

ABSTRACT

Title of thesis: TEXTURE-DETAIL PRESERVATION
 MEASUREMENT IN CAMERA PHONES:
 AN UPDATED APPROACH

Nitin Suresh, Master of Science, 2016

Thesis directed by: Professor Yu Chen,
 The Fischell Department of Bioengineering

Recent advances in mobile phone cameras have poised them to take over compact hand-held cameras as the consumer's preferred camera option. Along with advances in the number of pixels, motion blur removal, face-tracking, and noise reduction algorithms have significant roles in the internal processing of the devices. An undesired effect of severe noise reduction is the loss of texture (i.e. low-contrast fine details) of the original scene. Current established methods for resolution measurement fail to accurately portray the texture loss incurred in a camera system. The development of an accurate objective method to identify the texture preservation or texture reproduction capability of a camera device is important in this regard.

The 'Dead Leaves' target has been used extensively as a method to measure the modulation transfer function (MTF) of cameras that employ highly non-linear noise-reduction methods. This stochastic model consists of a series of overlapping circles with radii r distributed as r^{-3} , and having uniformly distributed gray level,

which gives an accurate model of occlusion in a natural setting and hence mimics a natural scene. This target can be used to model the texture transfer through a camera system when a natural scene is captured.

In the first part of our study we identify various factors that affect the MTF measured using the ‘Dead Leaves’ chart. These include variations in illumination, distance, exposure time and ISO sensitivity among others. We discuss the main differences of this method with the existing resolution measurement techniques and identify the advantages.

In the second part of this study, we propose an improvement to the current texture MTF measurement algorithm. High frequency residual noise in the processed image contains the same frequency content as fine texture detail, and is sometimes reported as such, thereby leading to inaccurate results. A wavelet thresholding-based denoising technique is utilized for modeling the noise present in the final captured image. This updated noise model is then used for calculating an accurate texture MTF. We present comparative results for both algorithms under various image capture conditions.

TEXTURE-DETAIL PRESERVATION MEASUREMENT
IN CAMERA PHONES: AN UPDATED APPROACH

by

Nitin Suresh

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Master of Science
2016

Advisory Committee:
Professor Yu Chen, Chair/Advisor
Professor Behtash Babadi
Dr. Quanzeng Wang

© Copyright by
Nitin Suresh
2016

*This work is dedicated to my parents, for their love, support and
constant encouragement.*

Acknowledgments

I would like to express my sincere gratitude to my advisor Prof. Yu Chen for his continuous support and guidance during the course of this research. I would like to thank my thesis committee members: Dr. Quanzeng Wang and Prof. Behtash Babadi for their constant encouragement and for their insightful questions, which generated countless avenues for research.

I would like to specially thank Dr. Quanzeng Wang and Dr. Joshua Pfefer for their guidance and for enabling me to carry out my thesis work at the Food and Drug Administration, White Oak. I would like to thank Dr. Pejman Ghassemi and Dr. Jianting Wang for their help and support. I would also like to acknowledge financial support from the Critical Path Funding at the Food and Drug Administration during the course of the projects discussed in this work.

I would like to thank my family and friends for their love and encouragement, and for supporting me in all of my endeavours.

Table of Contents

| | |
|----------------------------------|-----|
| List of Tables | vi |
| List of Figures | vii |
| List of Abbreviations | x |
| 1 Introduction | 1 |
| 2 Literature Review | 5 |
| 2.1 SR-IQA | 6 |
| 2.2 FR-IQA | 8 |
| 2.3 Summary | 9 |
| 3 The Dead Leaves Model | 10 |
| 4 Artificial Sharpening | 13 |
| 4.1 The CSF | 14 |
| 4.2 Unsharp Masking | 16 |
| 4.3 Results | 18 |
| 5 Distance from Camera to Target | 23 |
| 5.1 Results | 24 |
| 5.2 Conclusion | 27 |
| 6 Illumination Variation | 31 |
| 7 Exposure Time | 38 |
| 7.1 ISO32 | 39 |
| 7.2 ISO200 | 41 |
| 7.3 ISO800 | 42 |

| | | |
|--------|---|----|
| 8 | ISO Speed | 43 |
| 8.1 | Exposure Time Setting 1 | 43 |
| 8.2 | Exposure Time Setting 2 | 45 |
| 8.3 | Exposure Time Setting 3 | 46 |
| 9 | Noise Addition | 49 |
| 9.1 | Gaussian Noise Model | 49 |
| 9.2 | Poisson Noise Model | 52 |
| 9.3 | Salt and Pepper Noise Model | 55 |
| 9.4 | Conclusions | 57 |
| 10 | Implementation of the Wavelet Thresholding Approach | 60 |
| 10.1 | Region for estimating the noise variance | 61 |
| 10.2 | Wavelet Thresholding for Image Denoising | 68 |
| 10.2.1 | The Discrete Wavelet Transform | 69 |
| 10.2.2 | Denoising via Wavelet Thresholding | 73 |
| 10.2.3 | Results | 77 |
| 10.3 | Modification to the texture MTF algorithm | 87 |
| 11 | Conclusion | 93 |
| 11.1 | Variation of the texture MTF with various parameters | 93 |
| 11.2 | Modification of the calculation of the noise spectrum | 94 |
| | Bibliography | 95 |

List of Tables

| | | |
|------|--|----|
| 2.1 | Coefficients for the ideal PSD extension, as proposed by McElvain et al. | 7 |
| 5.1 | % coverage of HFOV at different distances, Canon DSLR | 25 |
| 5.2 | % Coverage of HFOV at different distances, iPhone 5s | 28 |
| 5.3 | % Coverage of HFOV at different distances, Nexus 5 | 28 |
| 10.1 | Noise standard deviations obtained for the gray region and texture regions from the $Image_{noise}$, obtained for 10 images, using the weighted average denoised estimate image | 64 |
| 10.2 | Noise standard deviations obtained for the gray region and texture regions from the $Image_{noise}$, obtained for 10 images, using the median-sorted denoised estimate image. | 64 |
| 10.3 | PSNR values of the denoised image obtained after denoising using wavelet thresholding, under variation of several parameters | 78 |
| 10.4 | Variation in the denoised image PSNR with the wavelet used for the wavelet transform | 84 |

