transformations. Section 2.1 presents relative scale Harris method for scale and rotation invariant interest point detection. Section 2.2 gives a threshold selection method for intensity invariant interest points. Section 2.3 deals with cluttered environment and partial occlusion.

2.1 Relative scale Harris method

The detection method depends on the **prior knowledge about relative scales** of the object, present in the images. We assign an arbitrary reference scale σ_R to the given reference object. A point (x', y') on the object present in the test image is related to a point (x, y) of the reference object by linear geometric transformations as follows:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} s & 0 \\ 0 & s \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} a \\ b \end{bmatrix}$$

where s is an arbitrary scaling factor, θ is an arbitrary rotation and (a,b) is an arbitrary shift. The **relative scale σ_I** of the object in the test image is a linear function of the scaling factor s.

$$\sigma_I = cs\sigma_R$$

where c is a constant. The value of c depends on the reference scale σ_R and the range of scale of consideration $[\sigma_{min} - \sigma_{max}]$, where $\sigma_{min} < \sigma_R < \sigma_{max}$.

In the original Harris method [Har88], the image is differentiated in two perpendicular directions and then integrated by a circular Gaussian window. In the improved Harris version [Sch00], a 1D Gaussian kernel is convolved with the image for differentiation. Here, we use the relative scale of the object σ_I as the variance of Gaussian for Harris integration. The variance of Gaussian for Harris differentiation is $\sigma_D = k\sigma_I$, where k is a constant. The scale normalized auto-correlation matrix of Harris detector [Mik04], [Mik01] at a point $X = (x, y)$ of the image I is given by

$$N(X, \sigma_I) = \sigma_D^2 g(\sigma_I) \otimes \begin{bmatrix} I_x^2(X,\sigma_D) & I_x I_y(X,\sigma_D) \\ I_x I_y(X,\sigma_D) & I_y^2(X,\sigma_D) \end{bmatrix}$$

Here, $g(\sigma_I)$ which is the circular Gaussian integration window at the scale σ_I is given by

$$g(\sigma_I) = \frac{1}{2\pi\sigma_I^2} e^{-\frac{x^2+y^2}{2\sigma_I^2}}$$

$I_x(X, \sigma_D)$ and $I_y(X, \sigma_D)$ are the partial derivatives of the given image in x and y direction respectively and can be found by convolving the image with the 1D Gaussian kernel.

$$I_x(X, \sigma_D) = h(\sigma_D) \otimes I(X)$$

$$I_y(X, \sigma_D) = (h(\sigma_D))^T \otimes I(X)$$

where $h(\sigma_D)$ is the 1D Gaussian first derivative kernel at the scale of σ_I *defined as*

$$h(\sigma_D) = -\frac{x}{\sigma_D^3\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma_D^2}}$$

The measure of **corner response** at the point X and scale σ_I is

$$R(X, \sigma_I) = \det(N(X, \sigma_I)) - \lambda \mathrm{tr}^2(N(X, \sigma_I))$$

where λ is a constant. $R(X, \sigma_I)$ is positive in corner region and a point is selected as a **corner point** if it is the local maximum of the measure of corner responses [Har88], i.e. point X is a corner point if

$$R(X, \sigma_I) > 0 \text{ and } R(X, \sigma_I) > R(X_w, \sigma_I) \; \forall \; X_w \in W$$

where W is the 8-neighborhood of the point X.

A corner point could be selected as an **interest point** if

$$R(X, \sigma_I) > T_c$$

where T_c is a constant threshold.

Experimental results for relative scale interest point detection are shown in Section 3.

2.2 Threshold selection for intensity invariant detector

In this section, we investigate the effect of intensity change on the detection method. The change in intensity may happen for different reasons such as different brightness of light source, direction of incident, and change in camera aperture etc. In this paper we only considered the effect of uniform intensity changes, i.e. intensity scaling and intensity shift given by

$$I'(x, y) = sI(x, y) + c$$

where s and c are the intensity scaling and shift parameters respectively.

