CIS 2 Paper Review

Author:
Jiayin Qu

Paper:
A Combined Corner and Edge Detector
Harris, C. G. and M. Stephens. "A Combined Corner and Edge Detector." *Alvey Vision Conference* (1988).

Paper Selection Reason:
This paper provides us with a novel method for detecting edges and corners. Unlike continuous texture, edges and corners are essential features in images. Recognizing these features and further matching can be applied in object recognition and stereo matching. In this paper, the Harris detector is a powerful tool for finding edges and corners based on intensity of the image. We will further show that this method can achieve relatively high accuracy and the computation load is not high.

Project Summary
The project originates from the idea that many of the operations on medical equipment can be performed by robots. When the environment in ICU is contagious, remotely control a robot to interact with a medical equipment can reduce time, protection gear and exposure risk.

An ideal robot would have 6 DOF, and the general flow of the robot would be:

1. Recognize the equipment that the user designated to interact with;
2. Change the relative pose based on live camera input;
3. After settled, the robot would be able to operate with various modalities including knobs, buttons, etc. by interacting with a user interface;
4. The robot would perform various functions and the user can get live camera feedback;

However, note that due to various limitations, our goal for the course would be building a 2D cartesian robot that can interact with an oscilloscope with working user interface. Other models will also be developed besides this including object recognition, stereo matching, hand-eye calibration, etc.

Background:
To perform stereo matching, we need to find feature points from the image and from the object. Continuous texture cannot serve as meaningful feature points. Before Harris's detector, early methods of tracking image features require the features to be discrete. Some methods (methods by Ayache, N and F lustman) represent edges as a set of straight-line fragments so that edges can still be discretely represented. An example would be an image of a bush and building. Figure 1 shows the original images. Figure 2 shows what the features look like when discretized.