List of Figures

| | | |
|-----|--|----|
| 3.1 | Generated dead leaves target. | 12 |
| 4.1 | Campbell-Robson CSF chart. | 15 |
| 4.2 | The Contrast Sensitivity Function. | 16 |
| 4.3 | Effect of varying the 'amount' factor, at various illumination intensity levels | 20 |
| 4.4 | Effect of varying both the 'amount' and 'radius' factors, at various illumination intensity levels | 21 |
| 4.5 | Results for texture MTF calculation for images sharpened varying the amount parameter, at various illumination levels. | 22 |
| 4.6 | Results for texture MTF calculation for images sharpened varying the radius and amount parameters, at various illumination levels. | 22 |
| 5.1 | Types of distortion | 24 |
| 5.2 | Variation in texture MTF with distance for the Canon DSLR camera. | 26 |
| 5.3 | Variation of acutance with distance of camera to target - Canon DSLR | 27 |
| 5.4 | Variation in texture MTF with distance for the iPhone 5s. | 28 |
| 5.5 | Variation of acutance with distance of camera to target, iPhone 5s. | 29 |
| 5.6 | Variation in texture MTF with distance for the Nexus 5. | 29 |
| 5.7 | Variation of acutance with distance of camera to target, Nexus 5. | 30 |
| 6.1 | Variation of texture acutance with illumination levels for the iPhone 5s, using the Eiko bulbs. | 32 |
| 6.2 | Exposure time and ISO sensitivity values for the images obtained at the different intensity levels, for iPhone 5s automatic metering mode, using Eiko bulbs. | 33 |
| 6.3 | Power Spectral Densities for the light sources utilized | 34 |
| 6.4 | Comparison of the texture acutance obtained for the iPhone 5s for two light sources, the Eiko Supreme and the Philipps Halogena, at different light intensities. | 35 |
| 6.5 | Exposure time and ISO sensitivity values for the images obtained at the different intensity levels, for the iPhone 5s, using the Philipps Halogena bulbs. | 36 |

| | | |
|------|---|----|
| 6.6 | Comparison of the texture acutance obtained for the Nexus 5 for two light sources, the Eiko Supreme and the Philipps Halogena, at different light intensities. | 37 |
| 6.7 | Overall texture acutance for the iPhone 5s and the Nexus 5. | 37 |
| 7.1 | Variation in texture MTF with exposure time, at ISO32. | 39 |
| 7.2 | Acutance values for different exposure times, at ISO32, ISO200 and ISO800. | 40 |
| 7.3 | Variation in texture MTF with exposure time, at ISO200. | 41 |
| 7.4 | Variation of texture MTF with exposure time, at ISO800. | 42 |
| 8.1 | Variation of texture MTF with ISO Speed, at 1/50 s exposure time. | 44 |
| 8.2 | Acutance values at different ISO speeds, at 1/50 s exposure time. | 45 |
| 8.3 | Sample images obtained at different ISO speeds, at 1/160 s exposure time. | 46 |
| 8.4 | Variation in texture MTF with ISO speed, at 1/160 s exposure time. | 47 |
| 8.5 | Acutance values for different ISO speeds, at 1/160 s exposure time. | 47 |
| 8.6 | Variation in texture MTF with ISO speed, at 1/550 s exposure time. | 48 |
| 8.7 | Acutance values for different ISO speeds, at 1/550 s exposure time. | 48 |
| 9.1 | The figure shows a comparison between the original image and image with Gaussian noise added. The mean and variance levels are stated with reference to image having intensity between 0 and 1. | 51 |
| 9.2 | Variation in texture MTF, Gaussian noise images. | 52 |
| 9.3 | Acutance values for images with added Gaussian noise. | 53 |
| 9.4 | Variation in texture MTF on addition of Gaussian noise of non-zero mean. | 54 |
| 9.5 | Acutance value comparison for images with added Gaussian noise. | 55 |
| 9.6 | Variation of texture MTF with addition of Poisson noise. | 56 |
| 9.7 | Change in acutance observed on addition of Poisson noise. | 57 |
| 9.8 | Noisy image generated with salt and pepper noise, with noise density $d=0.02$ | 58 |
| 9.9 | Variation in texture MTF with addition of salt and pepper noise with various noise densities. | 59 |
| 9.10 | Acutance values for images with added salt and pepper noise. | 59 |
| 10.1 | Example Dead Leaves chart. | 61 |
| 10.2 | The mean noise values obtained for the gray regions, for the different sets of images. | 66 |
| 10.3 | The noise standard deviations obtained for the gray regions, for the different sets of images. | 67 |
| 10.4 | The mean values obtained for the texture regions, for the different sets of images. | 68 |
| 10.5 | The noise standard deviations obtained for the texture regions, for the different sets of images. | 69 |
| 10.6 | Daubechies-4 wavelet and scaling functions | 71 |

| | | |
|-------|---|----|
| 10.7 | Symmlet-8 wavelet and scaling functions | 71 |
| 10.8 | Coiflet-4 wavelet and scaling functions | 72 |
| 10.9 | The Mallat algorithm for the DWT decomposition. | 72 |
| 10.10 | Description of the 2-D DWT decomposition in images showing the HH, HL, LH sub-bands, and further decomposed LL sub-band. | 73 |
| 10.11 | Functions for soft and hard thresholding. | 75 |
| 10.12 | Denoising performance for an image having Gaussian noise with mean=0 and variance=0.0005. | 78 |
| 10.13 | Denoising performance for an image having Gaussian noise with mean=0 and variance=0.005. | 79 |
| 10.14 | Denoising performance for an image having salt & pepper noise with noise density=0.02. | 79 |
| 10.15 | Texture MTF comparison using the proposed algorithm | 90 |
| 10.16 | Dead leaves target image at various noise levels | 91 |
| 10.17 | Texture MTF comparison using the proposed algorithm at various ISO levels and exposure times | 92 |