The measure of corner response $R'(X, \sigma_I)$, at the point X and scale σ_I for the image I', is independent of intensity shift but the intensity scaling affects the response. It can be easily shown that the change in corner response is not linear for the intensity scaling i.e.

$$R'(X, \sigma_I) = s^2 \det(N(X, \sigma_I)) - s^4 \lambda \mathrm{tr}^2(N(X, \sigma_I))$$

Here, we propose a heuristic based method to select the threshold for intensity invariant interest points. The threshold T_v' is selected based on the **normalized-corner response**. Consider, $\Lambda = \{X_i\}$ is the set of corner points for $I'(x, y)$ and $R_i'(X_i, \sigma_I)$ is the corner response at point X_i. The points in the set Λ are sorted in descending order i.e. $R_i'(X_i, \sigma_I) >= R_{i+1}'(X_i, \sigma_I)$ for $i = 1 \ldots. |\Lambda| - 1$; where $|\Lambda|$ is the number of corner points. We get the normalized-corner response $Q(X_i, \sigma_I)$, normalizing $R_i'(X_i, \sigma_I)$ by $max(R_i'(X_i, \sigma_I)) = R_1'(X_i, \sigma_I)$ for $i = 1, \ldots, |\Lambda|$.

$$Q(X_i, \sigma_I) = \frac{R_i'(X_i, \sigma_I)}{R_1'(X_i, \sigma_I)}$$

Plotting the normalized-corner response $Q(X_i, \sigma_I)$ against the detected corner points, we get **Q-Λ** histogram. From experiments, it can be seen that Q-Λ histograms always have almost similar shapes, irrespective of the objects contained in the image. Figure 1(a) and 1(b) show two images having

translation, rotation, scale change and intensity change, with respect to one another. Figure 1(c) shows the Q-Λ histogram for the image 1(b). X-axis of the histogram shows the sequence of detected corner points in the sorted set Λ and the Y-axis shows the corresponding $Q(X_i, \sigma_I)$. Now, we select a threshold T_v' for $Q(X_i, \sigma_I)$, where T_v' is sufficiently large (for example 0.01 - 0.04) to discard most points on the portion of the curve parallel to X-axis (Figure 1(c)). So, a corner point is an interest point if $Q(X_i, \sigma_I) > T_v'$.

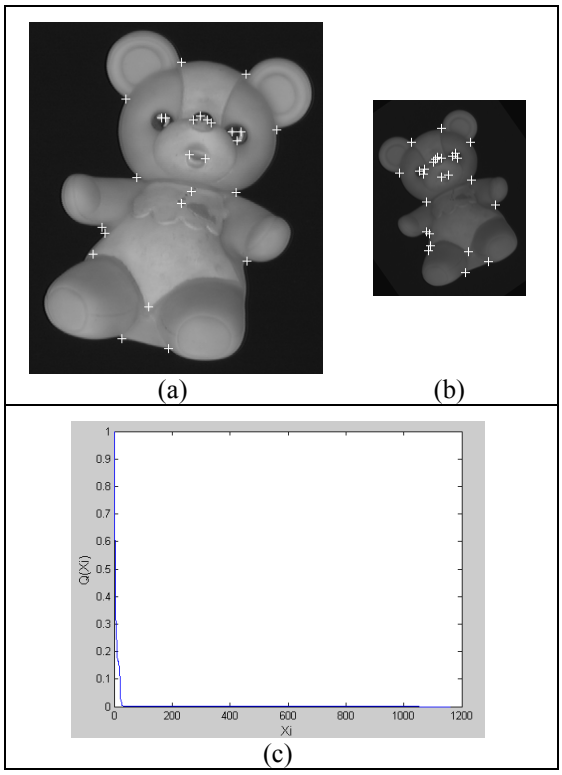

Figure 1: Threshold selection method a) The reference object and detected interest points. This image is borrowed from *RSORT* [Sel99]. b) Modified image (reference object is rotated (35^0), scale changed (50%), intensity scaled (0.5) and intensity shifted (+10)). c) Q-Λ histogram for the image in (b).