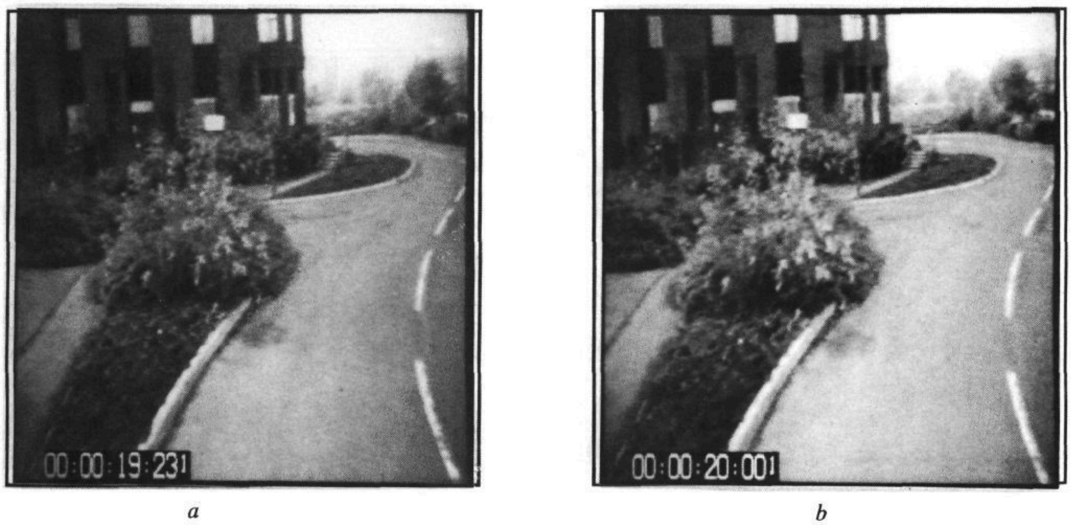

Figure 1

After a

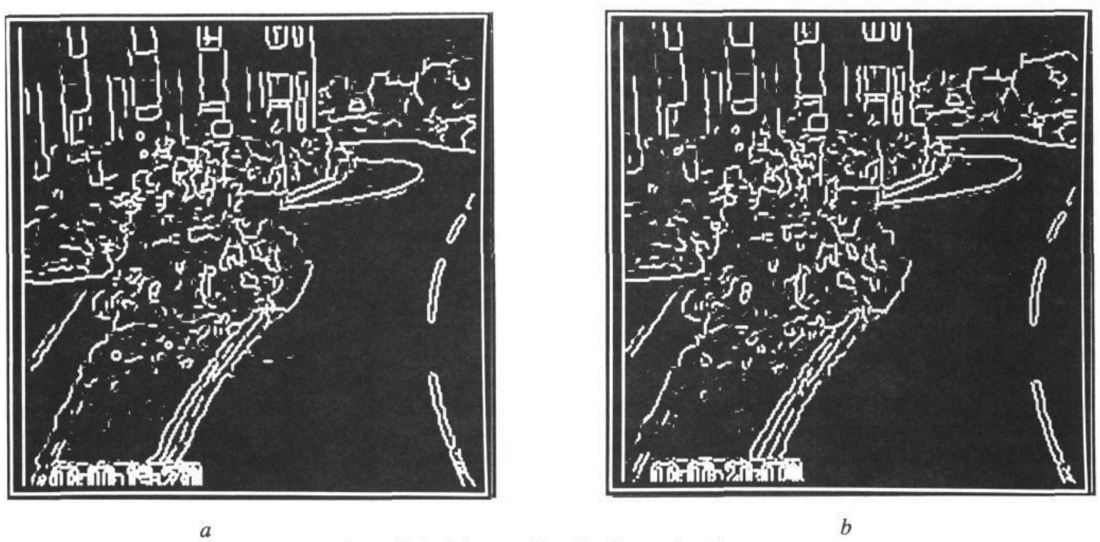

Figure 2

However, representing features in a discrete way would loss the original information regarding connectivity. Therefore, they behave badly in describing surfaces and objects. Later methods were developed to repair the feature image by doing junction completion. An example is as shown in Figure 3. Note that for objects like bush, it is not quite reasonable to describe them in line fragments. In fact, retaining the discretization is a better choice for such objects.

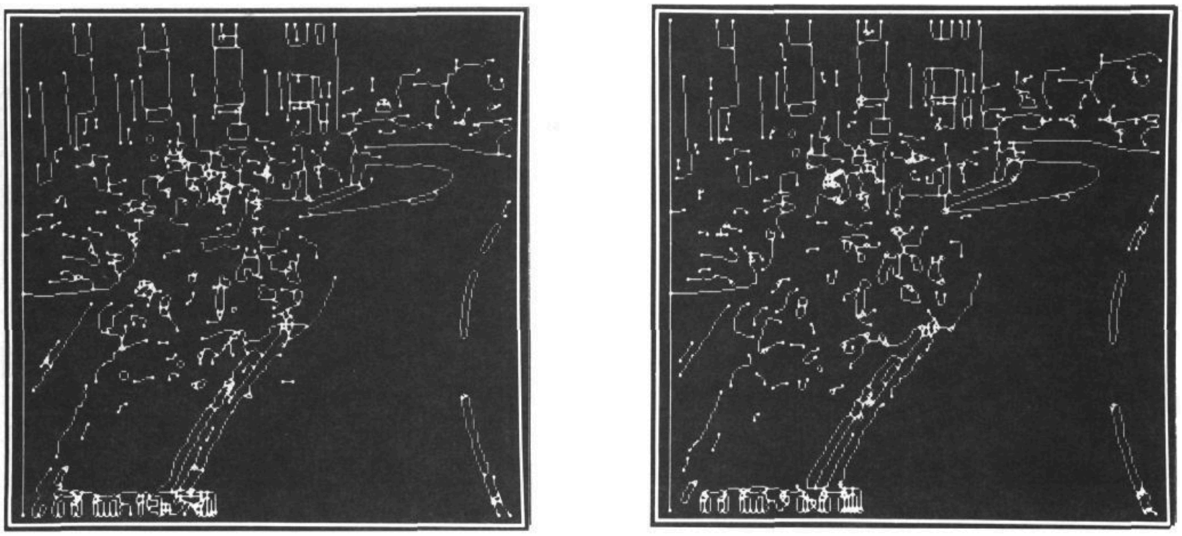

Figure 3

Algorithm in detail:
The basic idea of the Harrison edge and corner detector is to look through a small window (usually 3 by 3). While moving the window, we record the change of intensity within the window. If the intensity has no change, we know it is in a continuous texture area. If the intensity changes in one direction but not the other, we know that we have encountered an edge. If the intensity always changes no matter how the window moves, we know there is a corner in the window.

