Faculty of Engineering
Cairo University

Computer Vision

Task 3

Features Detection and Image Matching

Submitted by:

BN	Section	Name
10	1	Esraa Mohamed Saeed
12	1	Alaa Tarek Samir
15	1	Amira Gamal Mohamed
8	2	Fatma Hussain Wageeh
26	2	Mariam Mohamed Osama

Team 3

Submitted to:

Dr. Ahmed Badwy

Eng. Laila Abbas

Contact Email: amira.gamal.ag22@gmail.com

Harris

Harris Corner Detector is a mathematical approach to detect corners and edges in images. The idea is to locate interest points where the surrounding neighborhood shows edges in more than one direction.

The basic idea of algorithm is to find the difference in intensity for a displacement of (u,v) in all directions which is expressed as below:

$$E(u,v) = \sum_{x,y} \underbrace{w(x,y)}_{\text{window function}} \, [\underbrace{I(x+u,y+v)}_{\text{shifted intensity}} - \underbrace{I(x,y)}_{\text{intensity}}]^2$$

The above equation can be further approximated using Tayler expansion which gives us the final formula as:

$$E(u,v) \approx \begin{bmatrix} u & v \end{bmatrix} M \begin{bmatrix} u \\ v \end{bmatrix}$$

Where,

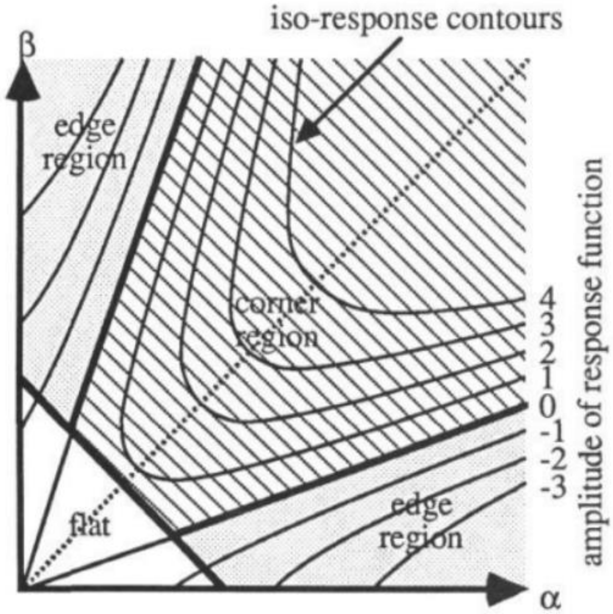
$$M = \sum_{x,y} w(x,y) \begin{bmatrix} I_x I_x & I_x I_y \\ I_x I_y & I_y I_y \end{bmatrix}$$

Ix and Iy are image derivatives in x and y directions respectively.

