TNM089
Evaluation of spatial and fourier focus measures for autofocusing

Johnny Elmér
Felix Lindgren

October 17, 2022

Abstract

In recent years camera companies have developed new and improved technologies to automatise a cameras focus point allowing the camera to calculate whether a point in an image is in focus with the help of a focus measure. Using this focus measure the camera is able estimate whether the current distance between the lens and sensor has produced a more or less focused image than the previous distance.

In this project six spatial focus measures; Sum-modified-Laplacian(SML), Energy of Laplacian of the image(EOL), Energy of image gradient(EOG), Spatial frequency(SF), Tenengrad and Variance, as well as two Fourier transform-based focus measures; Spectrum Band Ratio (SBR) and Spectrum were implemented. Which were subjectively evaluated from six criteria; Unimodality, Monotonicity, Defocus and noise sensitivity, Effective range, Computational efficiency and Variability.

The results showed that of the spatial focus measures SML had the best performance with EOL in close second. EOG, SF and Tenengrad were similar in performance, with some outlier cases. Variance showed itself to be the worst. For the Fourier transform-based measures Spectrum showed good results, whilst SBR had differing levels of effectiveness depending on the image and point in said image was being evaluated.

Contents

List of Figures

Chapter 1

Introduction

In recent years camera companies have developed new and improved technologies to automatise a cameras focus point. A camera is a sealed, light-tight box that is used to capture a visual image by exposing a digital sensor to light. The light is focused onto the sensor by using a curved lens to direct the light onto the sensor. A cameras focus point is decided by the distance between the lens and the sensor of the camera. Objects that are out of focus appear blurred in the final image [1], see Figure 1.1.

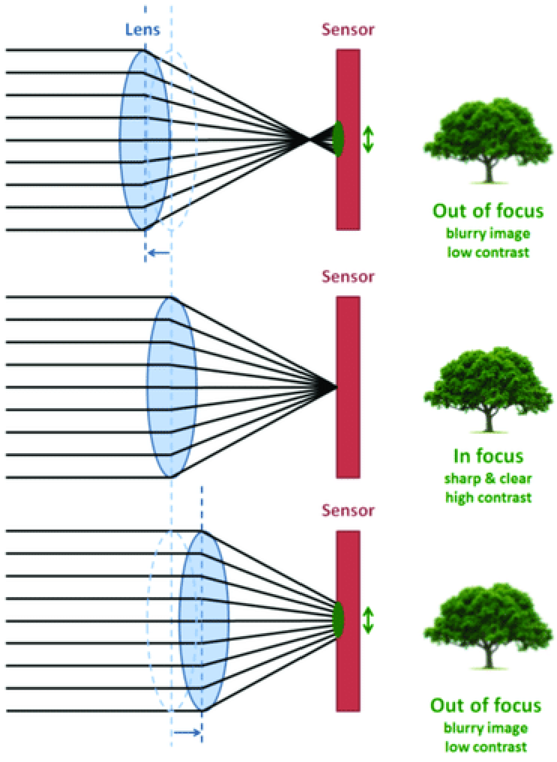


Figure 1.1: Phase detection auto focus

Companies have developed an autofocus technique that uses a motorized lens to move the lens closer or further away from the sensor depending on the distance between the camera and the object in the frame [1].

The camera can calculate whether a point in an image is in focus with the help of a focus measure, which can estimate whether the current distance between the lens and sensor has produced a more or less focused image than the previous distance. Adjustments can then be made back and forth until the best value, and with it the best possible focus, is found. A focus measure is one of the key components in any auto focus system [1].

This project aims to test several different focus measures of different types, mainly spatial and transform based measures, to see which give the best results, for this MATLAB 2021b was used to implement and test the measures.

Chapter 2

Theory

2.1 Focus measures

In order to measure the quality of the focus in a image a focus measure is used. There exists many different focus measures which can be separated into two different categories: focus measures in the spatial domain and transform-based focus measures. According to [2], in the field of autofocusing, focus measures have to satisfy six different criteria:

Unimodality: A focus measure should only have one extreme value that corresponds to the best-focused image.
Monotonicity: On each side of its extreme value the focus measure should be strictly monotone so that the focus measure values vary between different levels of defocus.
Defocus and noise sensitivity: A focus measure should strive to be sensitive to defocus and insensitive to noise. Unfortunately, defocus and noise tends to contain the same high spatial frequencies making it hard to separate the two.
Effective range: The defocus range should maintain a reasonable sensitivity over a broad range.
Computational efficiency: Since most use cases for focus measures requires computation in real-time the complexity of the focus measure cannot be too high.
Variability: A focus measure should not vary between different cases.

The six focus measures in the spatial domain which are focused on in this report are: Variance, Energy of image gradient (EOG), Tenenbaum's algorithm (Tenengrad), Energy of Laplacian of the image (EOL), Sum-modified- Laplacian (SML) and Spatial frequency (SF). According to Wei Huang and Zhongliang Jing [3] the best focus measure when excluding run-time is SML, Huang and Jing also suggests that EOL is quite similar in performance to SML. According to their evaluations the worst of the tested focus measurements seems to be variance.

Two transform based focus measures are also focused on. These focus measures are based on Fast Fourier Transforms which look at the sum of the spectrum without the DC component as well as the ratio between high and low spatial frequency in the image on a global scale. For this report these two focus measures are called Spectrum and Spectrum Band Ratio respectively in accordance to how they are named in [2].

In the following chapters the implementation and evaluation of the different focus measures will be described, along with what focus measures seem to work the best, respectively worst, in this project's use case.

Variance is a rather simple focus measure to implement as it only calculates the mean of whole image (or in this case cropped image), its equation can be written as follows:

$$Variance = \frac{1}{M \times N} \sum_x \sum_y (f(x,y) - \mu)^2 \tag{2.1}$$

where

$$\mu = \frac{1}{M \times N} \sum_x \sum_y f(x,y) \tag{2.2}$$

Energy of image gradient (EOG) operates locally and can be implemented in parallel, it is rather similar to that of the absolute image gradient [2] with the difference that it looks at the squared values of the image gradient rather than only the absolute values. EOG can be written as the following equation:

$$EOG = \sum_x \sum_y (f_x^2 + f_y^2) \tag{2.3}$$

where

$$f_x = f(x+1, y) - f(x,y) \tag{2.4}$$

and

$$f_y = f(x, y+1) - f(x,y) \tag{2.5}$$

Tenenbaum's algorithm (Tenengrad) works differently from the previous two focus measures in that it is based on using the Sobel operator (equation 2.7) to obtain the magnitude of the gradient (as seen in equation 2.6).

$$Tenengrad = \sum_{x=2}^{M-1} \sum_{y=2}^{N-1} [\nabla S(x,y)]^2 for \nabla S(x,y) > T \tag{2.6}$$

$$\nabla S(x,y) = [\nabla S_x(x,y)^2 + \nabla S_y(x,y)^2]^{1/2} \tag{2.7}$$

where T in equation 6 is a so called discrimination threshold value.