To calculate the change of intensity, we find the local gradient by using partial derivative with respect to x and y.

Gx and Gy can be calculated as follows:

$$G_\chi = \left(\frac{\partial I(u,v)}{\partial x}\right) \qquad \mathrm{G_y} = \left(\frac{\partial \mathrm{I(u,v)}}{\partial \mathrm{y}}\right)$$

Note that when calculating the partial derivative, we can calculate it by convolving the original image with a derivative kernel. Derivative kernels in x and y direction are as follows:

$$\partial_{\mathrm{x}} = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \qquad \partial_{\mathrm{y}} = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

The local structure matrix with respect to window M is computed as:

$$M = \begin{bmatrix} \sum_{u,v}^{W} G_x{}^{2} & \sum_{u,v}^{W} G_X \times G_{\mathrm{y}} \\ \sum_{u,v}^{W} G_X \times G_{\mathrm{y}} & \sum_{u,v}^{W} G_{\mathrm{y}} 2 \end{bmatrix}$$

We can then obtain the R-value, or Harris score, by taking the difference between the determinant of M and k times the squared trace of M. Note that k is a constant and by changing it, we have the freedom of retaining or discarding certain features that are below threshold.

One example of applying Harris detector would be setero matching. The general steps are as follows:

1. Input of the image I
2. Gradient calculator
3. Partial derivative unit
4. Finding corners
5. Non-maxima suppression
6. High and low threshold operator
7. Display corners
8. Feature matching

Evaluation of the algorithm:
To evaluate the performance of the algorithm, we test the repeatability of the detector. Repeatability is defined if the detection is independent of changes in the imaging conditions including parameters of the camera, camera position relative to the scene, and illumination conditions, etc. The independency is achieved if, as shown in figure 4, the detected point xi is within the epsilon neighborhood.

The following result is from *Evaluation of Interest Point Detectors* by C. Schimid. Several evaluations are made based on various factors.

The first is if the detector is invariant to rotation. An example is as shown in figure 4. As shown in figure 5, the Harris detector achieves the highest repeatability score among other detectors. Note that the right side of figure 5 has larger epsilon compared with the first one.