2.3 Cluttered environment and partial occlusion

Now, we will consider two other important factors for interest point detection: the given test image may contain other objects and the object of interest may be partially occluded (Figure 2). The threshold selection method is greatly affected by these two factors. In the case of cluttered environment, stronger corner points may come from other objects which affect the normalization process. For Partial occlusion some of the strongest corner points may be absent due to the occlusion, again affecting the normalization process. In both cases, our solution is to reduce the threshold, allowing more interest points. For a large number of detected corner points due to the cluttering, we want to reduce the the threshold slightly to ensure the higher repeatability rate. A particular choice for the final threshold T_v, for invariant interest point detection is

$$T_v = \frac{T_v'}{\ln(|\Lambda|)}$$

where $|\Lambda|$ is the number of detected corner points in the set Λ.

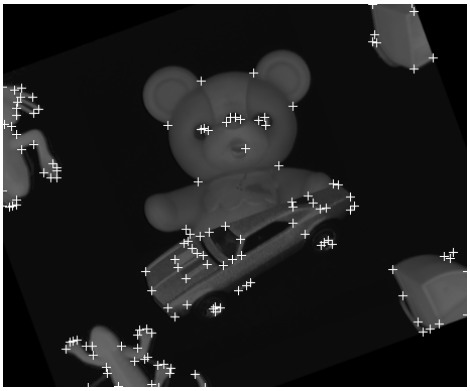

Figure 2: Detected interest points for partial occlusion and cluttered environment. Original images are taken from *RSORT* [Sel99] and then manipulated.

3. EXPERIMENTAL RESULTS

In order to evaluate our interest point detector we have used the repeatability rate criterion proposed by Schmid et al [Sch00]. Here, we can define the repeatability rate as the ratio of number of corresponding interest points pairs and the total number of points detected on the given reference object, with consideration of a localization error of ε. Let $\Lambda^{(r)} = \{X_i^{(r)}\}$ and $\Lambda^{(t)} = \{X_j^{(t)}\}$ be the sets of points detected on the reference object and the same object in the test image, respectively. If $C(X_i^{(r)}, X_j^{(t)})$ is the number of corresponding points pairs then the repeatability rate is

$$\boldsymbol{rr} = \frac{C(X_i^{(r)}, X_j^{(t)})}{|\Lambda^{(r)}|} \times 100\%$$

where $|\Lambda^{(r)}|$ be the number of points in the set $\Lambda^{(r)}$.

We carried out experiments on several objects from two data sets: *RSORT* [Sel99] and *COIL-20* [Mur95]. The four objects used for the experiments are shown in Figure 3. The results are summarized in Table 1. One of the main objectives of this work is to detect scale invariant interest points. We changed the scale of each of the objects from 150% to 50% with a simultaneous change in rotation (35^0), intensity scaling by a factor of 0.6, intensity shift +10. We

computed average repeatability rate (**avg-*rr*)** for all four objects. Figure 4 shows the avg-*rr* for scale change. On the average, we got 86% avg-*rr* which is much better than the improved Harris detector [Sch00] and better than the Harris-Laplacian method [Mik01], [Mik04]. Our relative scale method is as fast as the single step Harris method which is much faster than the scale-space method. We computed avg-*rr* for change of rotation angle (0^0 -180^0) and change of intensity (intensity scaling factor was changed from 1.6 to 0.4). We also carried out experiments for partial occlusion and cluttered environment as shown in Figure 2. Here, we have two objects of interest (bear and car) present in the image. Both of them have further scale, rotation and intensity change with respect to the given reference objects. The repeatability rate for the bear is 85% and for car is 70%. *Avg -rr* for these two objects is 77.5%.