Energy of Laplacian of the image (EOL) is mainly used to look at border sharpness with the help of analysis of the high spatial frequencies. It can be written as the following equation:

$$EOL = \sum_x \sum_y (f_{xx} + f_{yy})^2 \tag{2.8}$$

Sum-modified-Laplacian (SML) as described in [4] is from the beginning, as indicated by the name, based on the Laplacian. The Laplacian in itself can behave in a rather unstable way, which can be to some extent rectified by modifying it to use the absolute values for the partial derivatives in the Laplacian, which results in the modified-Laplacian. To furthermore get the Sum-modified-Laplacian one can simply compute the sum of all the values of the modified-Laplacian in an area of the image, which in turn gives the following equation:

$$\nabla^2_{ML} f(x,y) = |2f(x,y) - f(x-step,y) - f(x+step,y)| + |2f(x,y) - f(x,y-step) - f(x,y+step)| \tag{2.9}$$

$$SML = \sum_{i=x-N}^{i=x+N} \sum_{j=y-N}^{j=y+N} \nabla^2_{ML} f(i,j) for \nabla^2_{ML} f(i,j) \geq T \tag{2.10}$$

Spatial frequency (SF) in this report refers to a modified version of EOG, not the actual spatial frequency one might know in another context. The modified version can be written as follows:

$$SF = \sqrt{(RF)^2 + (CF)^2} \tag{2.11}$$

where RF and CF are row and column frequency

$$RF = \sqrt{\frac{1}{M \times N} \sum_{x=1}^{M} \sum_{y=2}^{N} [f(x,y) - f(x,y-1)]^2} \tag{2.12}$$

$$CF = \sqrt{\frac{1}{M \times N} \sum_{x=2}^{M} \sum_{y=1}^{N} [f(x,y) - f(x-1,y)]^2} \tag{2.13}$$

Spectrum (equation 2.14) and Spectrum Band Ratio (equation 2.15) are as described earlier based on Fast Fourier transform and look at the sum of the high and low spatial frequencies in different ways. Spectrum focuses only on the high spatial frequencies while SBR focuses the ratio between the high and low spatial frequencies.

$$Spectrum = \int_{0_+}^{+\infty} \int_{0_+}^{+\infty} ||I(u,v)|| dudv \tag{2.14}$$

$$SBR = \frac{\int_{u_2}^{u_3} \int_{v_2}^{v_3} ||I(u,v)|| dudv}{\int_{u_1}^{u_2} \int_{v_1}^{v_2} ||I(u,v)|| dudv} \tag{2.15}$$

where $(0,0) \leq (u_1, v_1) < (u_2, v_2) < (u_3, v_3) < (+\infty, +\infty)$.

2.2 The evaluation of focus measures

The evaluation of focus measures is often done differently depending on what is meant to be achieved by using the focus measure in the first place. According to [3] there is, for example, a difference between how the field of autofocusing and the field of multi-focus image fusion. Within autofocusing the criteria mentioned in Section 2.1 can be used to evaluate the different focus measures, whilst within multi-focus image fusion there often exists several images which can be used to compare with one another.

To be able to perform an evaluation it is easiest to create graphs to compare the results of the different focus measures visually. The evaluation is then done through visual analysis of the graphs in accordance to the criteria to ascertain which focus measure performs the best. In fields like multi-focus image fusion one can use methods such as the root means square error to get a mathematical value for the results by comparing a created image with that of an already existing one, but for autofocus this isn't really viable.

In a similar report written by Huang and Jing where they implemented and compared different focus measures it was found that their performance in decreasing order are as follows:

1. SML
2. EOL
3. EOG, SF and Tenengrad
4. Variance

If Huang and Jing also accounted for the time it took for each algorithm to calculate the values their performance is listed as follows in decreasing order:

1. EOG
2. SF
3. EOL
4. Tenengrad
5. SML
6. Variance

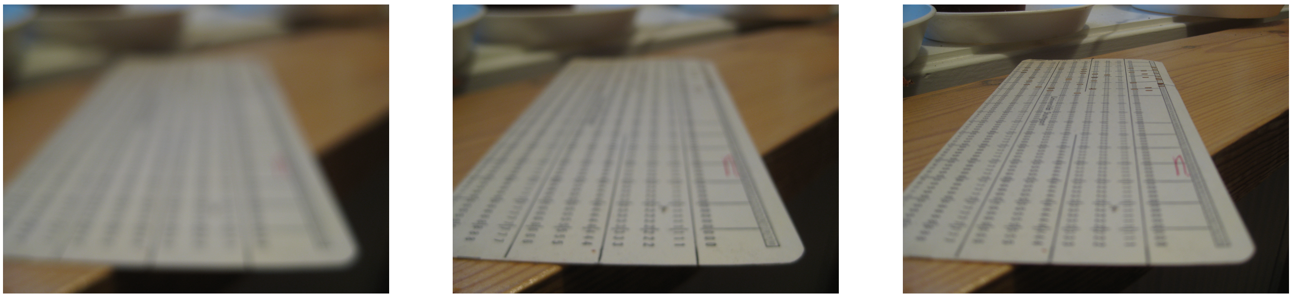
Chapter 3

Method

The first step of the implementation is to load in the desired image series. In this project three different image series were used, see figures 3.1 through 3.3 for selected images from the series.

(a) (b) (c)

Figure 3.1: Geode

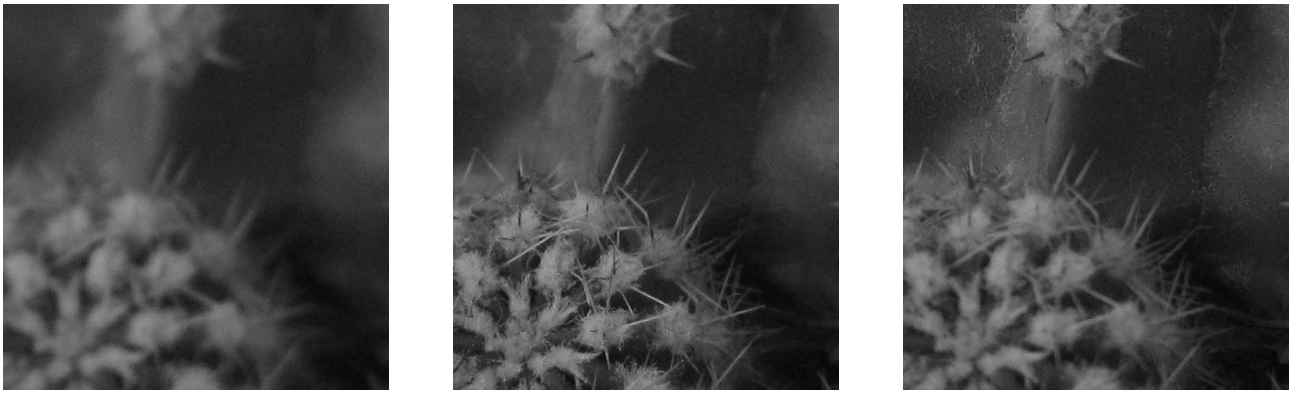


(a) (b) (c)

Figure 3.2: Card

After the images are loaded in one of them are shown using *imshow* and an area is selected with the help MATLAB's *roi drawrectangle* function. In return the location of the four corners of the rectangularly are given which are then used to calculate the width and height of the box as well as to crop the image if need be.