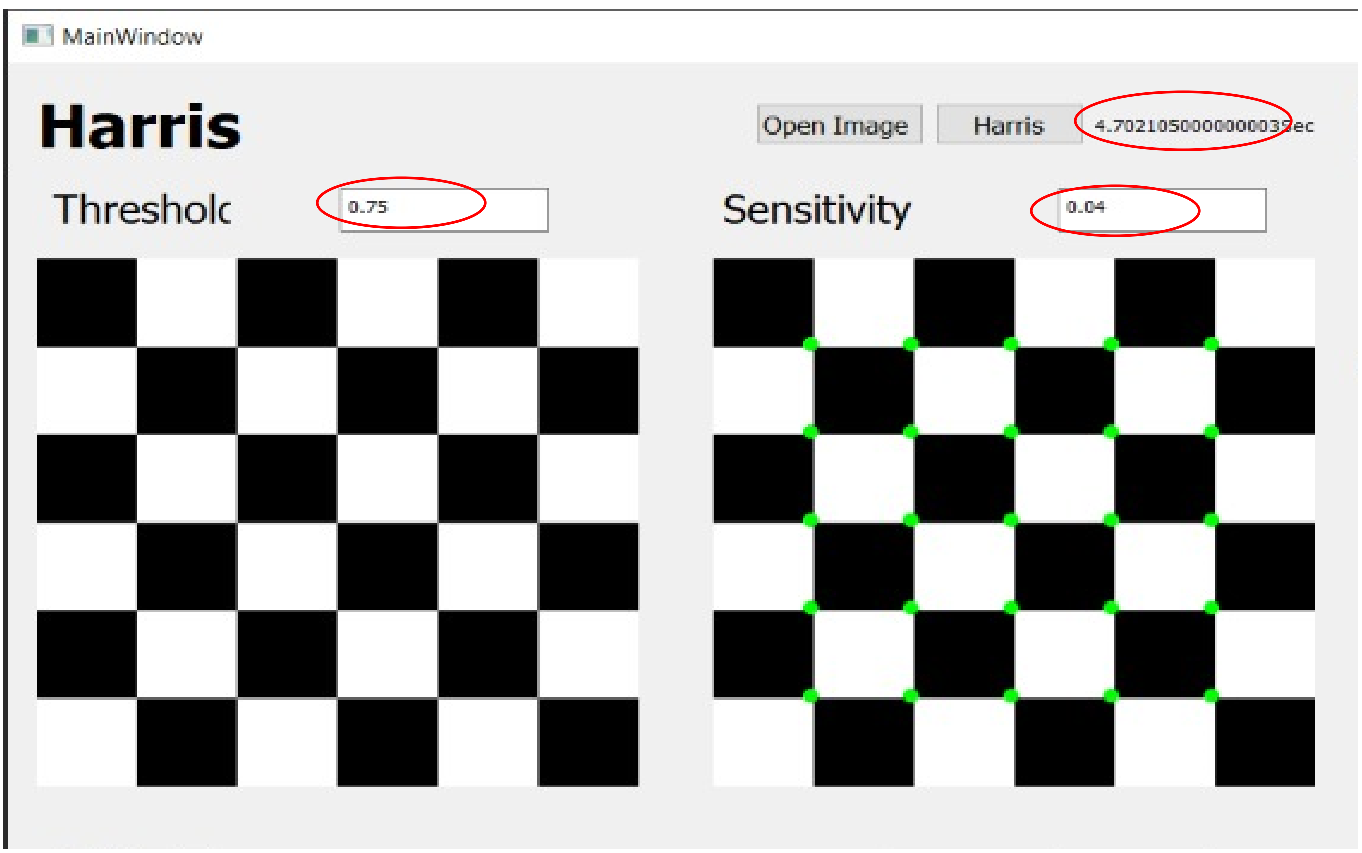
Then we finally find the Harris response R given by:

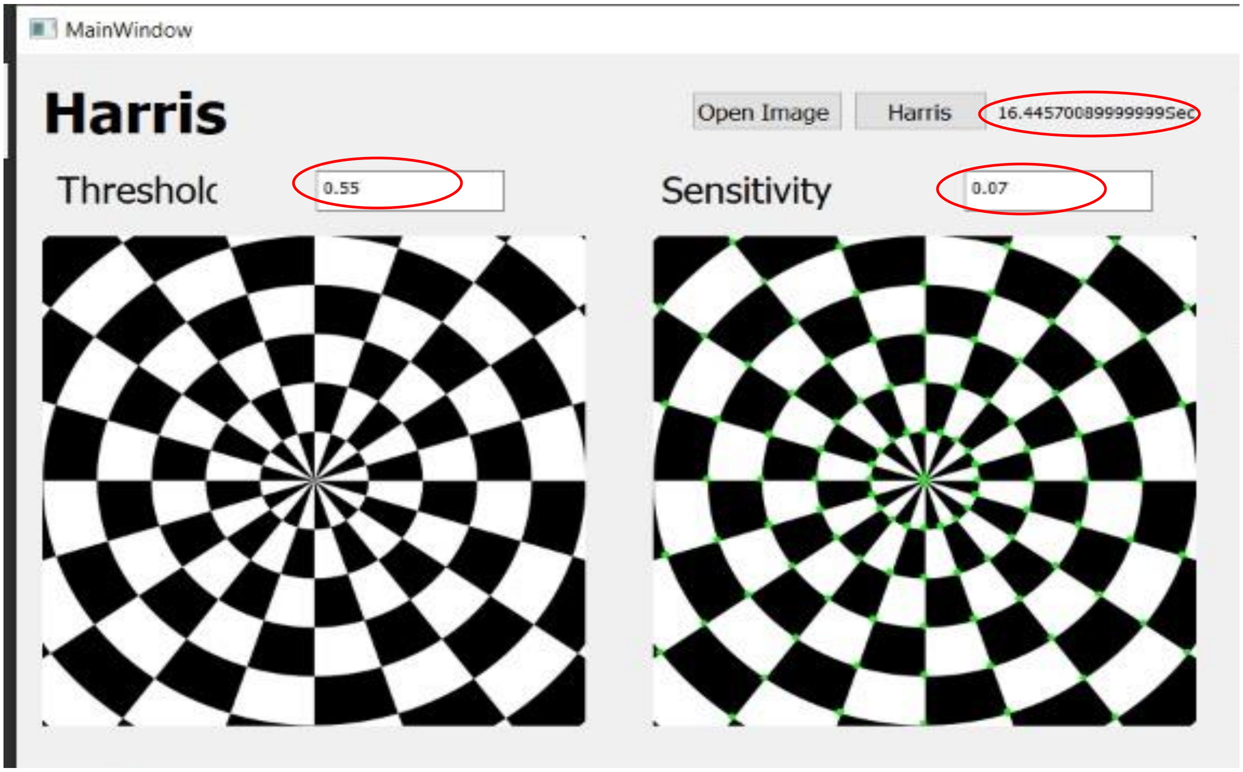
$$R = det(M) - k(trace(M))^2$$

We find the corners using the value of R.



Results





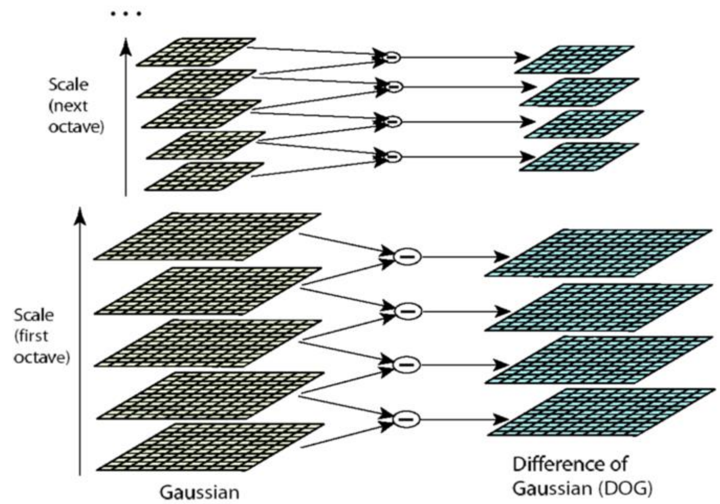
SIFT

SIFT key points are useful due to their distinctiveness. This distinctiveness is achieved by assembling a high-dimensional vector representing the image gradients within a local region of the image. The key points have been shown to be invariant to image rotation and scale and robust across a substantial range of affine distortion, the addition of noise, and change in illumination.

The fact that key points are detected over a complete range of scales means that small local features are available for matching small and highly occluded objects, while large key points perform well for images subject to noise and blur.

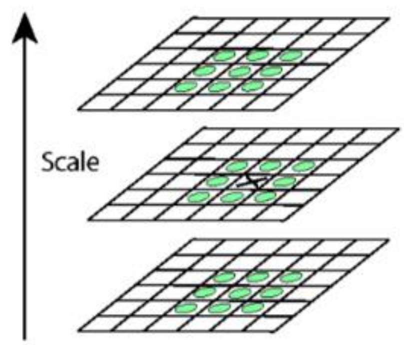
Scale Space Construction:

For each octave of scale space, the initial image is repeatedly convolved with Gaussians to produce the set of scale-space images shown on the left. Adjacent Gaussian images are subtracted to produce the difference-of-Gaussian images on the right. After each octave, the Gaussian image is down-sampled by a factor of 2, and the process is repeated.