List of Abbreviations

| | |
|------|-------------------------------|
| MTF | Modulation Transfer Function |
| SFR | Spatial Frequency Response |
| SNR | Signal to Noise Ratio |
| PSNR | Peak Signal to Noise Ratio |
| IQA | Image Quality Assessment |
| NSS | Natural Scene Statistics |
| USM | Unsharp Masking |
| PSD | Power Spectral Density |
| CSF | Contrast Sensitivity Function |
| FOV | Field of View |
| HFOV | Horizontal Field of View |
| MSE | Mean Squared Error |
| DWT | Discrete Wavelet Transform |

Chapter 1: Introduction

The current generation of mobile phone cameras have reached levels of technological advancement which make them comparable to the low-to-mid range consumer cameras. This factor combined with the ease of transport, multi-functionality, and attractive price point has led the mobile phone camera to be the first choice device for the everyday photographer. Noise-removal, sharpening, hot pixel identification, red-eye removal and several more features are commonplace in these devices. Along with normal everyday usage, specialized apps for medical image diagnosis or mobile dermatology have also become easily available in the market. This area has expanded because of the ease with which initial medical analyses can be made remotely with minimal time and resource expenditure. At the same time, the overall accuracy of the diagnosis depends, to a large extent, on the quality of the image produced. The quality of mobile phone camera images includes several parameters such as the spatial frequency response (SFR), color uniformity, visual noise, chromatic aberration, chroma level, texture preservation/texture reproduction etc. In this report, the main parameter that we consider will be the texture preservation metric of image quality in mobile smart phones.

The texture preservation measurement is important in camera phones that em-

ploy aggressive noise reduction algorithms in their image processing pipelines. These algorithms can mistake high-frequency texture detail for noise in several cases, and smooth out the image to an extent which removes the texture detail. A parameter is required to quantify this loss and to identify devices in which there is a high degree of this texture loss. The generally used metric image resolution, the slanted edge spatial frequency response, is sometimes fooled by such image processing algorithms, since edges are detected and processed differently in the image. As a consequence, the dead leaves target and related algorithms have been developed for the texture preservation measurement.

The general technique used for the evaluation of image quality is to use particular targets that are specially designed to evaluate particular metrics of image quality. Examples include the slanted edge chart for spatial resolution, the dot chart for chromatic aberration, the dead leaves target for texture preservation etc. Images of the target chart are captured by the camera system under certain specified conditions, and the properties of the images are evaluated in order to estimate the metrics under study. The ‘Dead Leaves’ target has been used extensively to measure the modulation transfer function (MTF) of cameras that employ highly non-linear noise-reduction methods. This stochastic model consists of a series of overlapping circles with radii r distributed as r^{-3} , and having uniformly distributed gray level, which gives an accurate model of occlusion in a natural setting and hence mimics a natural scene. This property of the dead leaves target makes it suitable for evaluation of the texture preservation aspect in natural images, which mainly consist of textures and images with fine detail.

We present a literature review of the existing algorithms for the measurement of the texture preservation metric using the dead leaves chart, along with extensions of the same. The algorithms mainly utilize the variation in the power spectrum of the ideal target image and the captured image, since the aberrations in the power spectral density (PSD) of the captured image are due to the image processing occurring in the camera system.

In the first part of the report we identify various factors that affect the texture MTF measured using the ‘Dead Leaves’ chart. These include variations in illumination, distance, exposure time and ISO sensitivity, sharpness and the effect of added noise models. An in-depth study of the effects of varying each of the parameters would give an idea as to the ideal conditions under which the best performance of the camera system can be expected. Easily reproducible conditions would mean that the test procedure can be carried out for a wide array of devices, with results that can be accurately compared. We also examine the usage of the texture acutance measurement as a single-value measure. This enables a numerical comparison of the texture preservation quality between camera systems.

An improvement to the texture MTF algorithm is proposed in the second part of the report. Images captured at very high ISO values, or under non-uniform illumination conditions can have considerable amounts of noise in them. These high-frequency variations in image intensity contain the same frequency content as some fine texture details. Due to this similarity in frequency information, the algorithm can artificially increase the texture MTF at high frequencies in some situations when noise in the image can be misinterpreted as texture detail. We examine the effec-

tiveness of using an estimation of the noise level in the texture image by using the uniform gray region, and propose an alternative noise estimation technique utilizing wavelet thresholding. The effects on the denoised image peak signal-to-noise ratio (PSNR) of various parameters in the thresholding step are considered, and suitable parameters are estimated based on these. After incorporating the modification in noise estimation, the texture MTF is calculated for a range of images with variations in noise level, ISO sensitivity and exposure time.

Chapter 2: Literature Review

Evaluation of the image quality of a camera with respect to its texture preservation aspect requires a different approach than just evaluating the sharpness measure. The dead leaves target is used for measurement of the texture preservation, as it models the occlusion phenomenon commonly seen in naturally occurring textures.