In a loop each image in the series is picked out, converted into double and made grayscale before being sent into a function. Depending on what function is being used, as there is one function for each focus measure, another step might be taken where the image is cropped so that only the selected square is sent. Alternatively the coordinates of the middle of the square and the desired width and height are sent.



(a) (b) (c)

Figure 3.3: Cactus

Each function calculates the value of the focus measure according to the equations given in 2.1 and the result is sent back. The result is a 1x27(Geode and Card images) or 1x26(Cactus images) vector containing the focus measure values of each of the 27 or 26 images in the image series, which can later be normalised between 0 to 1 and then plotted against each other for comparison.

3.1 Focus Measures

Eight different focus measures have been implemented in this project, six spatial focus measures and two Fourier based focus measures. Each focus measure has been implemented in its own MATLAB function.

The spatial focus measures have been implemented with the help of the equations in Chapter 2, whilst the Fourier based focus measures were implemented by using a threshold mask to separate high and low frequencies.

First the images are transformed over to the frequency domain then a threshold mask is defined as a part of the images original size, where all values outside of the mask are high frequencies and every value inside of the mask are low frequencies. The Spectrum measure is the sum of the real high frequency values, and the Spectrum Band Ratio (SBR) measure is the ratio of higher frequencies and the lower frequencies.

Chapter 4

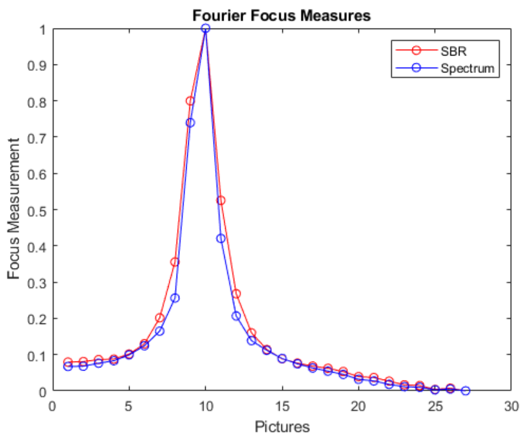
Results

In the following chapter the results of the testing are presented, which are later discussed in Chapter 5.

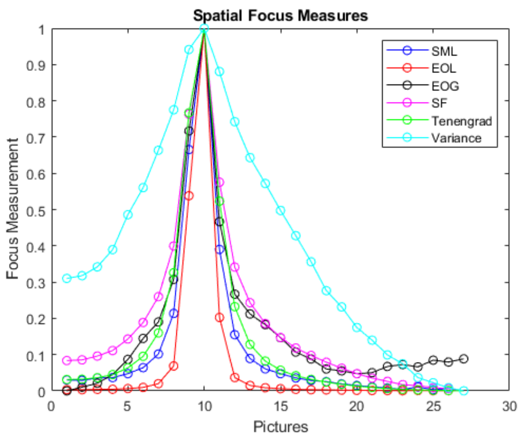
(a)

(b)

Figure 4.1: Geode result; (a) Chosen area, (b) Best image focus

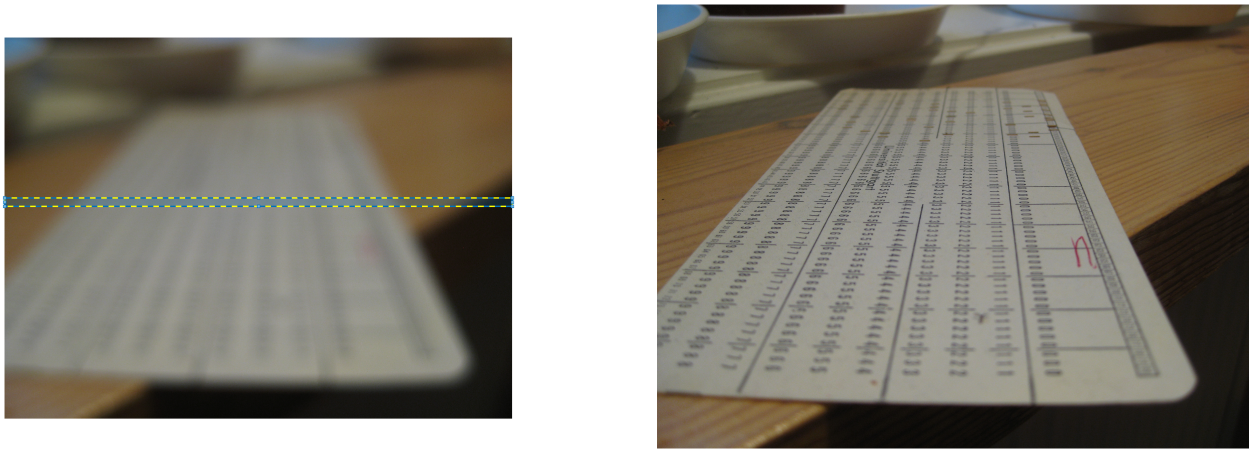


(a)



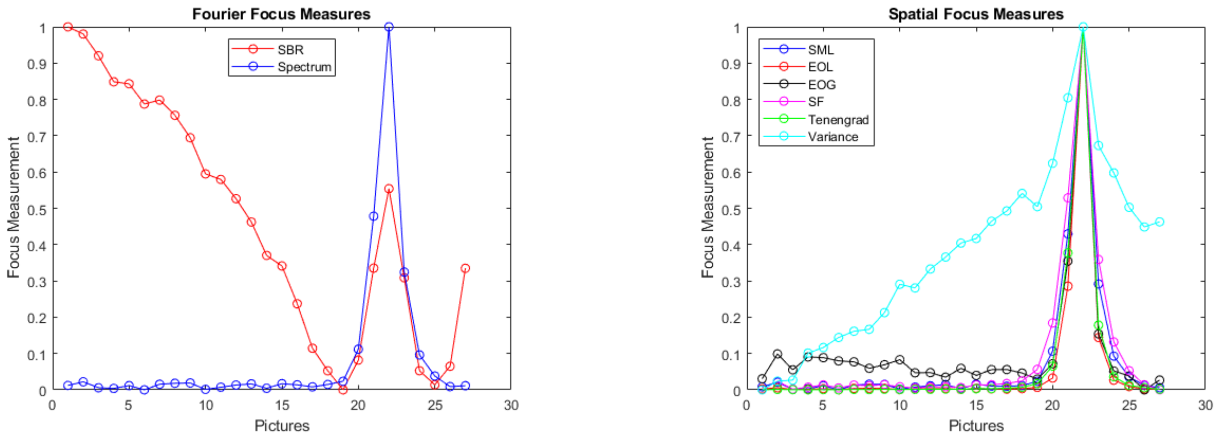
(b)