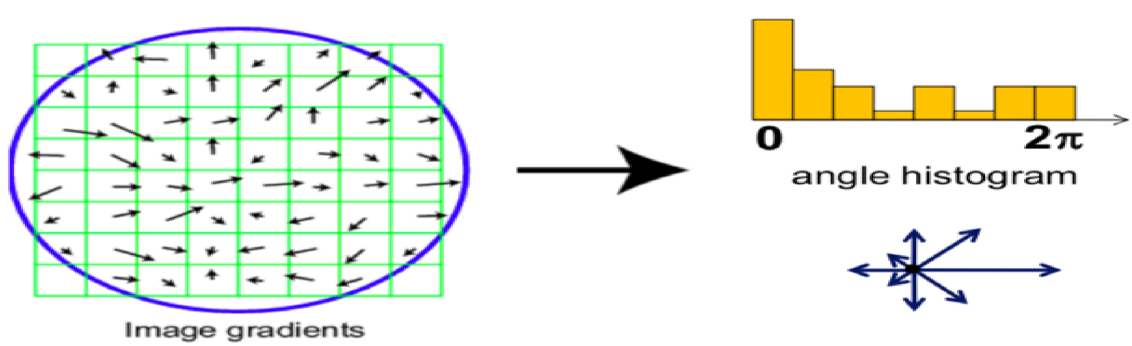
Scale Space Extrema Detection:

Maxima and minima of the difference-of-Gaussian images are detected by comparing a pixel (marked with X) to its 26 neighbors in 3x3 regions at the current and adjacent scales (marked with circles).