Various algorithms have been proposed to measure the texture Modulation Transfer Function (MTF) using the dead leaves target. The algorithms for image quality assessment (IQA) can mainly be classified into two categories:

1. *Semi-reference IQA (SR-IQA)* methods, which estimate the quality of distorted images with respect to subjective perceptual measures without having to use a reference of the original image; and
2. *Full-reference IQA (FR-IQA)* methods, which measure local pixel-wise disparity between the reference image and the obtained distorted image, and then obtain a scalar representing the total image quality metric of the distorted image.

2.1 SR-IQA

The SR-IQA methods for measuring texture preservation utilize the Natural Scene Statistics (NSS) approach: some of the statistics of the dead leaves target follow the same statistics as observed in natural images. Cao et al. [1] [2] proposed that the power spectrum of the dead leaves target follows a power law. By using this property, the power spectral density (PSD) of an ideal dead leaves model at any size can be obtained, as given in Equation 2.1. By obtaining the PSD of the captured image, and then dividing it by the calculated ideal PSD, the texture MTF of an imaging system can be calculated.

$$|\hat{U}_{mn}|^2 = \begin{cases} L^4 \langle u \rangle^2, & m = 0, n = 0 \\ \frac{A(L)}{(\frac{\sqrt{m^2+n^2}}{L})^p}, & otherwise \end{cases} \quad (2.1)$$

where p is the empirical power law coefficient, measured as $p \sim 1.857$ [2], $\langle u \rangle$ the average image intensity, m and n are the coordinates considered, L the image length (in numbers of pixels), and $A(L)$ the normalization coefficient. This model is for the case when the Fourier spectrum is shifted and the location $(m, n) = (0, 0)$ appears at the center of the transformed image. The normalization coefficient $A(L)$ is approximated by $71.0156 * L^{1.8905}$. The 2-D power spectral density was then converted to 1-D power spectral density using radial averaging, by averaging over all orientations. The texture MTF was then calculated as given in Equation 2.2.

$$MTF_{texture} = \sqrt{\frac{PSD_{measured}}{PSD_{ideal}}} \quad (2.2)$$

McElvain et al. [3] provided several extensions to the method proposed by

Cao et al. It was shown that the power spectrum of the dead leaves model does not follow an exact power law across the entire spatial frequency range. The deviation from the power law was attributed to the limitation in bounding of the radii of the circles in the dead leaves model. A modified model for the ideal PSD was proposed, as given in Equation 2.3.

$$\ln[PSD_{ideal}(f)] = \ln[A(L)] - B(s) \ln(f) - C(s)[\ln(f)]^2 \quad (2.3)$$

The $\ln[A(L)]$, B and C coefficient values depend on the captured image size L as given in Table 2.1. The factor s is the scale factor obtained by N/L , where N is the rendered image length (in numbers of pixels).

Table 2.1: Coefficients for the ideal PSD extension, as proposed by McElvain et al.

| L=N/s | ln[A(L)] | B(s) | C(s) |
|--------------|-----------------|-------------|-------------|
| 128 | 12.531 | 2.295 | 0.09991 |
| 256 | 13.985 | 2.400 | 0.12613 |
| 512 | 15.476 | 2.407 | 0.12067 |
| 1024 | 16.690 | 2.718 | 0.19723 |
| 2048 | 17.253 | 3.601 | 0.39951 |

The second important extension was in the identification that the captured image PSD at higher spatial frequencies would be dominated by the high frequency noise and artifacts, thereby causing an artificial and false increase in the captured image PSD at higher spatial frequency levels. The noise is generally sensor noise and

JPEG artifacts. Correction of this parameter is performed by first subtracting the noise PSD from the measured PSD before using it for the texture MTF calculation, as given in Equation 2.4. The noise PSD is measured from the PSD of a 50% uniform grey patch, using the same camera system and shooting conditions.

$$MTF_{texture} = \sqrt{\frac{PSD_{measured} - PSD_{noise}}{PSD_{ideal}}} \quad (2.4)$$

2.2 FR-IQA

The FR-IQA methods perform local methods comparing each pixel in the reference image to the corresponding pixel in the captured image, for obtaining the difference. Kirk et al. proposed a technique for texture preservation measurement using the dead leaves target that utilizes the phase information as well, using the cross-correlation power density $\phi_{YX}(f)$ between the input and output signals, and the auto power density $\phi_{XX}(f)$ [4]. The full transfer function $H(f)$ is obtained as in Equation 2.5.

$$H(f) = \frac{\phi_{YX}(f)}{\phi_{XX}(f)} \quad (2.5)$$

Utilizing registration of the measured pixel pattern to the original pixel pattern is performed using projective transformation. Fourier transformation of these images provides the spectrums of the ideal and measured images, from which the cross-correlation power density, the auto power density and the texture MTF are then obtained by Equation 2.6.

$$MTF_{texture} = average_{radial}(| H(f) |) \quad (2.6)$$

2.3 Summary

This chapter presented an overview of the existing algorithms for texture preservation measurement in camera devices. The two main categories are semi-reference and full-reference methods. The semi-reference methods calculate the texture MTF by modelling the ideal power spectral density according to certain parameters. The full-reference methods utilize additional spatial information of the captured and ideal image to obtain the texture MTF.

The model generation data for the original target is required in order for the utilization of a full-reference method. This model data may not be available at the testing site. There is also the possibility of errors being introduced in the measurement pipeline during the process of spatial registration between the measured and the ideal images. Due to these reasons, we concentrate on the semi-reference method for texture preservation measurement in this study.