Figure 4.2: Geode graphs; (a) Fourier measures, (b) Spatial measures



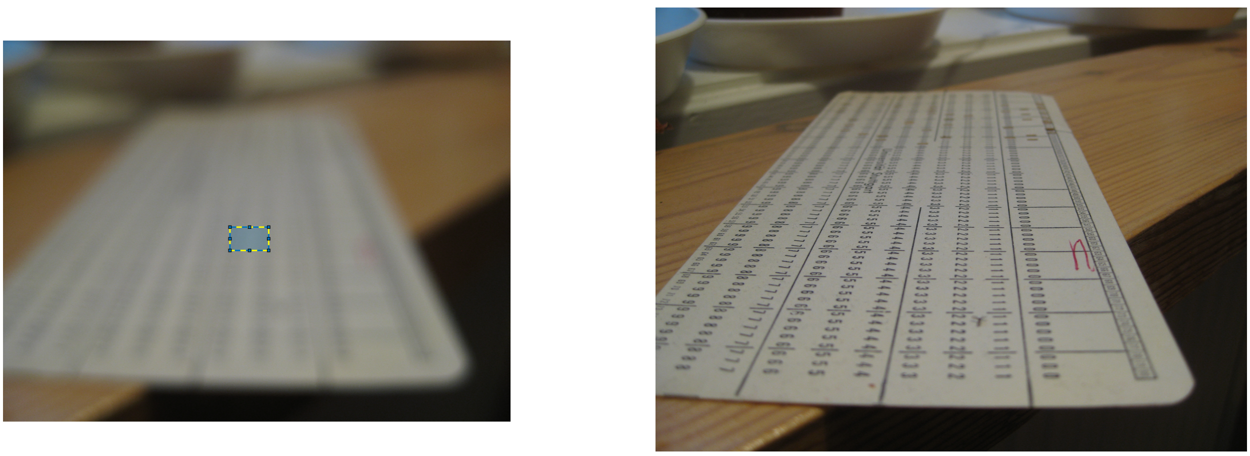
(a) (b)

Figure 4.3: Card 1 results; (a) Chosen area, (b) Best image focus



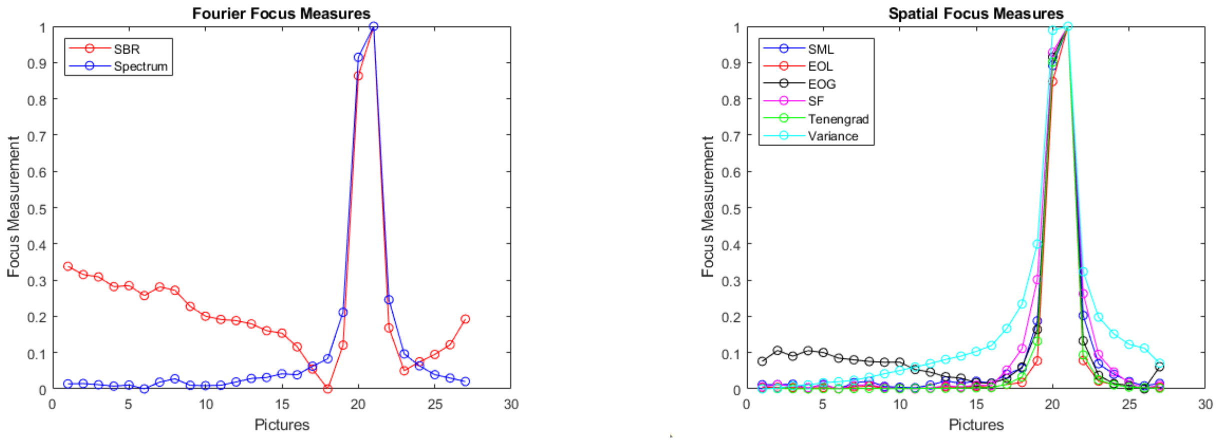
(a) (b)

Figure 4.4: Card 1 graphs; (a) Fourier measures, (b) Spatial measures



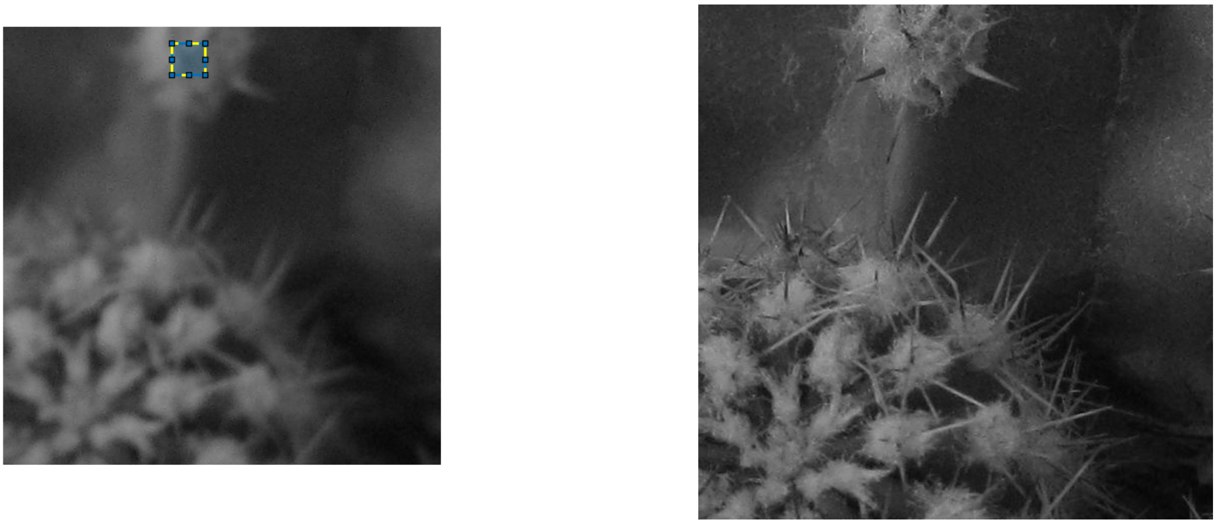
(a) (b)

