Square Detection

Introduction

The general idea for the coursework outlined an image processing pipeline which involved multiple stages. The first of which involved generating an edge map of an image. From here the edge image is mapped in order to determine the most likely lines and their polar coordinates (p theta parameterization of hough transform). From here the aim of the pro forma was to provide a means of processing the processed image in order to be able to vote for the likelihood of a square occurring at a set pixel. Using more than one aspect, e.g. orientation position etc.

Difference of gaussian

The first step in the process applying a difference of gaussian edge detection. This is an estimation of the laplacian of gaussian(See fig.1) edge detection, however it is moreso computationally feasible. Initially the image is blurred, blurred again with a different sigma, and the two resultant arrays are subtracted from one another to produce the desired effect. The manner in which this is done is governed by the one(See fig.2) and two dimensional gaussian distribution functions; The process of blurring using the gaussian function is done discreetly using kernels, empty matrices populated with the result of the gaussian function. Generally speaking this can be done using a 2 dimensional matrix however for computational efficiency it is better to incorporate two separate 1D kernels for both the X and Y axis respectively. Either kernel is convolved with sub-matrices of the image in size 2*3*sigma, and the multiplicative sum of both kernel and sub-matrix is set to the central pixel. The resultant image is one that has a spectrogram with far fewer peaks above the threshold set by the deviation, this is often done in order to remove noise in an image.

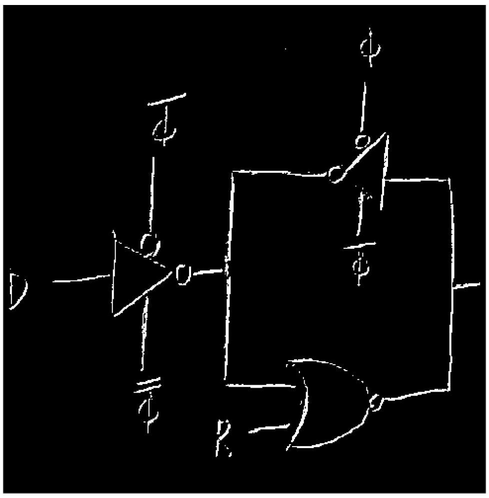
$$LoG(x,y) = -\frac{1}{\pi\sigma^4}\left[1 - \frac{x^2+y^2}{2\sigma^2}\right]e^{-\frac{x^2+y^2}{2\sigma^2}}$$

<F.1 F.2>

$$G(x,y) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{x^2+y^2}{2\sigma^2}}$$

The edge detection that is provided from the difference of gaussian operation can then simply be calculated by index-wise subtraction of the matrices resulting from the aforementioned method. As can be seen in the image (See fig.3) the image has been edge detected using the difference of gaussian edge detection method.

The effect of the difference of gaussian edge detection is similar to that of a band-pass filter in signal processing, as an image is also a signal of a spatial field. Therefore it can be said the difference of gaussian edge detection is a spatial band pass filter of distance sigma1-sigma2.



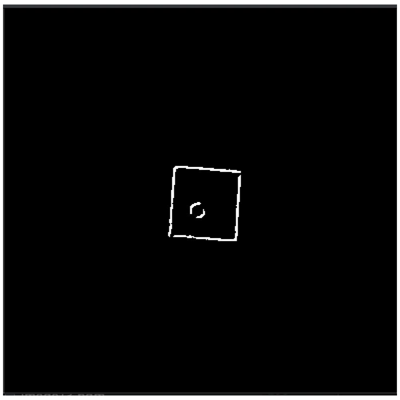
F.3

Hough Transform, Moment calculation and Local detection of maxima and plotting of lines.

The hough transform is an image processing technique used in order to extract edge features out of images. The transform can be thought of as a mapping function from a point in the pixel array of the image, to a curve or a line in the hough space. The parameterization of the hough transform can take most mathematical representations of lines, such as y = mx+c etc; this method however has its fallbacks, such that when the line is vertical, the y intercept values can be infinite. To counter this, often the (p,θ) parameterization is used, p representing the lines distance from the origin, and theta representing the angle from the origin.

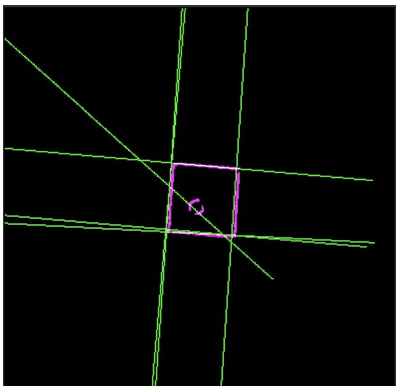
In this application, the hough transform variant is used as you can calculate the orientation of local windows using geometric moments. Using the orientation from a given coordinate on the image, you can reduce the degrees of freedom in the image as peaks will have higher accumulator votes where they are detected given the angle. The figure on the right(See fig.4) is the output from the hough transform after incrementing the most likely candidate for lines to occur in the image space. The hough space pictured on the right can be used to illustrate the highest peaks in the space. Plotting these peaks given a threshold inputted at the beginning of the program via a factor double which indicates as to what factor of the highest peak shall be the threshold of the local maxima search.

In order to establish the lines in the accumulator space, the conversion from polar to cartesian coordinates must occur. This ensures the images' strong lines do not turn out deformed and misshapen due to the wrong width and height parameters being used in the drawing of the lines. The way in which this is done given the 19x19 local search size, ensures that the lines being found are definite, and strong edges in the image, as the peaks are comparatively checked against every index in the 2 dimensional sub-matrix given by the constant neighborhood size. The values of rho and theta where these peaks occur are saved. These values are then used in order to draw lines across the image(See fig.6).



F.4

F.5



F.6

Name: Korab Berisha
StudentID: c1825847