Chapter 3: The Dead Leaves Model

The circular dead leaves target is a reproducible model for the occlusion phenomenon commonly observed in natural images. It consists of a number of overlapping discs with the overall image representing a collection of dead leaves, i.e. a collage of discrete objects that partially occlude one another. Lee et al. [5] provide a model for the generation of the dead leaves target. Consider a uniform Poisson process $s_i = \{x_i, y_i, z_i\}$, and the closed sets T_i in \mathbb{R}^2 , having random size r_i and centered at the origin. The coordinates given by the Poisson process $\{x_i, y_i, z_i\}$ are utilized for locating the discs. The dead leaves model is then generated as in Equation 3.1.

$$\bigcup_i (T_i + s_i) \tag{3.1}$$

Thus the set $T_i + s_i$ is an object at a time i . The variation in the z_i parameter causes the occlusion of the objects or leaves. Consider that the viewer is situated at the point $z_i = 0$, and that all the generated values of z_i are positive in the direction away from the viewer. The foremost leaf would be the one generated at the least z_i distance from the viewer. In the theoretical model, the discs that are generated later in time occlude the discs generated earlier. For the purpose of computer generation of the dead leaves model, we assume that the discs generated later in time are at a lesser distance, and thus occlude the earlier generated discs. The random size R

of the objects in the dead leaves chart follows the distribution as given in Equation 3.2.

$$P(R = r) \propto \frac{1}{r^\alpha} \quad (3.2)$$

where $\alpha > 1$.

The maximum and minimum possible values of the size are set in order to obtain non-trivial images. For a very low value of r_{min} the image is almost completely covered by microscopic objects, and for very large r_{max} , the image may completely consist of only one large object. Gray levels are assigned to the discs using a uniform distribution over the total range, with each point (x, y) being assigned the gray level of the visible disc at that point. Also, in order to have approximately full scale invariance, the exponent of the power law for the radii has to be -3 . For the texture preservation measurement using the dead leaves target, Cao et al. [2] suggest a contrast range for the gray levels of the target to be between $[0.25, 0.75]$. This low-contrast is used since it is the image condition in which the processing limitations in camera devices are most clearly observed. An example of the generated dead leaves target is shown in Figure 3.1. The statistics followed by the power spectral density of the dead leaves model have been discussed previously.

The other properties of the considered dead leaves target include:

1. Statistical rotation invariance. Since the objects in the model are discs, arbitrary rotations to the target do not change the measured power spectral density;
2. Statistical shift invariance. This property arises as a consequence of the Pois-

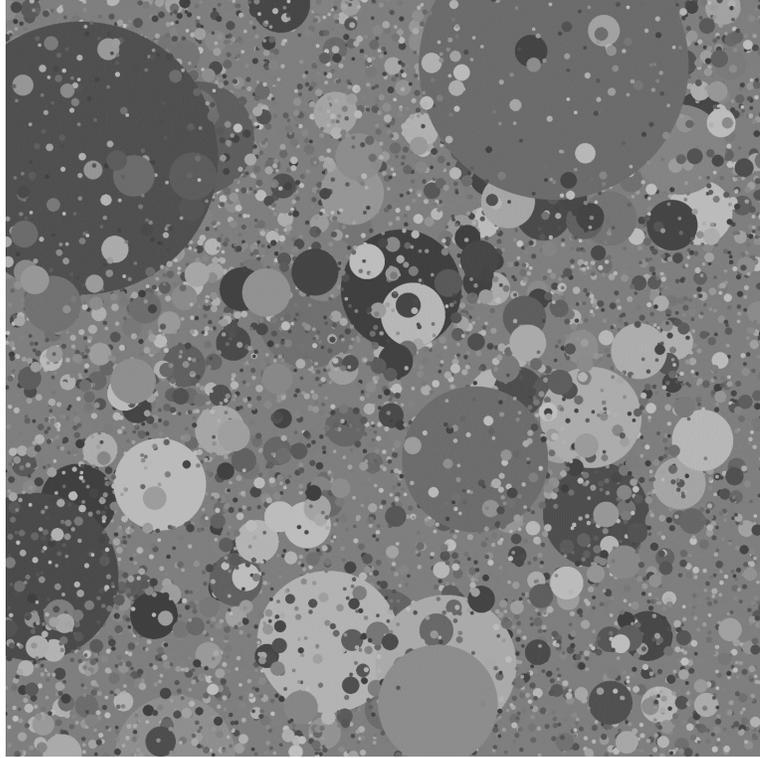


Figure 3.1: Generated dead leaves target.

son distribution used for locating the centers of the discs away from the plane origin. If the origin is shifted and the discs are generated again, the statistical properties of the completed target are maintained;

3. Statistical scale invariance. Ideal scale invariance would mean that the statistical properties of the target would not depend on the viewing distance. The dead leaves target is fractal in nature. We do however observe a variation in the modulation transfer function (MTF) with varying camera-target image capture distances.

Chapter 4: Artificial Sharpening

The two factors that determine the perceived sharpness of an image are the resolution and the acutance. Image acutance is a measure of perceived display sharpness, which can enable single-number comparisons between different camera systems, in place of plot comparisons using the texture modulation transfer function (MTF). It is a relatively recent measurement technique, introduced by the IEEE Camera Phone Image Quality (CPIQ) group [6]. The factors that determine the acutance include the system MTF, the contrast sensitivity function (CSF) of the human visual system, the viewing distance, and the image print height, according to Equations 4.1 & 4.2.

$$Acutance = \frac{\int_0^{\infty} MTF(v)CSF(v)dv}{\int_0^{\infty} CSF(v)dv} \quad (4.1)$$

where,

$$CSF(v) = \frac{av^c \exp(bv)}{K} \quad (4.2)$$

where, $a = 75$, $b = 0.2$, $c = 0.8$, $K = 34.05$, and v is angular frequency (*cycles/degree*).

The relation between spatial frequency in cycles/pixel and angular frequency in *cycles/degree* is given by Equation 4.3.

$$v\left(\frac{cycles}{degree}\right) = f\left(\frac{cycles}{pixel}\right)\left(\frac{\pi}{180}\right)\left(\frac{n_{ph}}{PH}\right)(d) \quad (4.3)$$

Where, n_{ph} is the number of vertical pixels along the picture height, d is the viewing distance, and PH is the picture height in the same units as the viewing distance.