Figure 4.5: Card 2 results; (a) Chosen area, (b) Best image focus



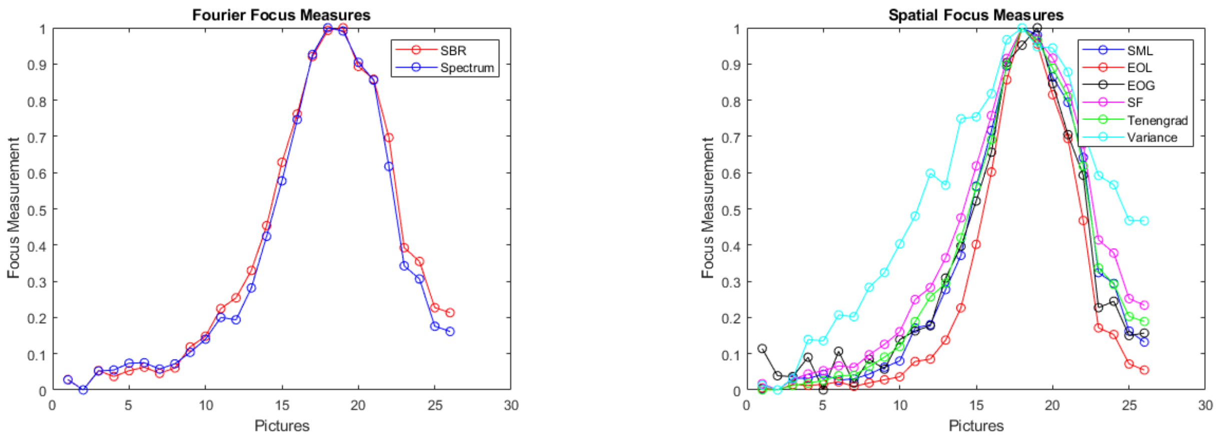
(a) (b)

Figure 4.6: Card 2 graphs; (a) Fourier measures, (b) Spatial measures



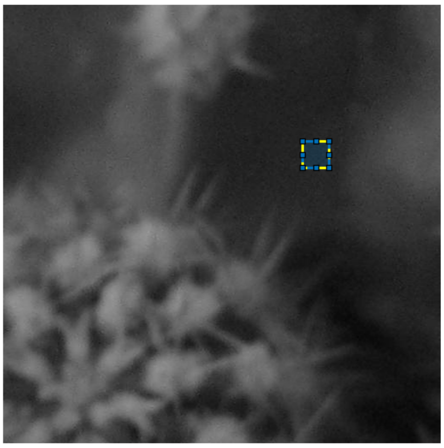
(a) (b)

Figure 4.7: Cactus results; (a) Chosen area, (b) Best image focus

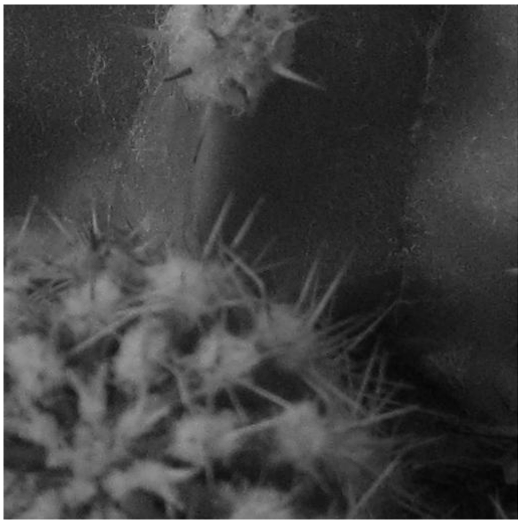


(a) (b)

Figure 4.8: Cactus graphs; (a) Fourier measures, (b) Spatial measures

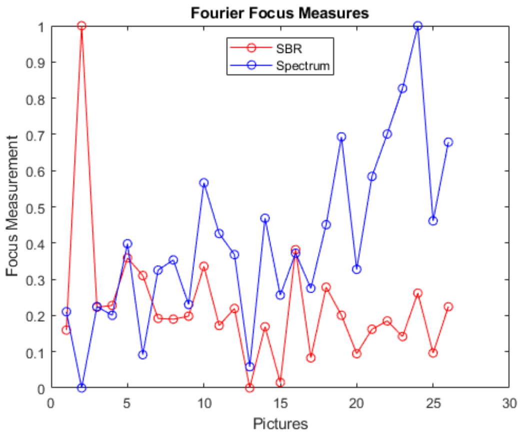


(a)

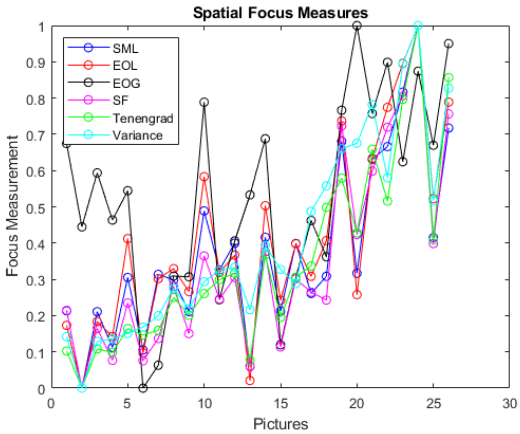


(b)

Figure 4.9: Cactus Background Results



(a)



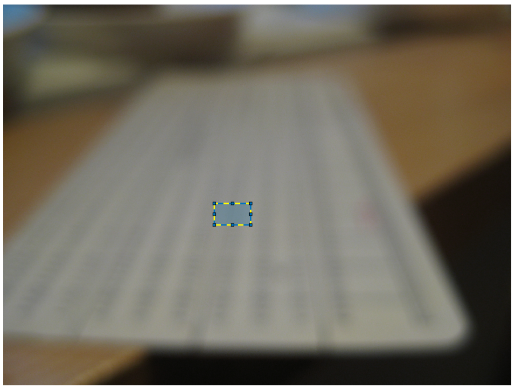
(b)

Figure 4.10: Cactus Background Results

4.1 Tests with the radius of the mask

The radius of the threshold mask decides the ratio between high and low frequencies. Changing the radius of the mask changes this ratio and affects the results of the Fourier based focus measures. In this section results with varying radius of the mask are presented.

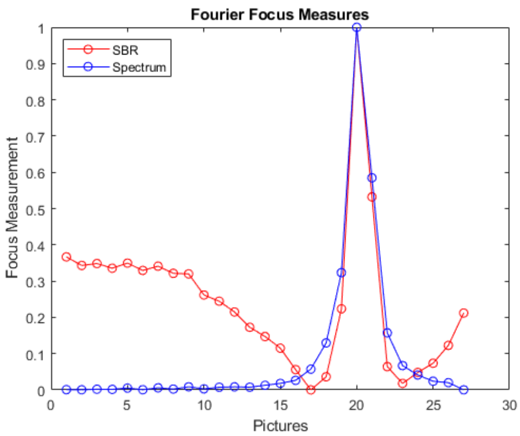
All pictures in the section above have SBR with radius 1/15 of the image size and Spectrum with radius 1/5 of the image size.