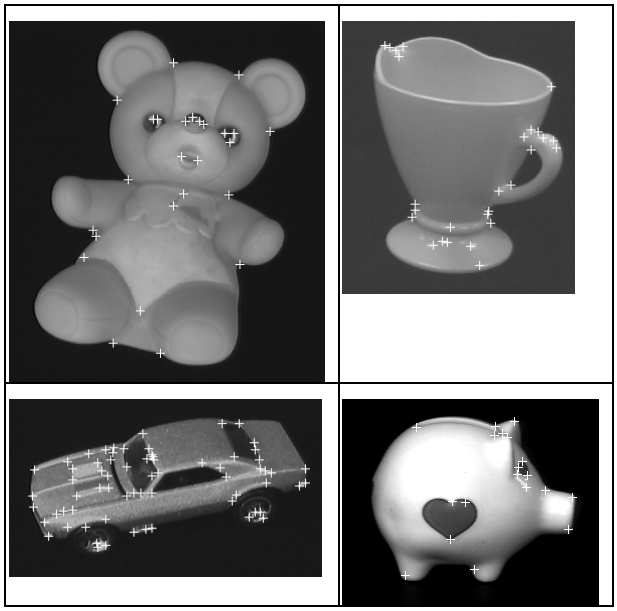

Figure 3: Four reference objects used for experiment. Detected interest points are also shown. Images are taken from *RSORT* [Sel99] and *COIL-20* [Mur95] dataset.

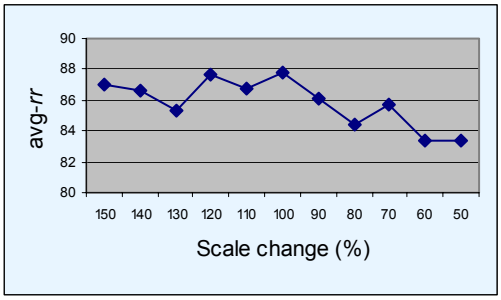

Figure 4: Average repeatability rate (avg-*rr*) for scale change (150% - 50%) for the four objects in Figure 3.

Transform-ation /cond-ition type	**Range of change**	**Additional simul-taneous transfor-mation/condition**	**Average and range of avg-*rr* (%)**
Rotation	0^0 - 180^0	Scale 80% Intensity scaling 1.2 Intensity shift -5	86.86 [83.3 - 91.6]
Scale	150%-50%	Rotation 35^0 Intensity scaling 0.6 Intensity shift +10	85.84 [83.3 - 87.8]
Intensity scaling	1.6 – 0.4	Scale 120% Rotation 25^0	83.5 [83.5 - 83.5]
Cluttered & Partial Occlusion	Figure 2	Scale 80% Rotation 20^0 Intensity scaling 0.5	77.5 bear- 85% car – 70%

Table 1: Results for different transformations and conditions.

4. DISCUSSIONS

In this paper, we introduced the relative scale interest point detection method for scale invariant interest points. A heuristic based method was proposed to select thresholds for invariance to intensity change, cluttered environment and partial occlusion. Our detection method gives better performance than the existing methods in term of repeatability rate and computing time. The method could be useful for many applications. For the relative scale method we assumed that the scale of the object is known in advance. Our method only deals with linear geometric transformations. Consideration of affine scale change is important for real-time applications. As future work we like to find invariant descriptors for the detected interest points. Such descriptors could be helpful for point-to-point matching.

5. REFERENCES

[Har88] Harris, C. and Stephens, M. A combined corner and edge detector. 4th Alvey Vision Conference, Manchester, UK, 1988.

[Mik04] Mikolajczyk, K. and Schmid, C. Scale & affine invariant interest point detectors, International Journal of Computer Vision, vol. 60, no. 1, pp. 63-86, 2004.

[Mik01] Mikolajczyk, K. and Schmid, C. Indexing based on scale invariant interest points. ICCV01, pp. 525-531, 2001.

[Mur95] Murase, H. and Nayar, S. Visual learning and recognition of 3D objects from appearance. International Journal of Computer Vision, vol. 14, no. 1, pp. 5-24, Jan 1995.

[Sch00] Schmid, C., Mohr, R. and Bauckhage, C. Evaluation of interest point detectors. International Journal of Computer Vision, vol. 37, no. 2, pp. 151-172, 2000.

[Sel99] Selinger, A. and Nelson, R.C. A perceptual grouping hierarchy for appearance-based 3D object recognition. Computer Vision and Image Understanding, vol. 76, no. 1, pp. 83-92, 1999.