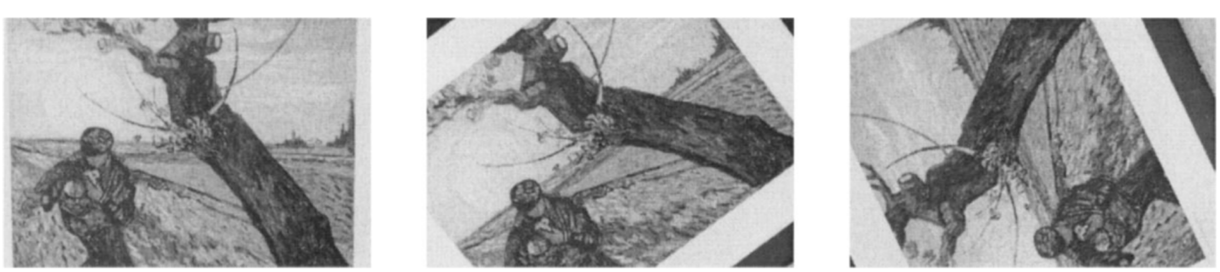

Figure 4

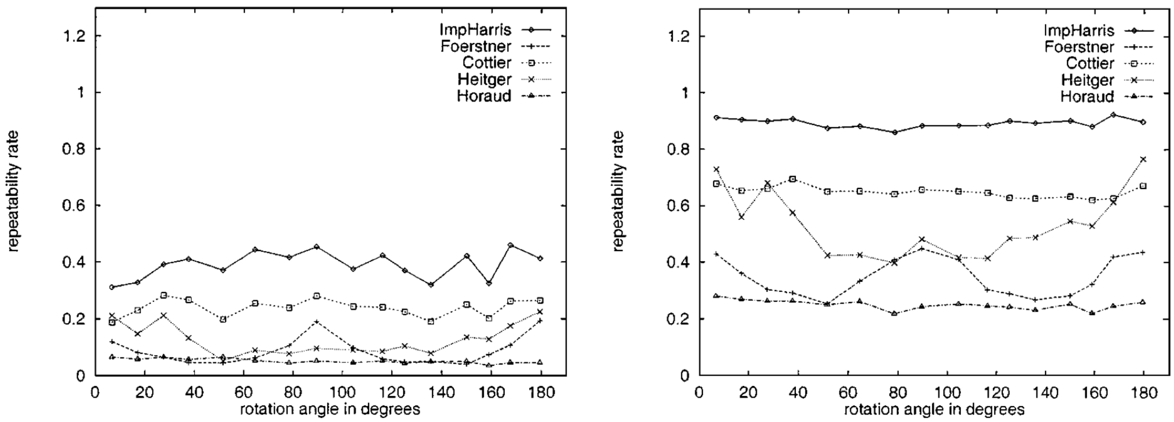

Figure 5

The second evaluation is based on testing the invariance to scaling of the detector. As shown in the figure 6, the performance of Harris detector is not quite good. The underlying reason is rather explicit since if we shrink the image, what is initially detected as a combination of edges and corners might be evaluated as corners alone.

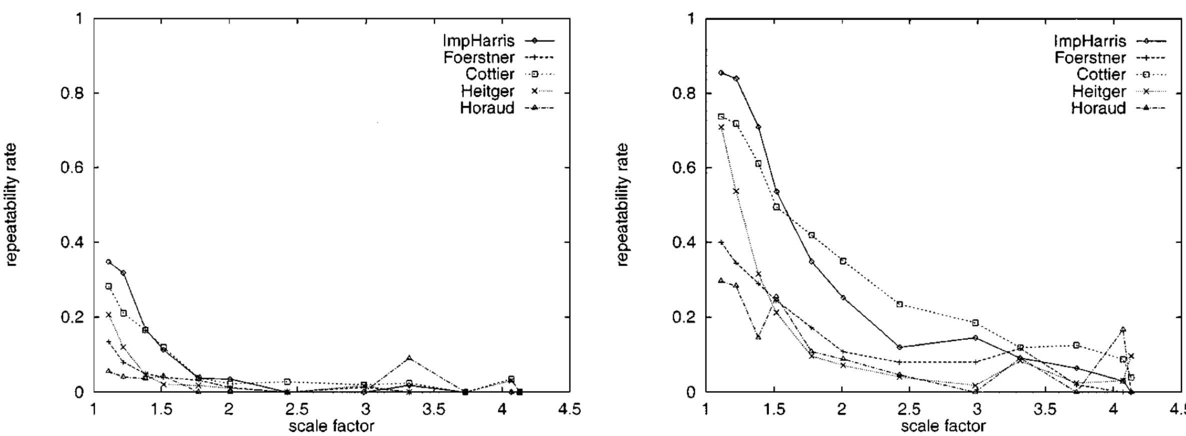

Figure 6

The third evaluation is based on the invariance to intensity change of the detector. As shown in figure 7, Harris detector again achieves the best performance among all. The underlying reason is that Harris detector is obtained based on gradients. In other words, local maxima and minima are invariant to linear shift of intensity.