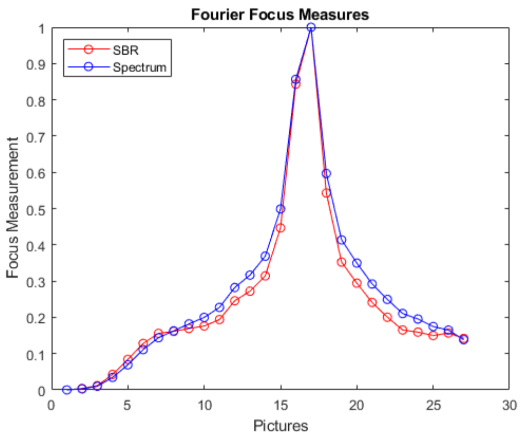
(a)

(b)

Figure 4.11: Image areas chosen

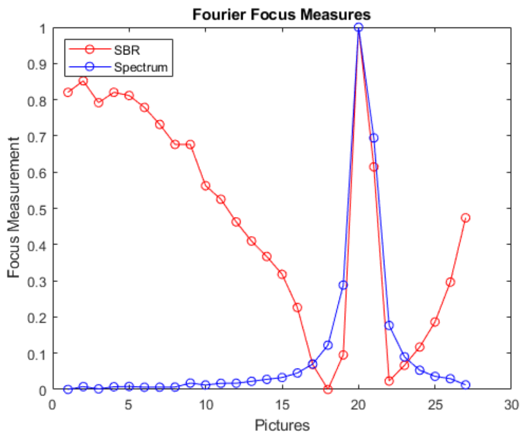


(a)

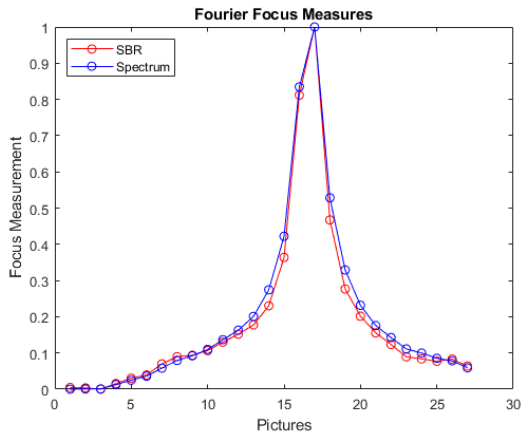


(b)

Figure 4.12: With a radius of 1/15 of image size; (a) Card square, (b)Geode cat figure

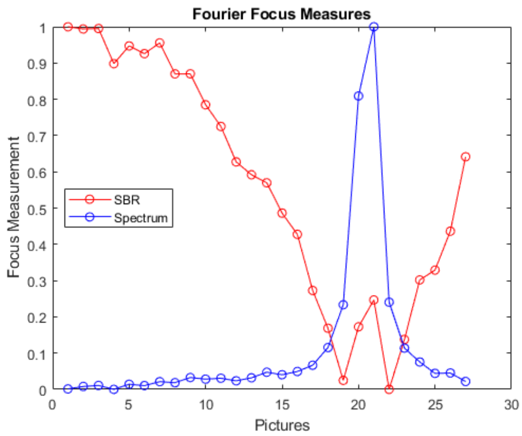


(a)

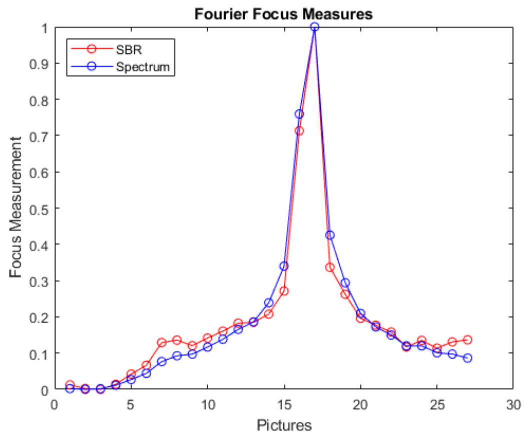


(b)

Figure 4.13: With a radius of 1/10 of image size; (a) Card square, (b)Geode cat figure



(a)



(b)

Figure 4.14: With a radius of 1/5 of image size; (a) Card square, (b)Geode cat figure

Chapter 5

Discussion

In the following chapter we will discuss the results presented in Chapter 4.

5.1 Spatial

According to Huang and Jing's research paper [3] SML had the best performance with EOL in close second. In this project we have performed a subjective human evaluation according to the six criteria presented in Section 2.1. We think that SML and EOL have similar performance and it varies between cases which one is the best. EOL usually has the best accuracy of the two but SML has a slightly better effective range. The difference between SML and EOL is minimal.

Out of the remaining spatial focus measures EOG, SF and Tenengrad are similar in performance, with some outlier cases, most of which can be found in the Appendix. It is of our subjective opinion that Variance is the worst out of all the focus measures tested, as it often times picks the incorrect image and fails many of the criteria given in Section 2.

5.2 Fourier

Unlike in [3] there isn't a clear ranking in [2] of which of the two Fourier transform based focus measures is the better one, simply stating that they are more computationally efficient than their spatial counterparts. Because of this we will discuss the criteria somewhat closer than we did for the spatial measures.

In most cases both SBR and Spectrum show Unimodality, but the former struggles with Monotonicity in certain images. Both measures have slight issues when trying to evaluate the background of the cactus image, which most likely is the result of a sensitivity to noise. SBR and Spectrum both, for the most part, give a smooth curve without jumps that are too sudden, suggesting an effective range. The only criteria that they both falter on is the variability criteria, as a change in radius might have to be done to optimise results.

However, the clearest difference between the two is quite clearly how even though they often pick the same image as that of the Spatial focus measures, SBR sometimes picks a completely different image. This might be for a number of reasons, but we have been unable to identify a definitive one. However, before we clearly state which of the two is better than the other, we first need to discuss the result of the radius tests.

5.3 Radius tests

As seen in Section 4.1 there are certain cases where changing the radius of the mask in SBR and Spectrum will lead to different results. When using SBR on Figure 4.11(a) a smaller radius like 1/15 gives a worse result than a larger radius, whilst Spectrum only somewhat changes shape (the square chosen isn't in the exact same location for each test, therefore the result will not be the exact same image). Meanwhile in Figure 4.11(b) a smaller radius simply gives a more narrow result of the curve for both SBR and Spectrum.

This difference shows that whilst Spectrum doesn't get affected much when adjusting the radius SBR can, depending on the image, become less accurate. Because of this issue it is difficult to know when to use and how much one can trust the result of SBR and we would therefore not recommend using SBR on its own. If you still wish to use it we recommend doing so in combination with other focus measures to be able to compare the results to avoid identifying focus incorrectly.

5.4 Comparison

According to our subjective evaluation we believe SML and EOL to be the best out of all the focus measures that have been tested during this project. Spectrum, EOG, SF and Tenengrad we place on approximately the same level in terms of performance. As SBR's positioning in the ranking heavily depends on what image is being analysed and the radius of the mask, we have decided to rank it on the next to lowest spot in the ranking, with Variance at the very bottom. Our final subjective ranking is therefore as follows:

1. SML and EOL
2. Spectrum, EOG, SF and Tenengrad
3. SBR
4. Variance

Bibliography

[1] Todd Vorenkamp. How focus works. `https://www.bhphotovideo.com/explora/photography/tips-and-solutions/how-focus-works`, 2015.