From Equations 4.1 to 4.3 it is clear that the acutance measures the perceived sharpness as a function of print height and viewing distance. By maintaining standard experimental values for these variables, different camera systems can be compared on the basis of image quality. This chapter evaluates the effect of sharpening on the measured acutance value.

4.1 The CSF

The contrast sensitivity describes the ability of the visual system to distinguish bright and dark components or areas of an image. Campbell et al. provided one of the first charts displaying the sinusoidal grating pattern that can be used to determine the contrast thresholds required to view the pattern at different spatial frequencies (Figure 4.1). In the chart, the luminance of the pixels is modulated sinusoidally, with logarithmically increasing frequency, along the horizontal direction. The contrast also varies logarithmically from the bottom to the top of the image. In general, the sensitivity of the visual system at a particular spatial frequency is the contrast threshold required to view the pattern at that frequency. Hence, by measuring the contrast thresholds at different spatial frequencies, the contrast sensitivity function is obtained. The area below the obtained CSF curve determines the combinations of contrast and spatial frequency at which the pattern is visible. The CSF of the human visual system can be modelled as in Equation 4.2. The results

are shown in Figure 4.2.

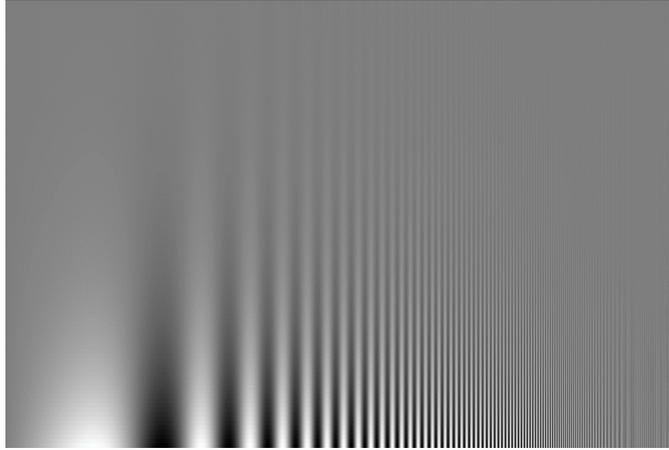


Figure 4.1: Campbell-Robson CSF chart.

The human visual system is less sensitive to low spatial frequencies, than it is to intermediate spatial frequencies (i.e. 2 cycles/degree to 5 cycles/degree) [7]. Thus maximum sensitivity is attained for a range of intermediate frequencies, with higher contrast thresholds required for gratings having spatial frequencies outside this range. For the estimation of the CSF from the chart, the illumination level also has a role on the contrast sensitivity, with low illumination conditions causing a drop in visual sensitivity at primarily high spatial frequencies. The low spatial frequencies are not affected the same at such low illumination levels. In case of modelling the CSF from the equation, the assumptions of the model include uniform sufficient illumination over the chart.

As given in Equation 4.1, the acutance weights the SFR or MTF value with the corresponding value of the contrast sensitivity, and integrates it over all possible frequencies (normalized by the denominator), to define a single number for the

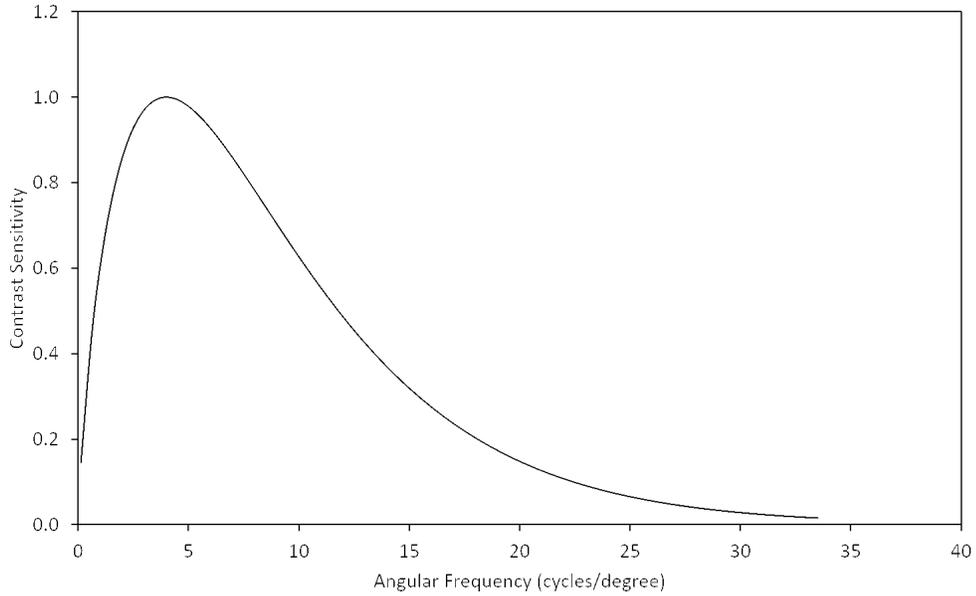


Figure 4.2: The Contrast Sensitivity Function.

camera system in question. The CPIQ group defines an acutance value of 0.8851 beyond which there is no perceptible improvement in image quality.

In the context of image sharpness, the acutance measures the transitions in the image from bright to dim regions, or the overall contrast in the image edges. The resolution determines the cameras ability to distinguish between closely spaced detail.

4.2 Unsharp Masking

Sharpening a digital image by increasing its acutance effectively implies increasing the contrast along the edges in the image - a form of contrast enhancement. A widespread technique for performing this is by the Unsharp Masking method [8].