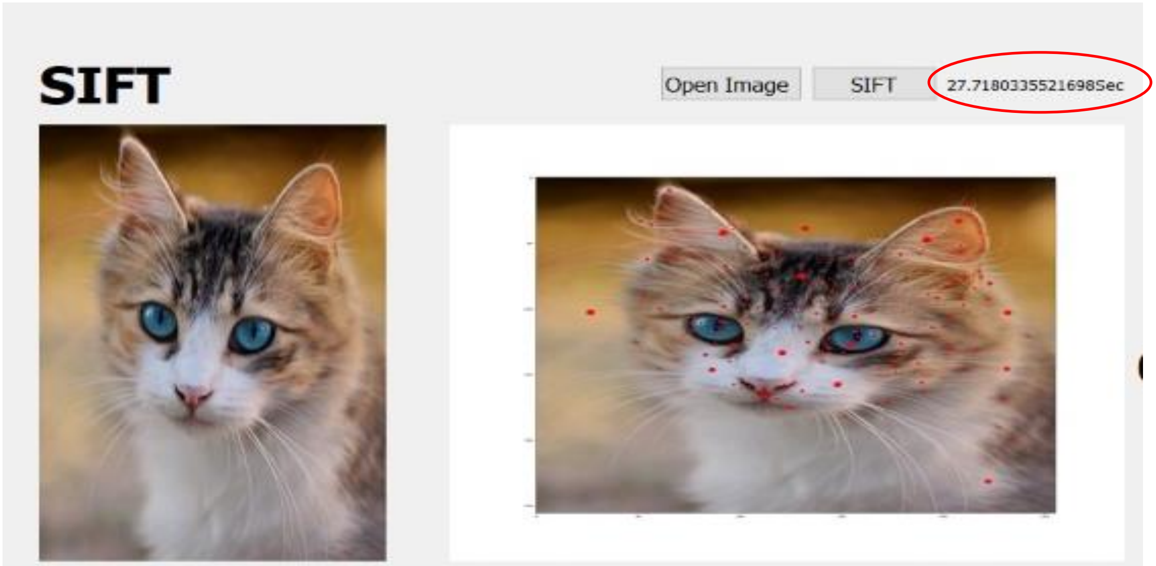


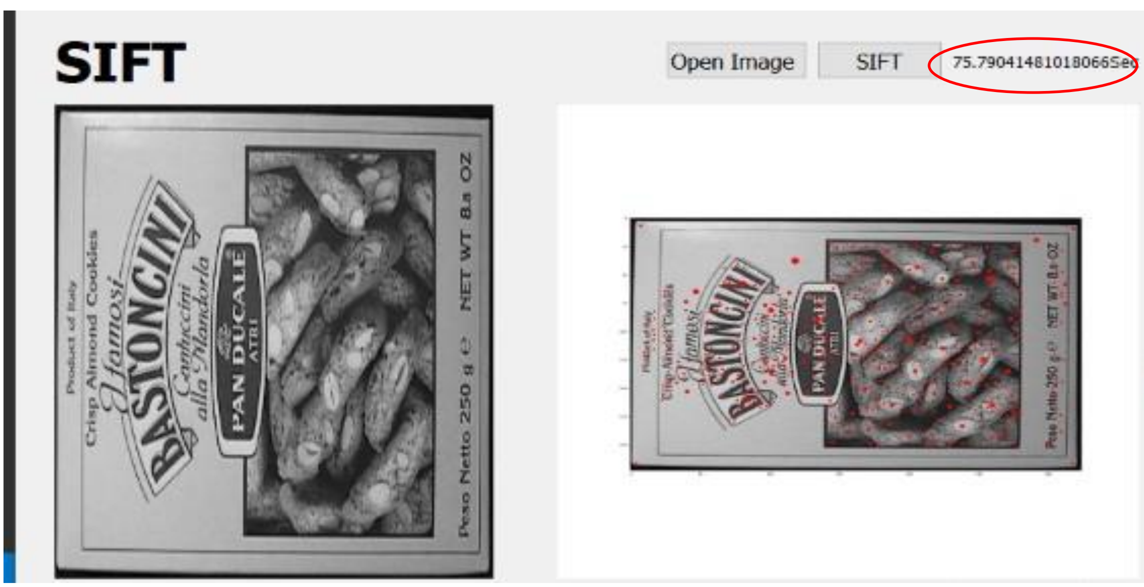
Key point descriptor:

- Take 16x16 square window around detected feature
- Compute edge orientation (angle of the gradient - 90°) for each pixel
- Throw out weak edges (threshold gradient magnitude)
- Create histogram of surviving edge orientations