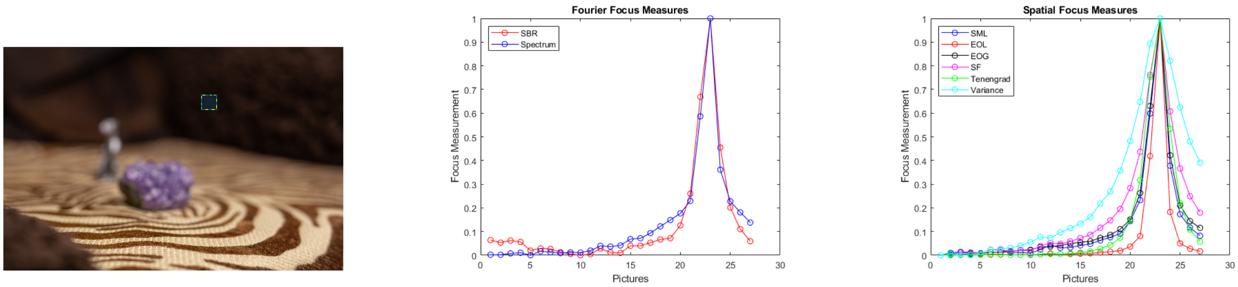
[2] Yibin Tian. Monte Carlo evaluations of ten focus measures. In Russel A. Martin, Jeffrey M. DiCarlo, and Nitin Sampat, editors, *Digital Photography III*, volume 6502, page 65020C. International Society for Optics and Photonics, SPIE, 2007.

[3] Wei Huang and Zhongliang Jing. Evaluation of focus measures in multi-focus image fusion. *Pattern Recognition Letters*, 28(4):493–500, 2007.

[4] S.K. Nayar and Y. Nakagawa. Shape from focus. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(8):824–831, 1994.

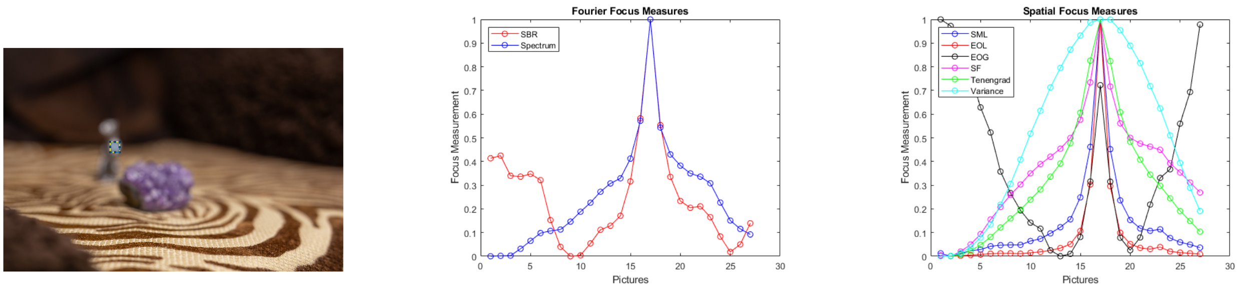
Chapter 6

Appendix



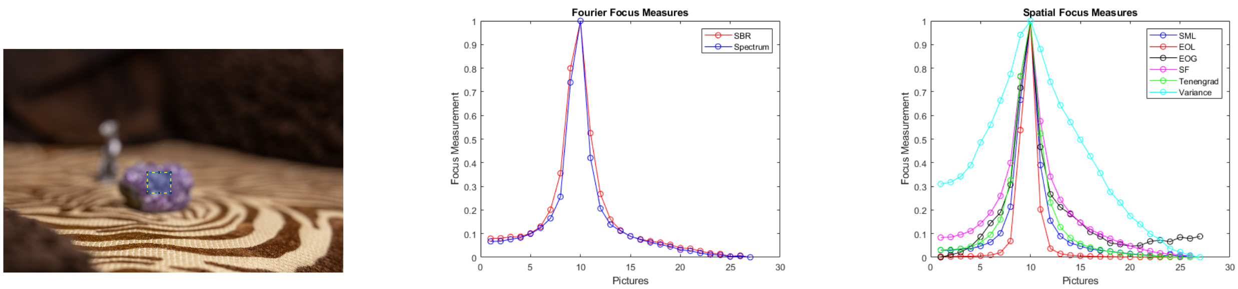
(a) (b) (c)

Figure 1: Geode Background Results



(a) (b) (c)

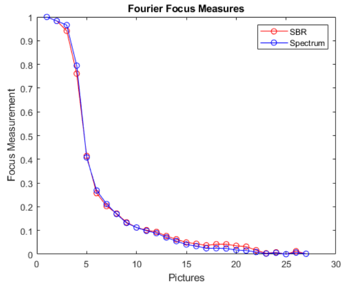
Figure 2: Geode Cat Results



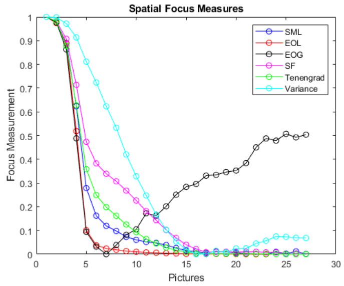
(a) (b) (c)

Figure 3: Geode Geode Results

(a)

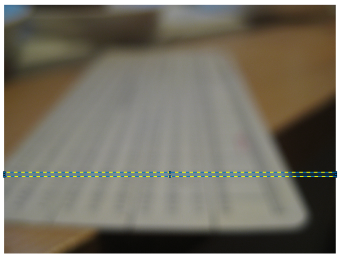


(b)

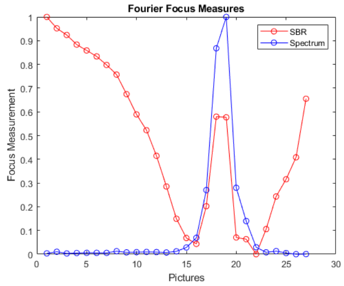


(c)

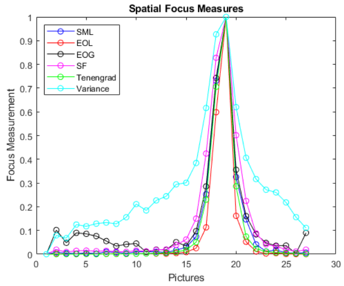
Figure 4: Geode Foreground Results



(a)

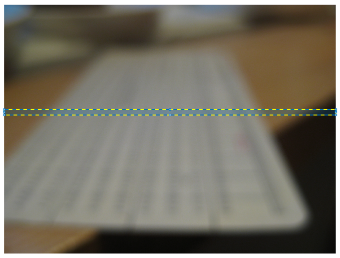


(b)

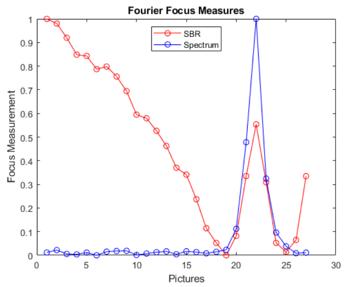


(c)