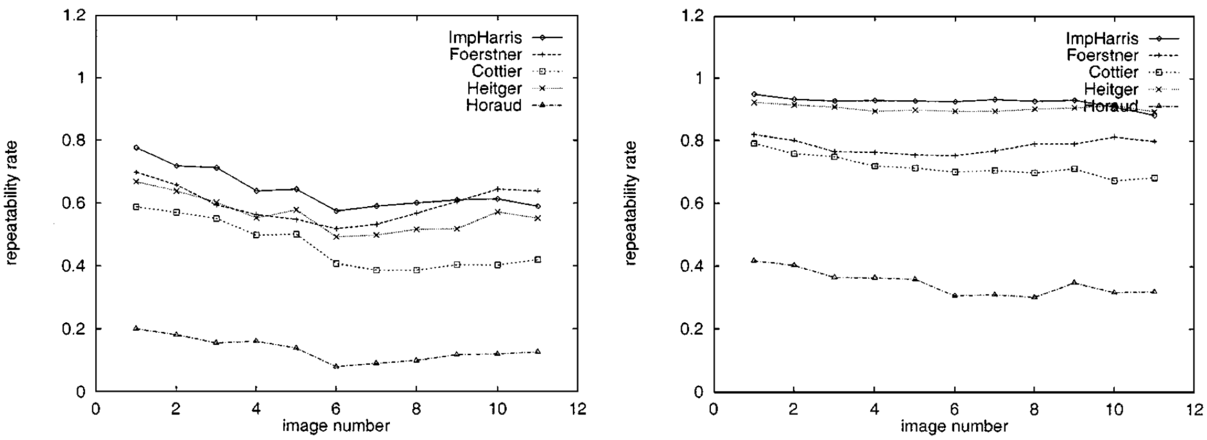

Figure 7

The final evaluation is based on viewpoint change. An example of viewpoint change is as shown in figure 8. And the result is as shown in figure 9.

Figure 8

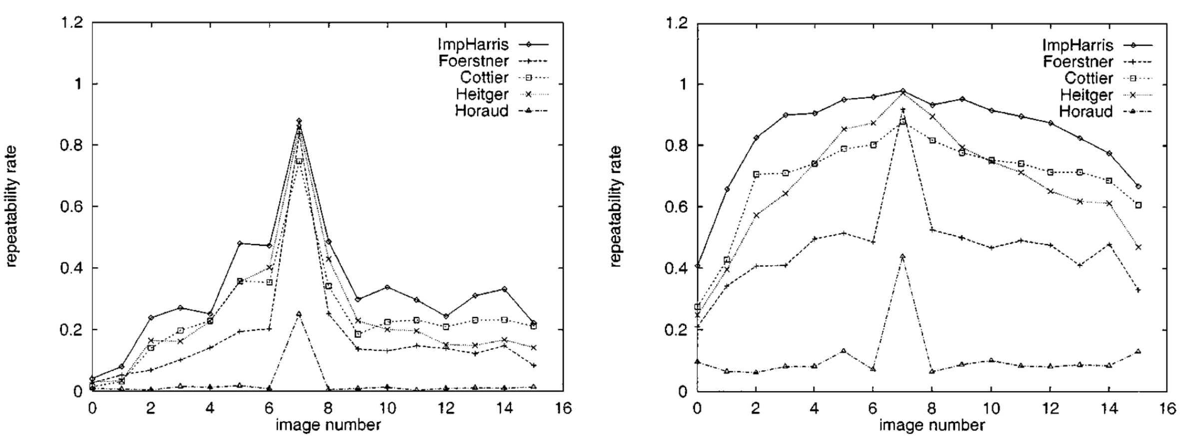

Figure 9

Significance:
The Harris detector allows for flexible threshold operator for exact corner detection. Recall this can be achieved by setting different k values and threshold for R values. The latency as well as the computation load of Harris detector are also proved to be low. As shown in the evaluation parts above, Harris detector is almost invariant to rotation and change of intensity. It also responds well to viewpoint change.

Critiques:
Despite of the merits of Harris detector shown above, several problems remain for the detector. The first one is that the Harris detector responds bad to scaling. It also behaves bad when there is there exists occlusion for the object.

Other developments:
Various developments were made regarding to Harris detector. An improved Harris detector is suggested by changing the way of calculating gradient. Initially it is similar to taking a derivative of a squared box. In ImpHarris, the derivatives of a Gaussian filter are taken instead to make the curve smoother. Applying Gaussian filter also enables fast corner detection due to the recursive implementation algorithm of the Gaussian filter developed by Deriche. Besides Harris detector, other forms of detectors are also developed including descriptor extraction and regional feature detection.

Conclusion:
Harris edge and corner detector performs well in extracting useful feature information from the image. The detector is also proved to be relatively fast and stable. Therefore, it is a powerful tool for feature finding and would further enable other operations like stereo matching and object recognition.