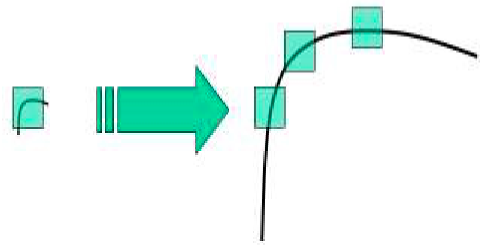
Results


SIFT
Open Image
SIFT
27.71803355216985ec

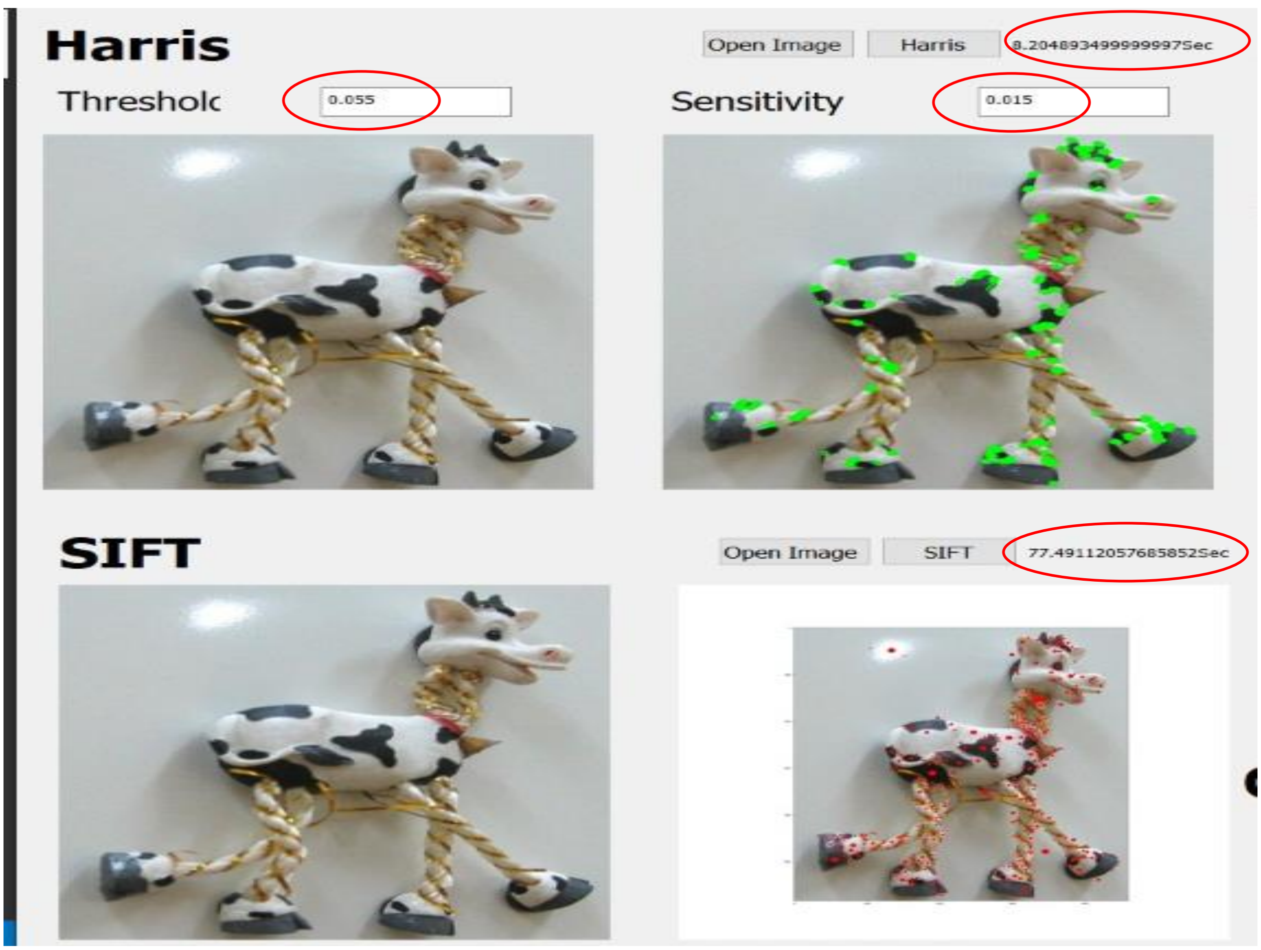

SIFT
Open Image
SIFT
75.790414810180665ec

The Difference between Harris and SIFT

SIFT has the property of scale invariance, which makes it better than Harris. Harris is not scale-invariant, a corner may become an edge if the scale changes, as shown in the following image.



SIFT Vs. Harris



Feature Matching

We used two methods to match the features extracted from SIFT one on them is Sum of Squared Difference (SSD) which calculates the summation of squared for the product of pixels’ subtraction between two images, and its equation is:

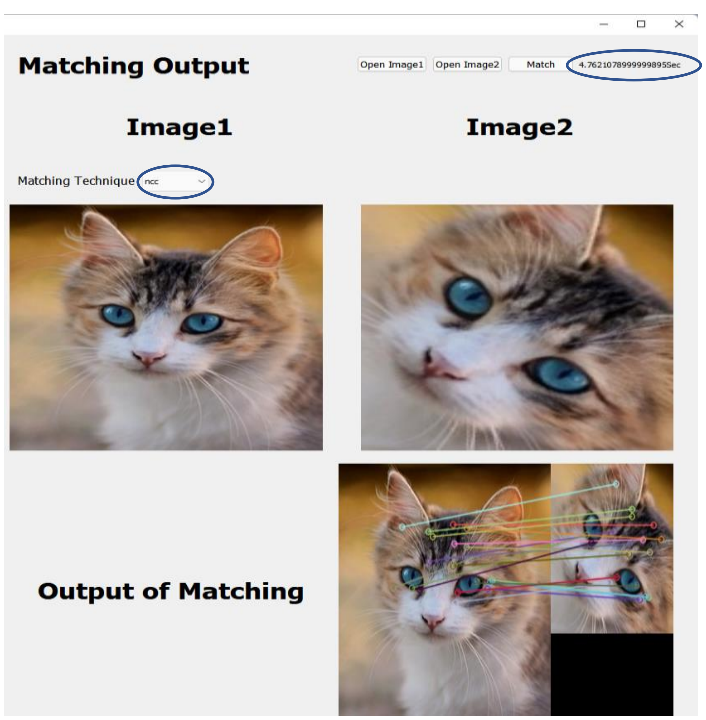
$$\sum_i (I_1(\mathbf{x}_i) - I_0(\mathbf{x}_i))^2$$

And the other method is Normalized Cross Correlation (NCC) which has more complex computation compared to (SSD) as it involves numerous multiplication, division and square root operations, and its equation is:

$$\frac{1}{n-1} \sum_i \frac{(f(x_i) - \bar{f})(g(x_i) - \bar{g})}{\sigma_f \sigma_g}$$

After applying these two methods we noticed that matching computation time for NCC is less than that for SSD

NCC



SSD