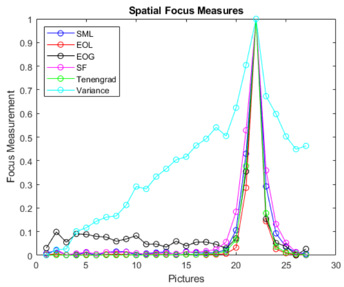
Figure 5: Card Foreground Results



(a)

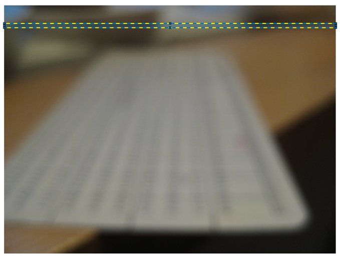


(b)

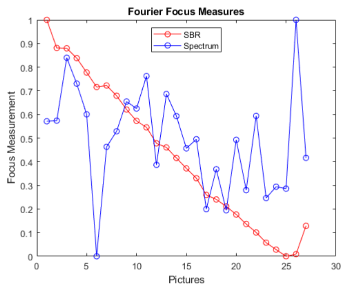


(c)

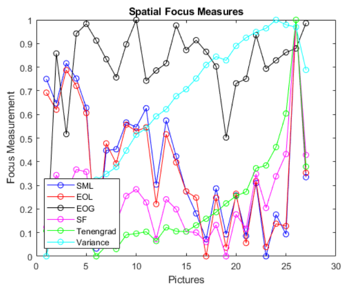
Figure 6: Card Middle Results



(a)

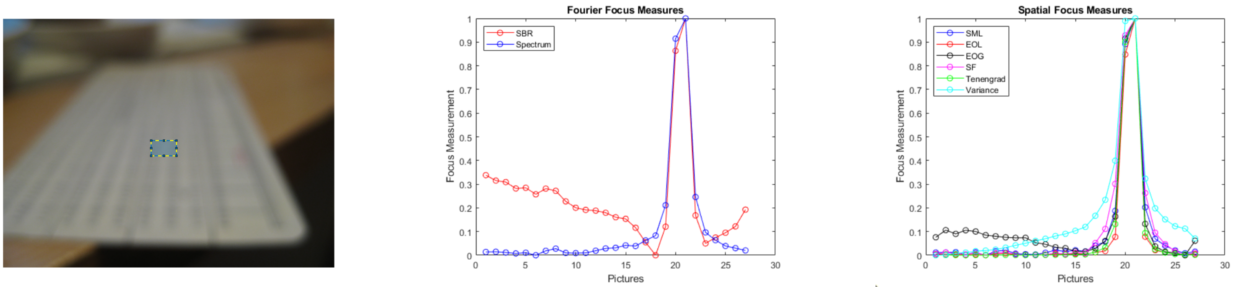


(b)



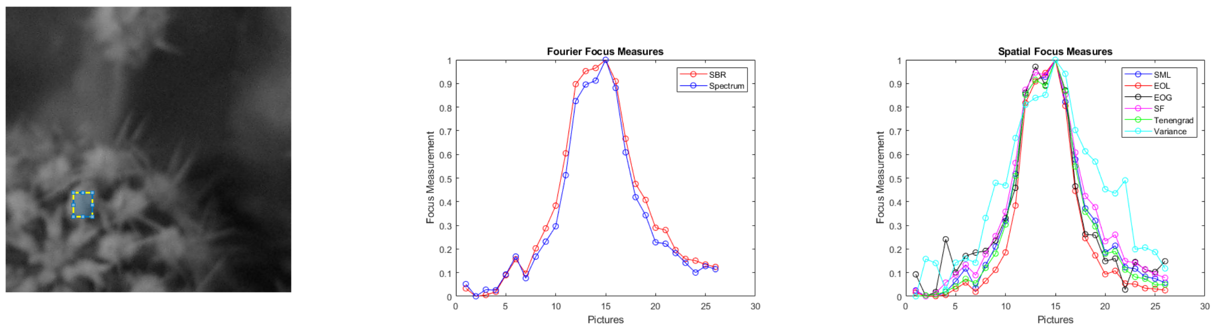
(c)

Figure 7: Card Background Results



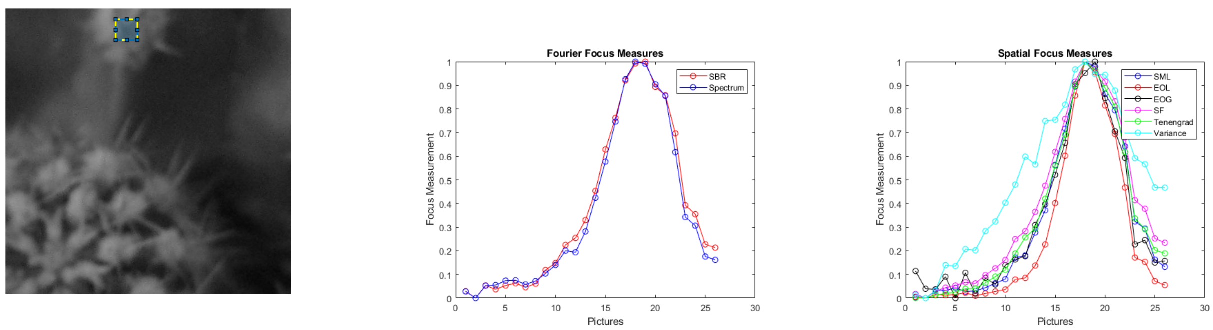
(a) (b) (c)

Figure 8: Card Square Results



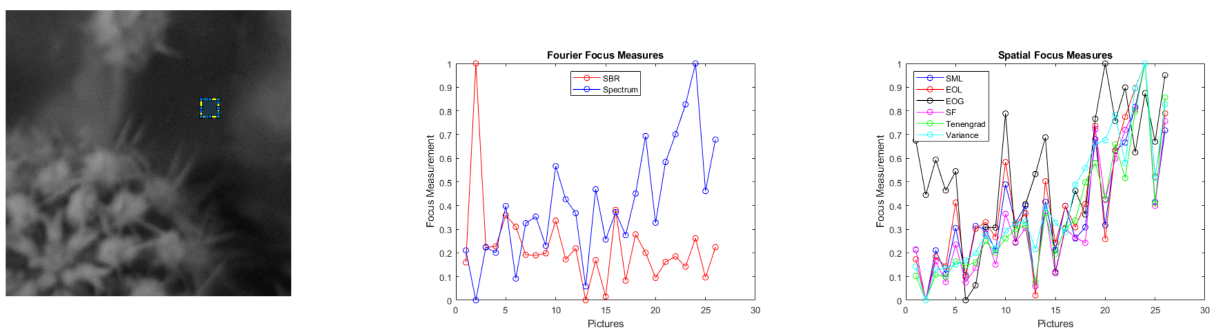
(a) (b) (c)

Figure 9: Cactus Foreground Results



(a) (b) (c)

Figure 10: Cactus Bulb Results



(a) (b) (c)

Figure 11: Cactus Background Results