A sharpened image is created using a blurred version of the original image, as given in Equation 4.4 & 4.5.

$$f_{edge}(x, y) = f(x, y) - f_{smooth}(x, y) \quad (4.4)$$

where $f(x, y)$ is the original image, $f_{edge}(x, y)$ contains the edge information or high frequency detail of the original image, and $f_{smooth}(x, y)$ is obtained by blurring the original image (using a Gaussian filter of a specified kernel size, for example).

$$f_{sharp}(x, y) = f(x, y) + k f_{edge}(x, y) \quad (4.5)$$

Subtracting the low-frequency information from the original leaves the edge image which contains the high frequency components. Adding a fraction k of this edge image to the original image thus amplifies the high frequency components, and results in a sharpened image. A Gaussian low-pass filter can generally be used for the purpose of obtaining the blurred image. A threshold value also exists which specifies how much intensity difference is required between a pixel in the original image and the blurred image, before sharpening is applied.

The *amount* parameter in Unsharp Masking determines the strength of the sharpening effect, and is related to the fraction of the edge image that is added to the original image (the parameter k in Equation 4.5). Figures 4.3(a) to 4.3(i) shows the effect of varying the *amount* parameter, for different illumination intensity levels. The effect of sharpening is clearly visible in the texture region.

The radius parameter determines the distance out from the edge on the dark and light sides, that are affected by sharpening. It is related to the kernel size of the Gaussian filter used for obtaining $f_{smooth}(x, y)$ via blurring. Using a high radius of

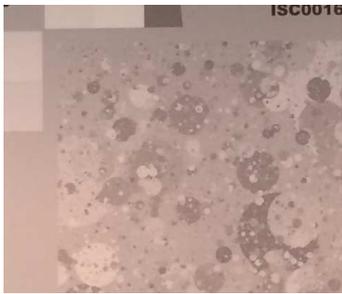
sharpening can lead to halo effects observed on the edges of images. Figures 4.4(a) to 4.4(f) show sample images showing variation when both the radius and amount parameters are varied, at different illumination levels.

4.3 Results

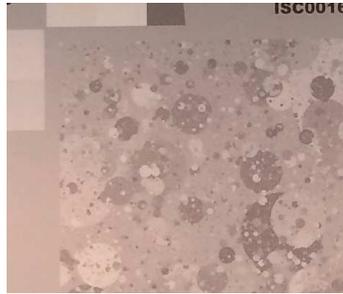
The image processing occurring in mobile camera phones includes an amount of denoising and sharpening, which would increase the visual quality of the image as a whole. The sharpening may also add an amount of artificial texture information to the image which would not have been initially present in the scene. Since acutance is a perceptive measure of sharpness, artificial sharpening may be used to increase the edge contrast differences and hence the acutance. This makes an objective measure of sharpness all the more important. In this section we obtain the texture MTF results of the dead leaves target image at various illumination levels, that have been sharpened to different extents.

Figure 4.5 and Figure 4.6 show the results for the texture MTF calculation for artificially sharpened images. The texture MTF increases with illumination level and stabilizes at around 1250 lx illumination intensity. This would indicate the minimum level of illumination required for accurate readings. Too low illumination levels lead to underexposed images which lack important texture structure and information. The interesting observation is in the case of the acutance values for the sharpened images. Sharpened images result in higher acutance for all the illumination intensity levels, with a correlation between the extent of sharpness and the

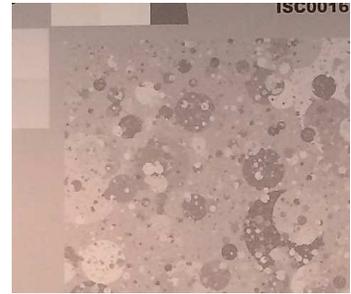
observed acutance. Increasing both the radius and the amount of sharpness lead to non-linear increases in the calculated texture MTFs. As was discussed earlier, high levels of sharpening in the texture images lead to a perceived improvement in image quality. The dead leaves algorithm as such is unable to distinguish between the natural texture information contained in an unaltered image, and artificially enhanced edges and other texture components, which increase the amount of high-frequency information in the image.



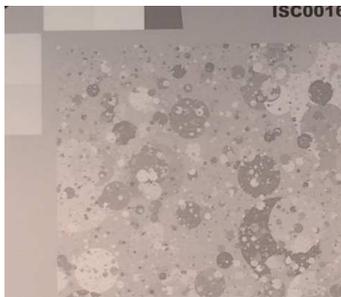
(a) Normal unsharpened image at 150 lx illumination intensity



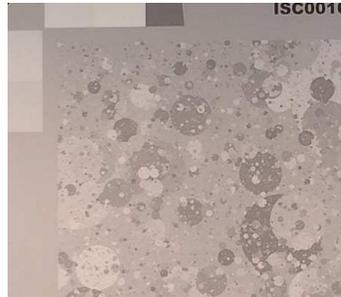
(b) Sharpened image at 150 lx illumination intensity, radius=1, amount=0.8



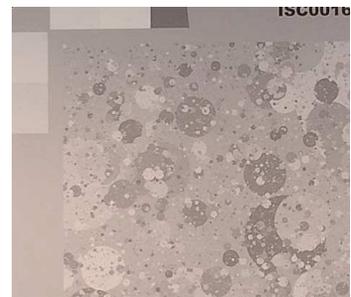
(c) Sharpened image at 150 lx illumination intensity, radius=1, amount=1.5



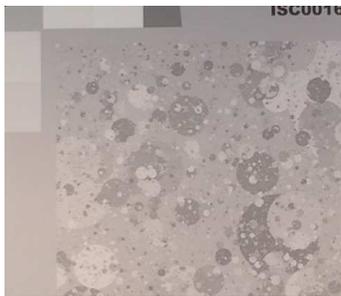
(d) Normal image at 1250 lx illumination intensity



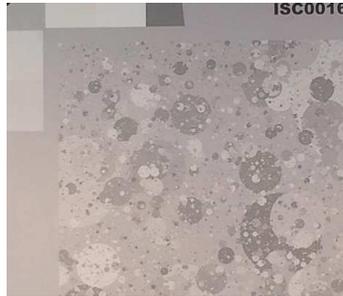
(e) Sharpened image at 1250 lx illumination intensity, radius=1, amount=0.8



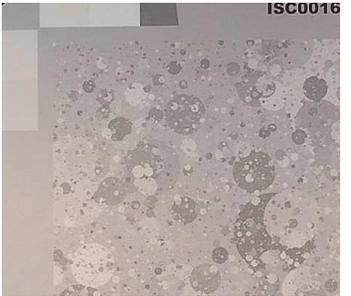
(f) Sharpened image at 1250 lx illumination intensity, radius=1, amount=1.5



(g) Normal image at 3000 lx illumination intensity



(h) Sharpened image at 3000 lx illumination intensity, radius=1, amount=0.8



(i) Sharpened image at 3000 lx illumination intensity, radius=1, amount=1.5

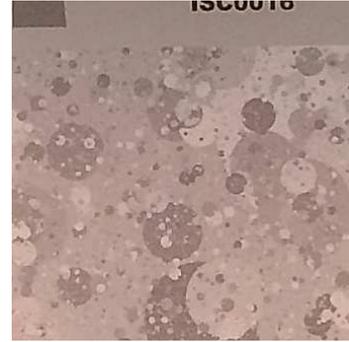
Figure 4.3: Effect of varying the 'amount' factor, at various illumination intensity levels



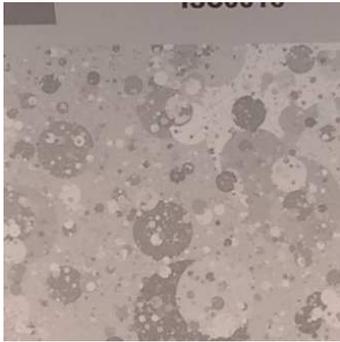
(a) Normal image at 150 lx illumination intensity



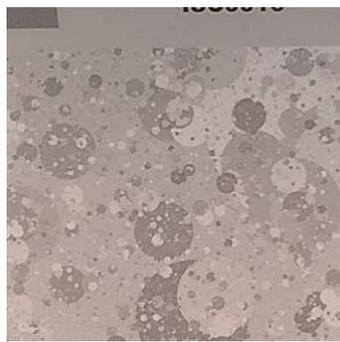
(b) Sharpened image at 150 lx illumination intensity, radius=1.5, amount=0.8



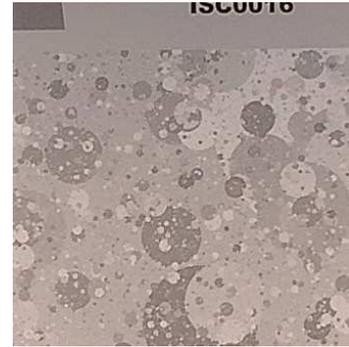
(c) Sharpened image at 150 lx illumination intensity, radius=1.5, amount=1.5



(d) Normal image at 1250 lx illumination intensity



(e) Sharpened image at 1250 lx illumination intensity, radius=1.5, amount=0.8



(f) Sharpened image at 1250 lx illumination intensity, radius=1.5, amount=1.5

Figure 4.4: Effect of varying both the 'amount' and 'radius' factors, at various illumination intensity levels

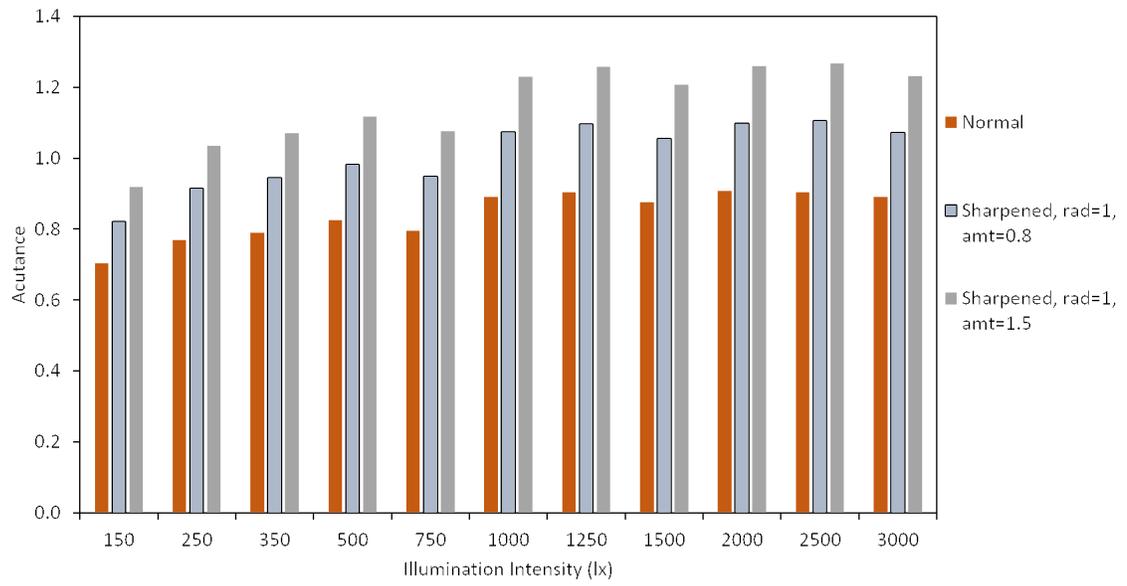


Figure 4.5: Results for texture MTF calculation for images sharpened varying the amount parameter, at various illumination levels.

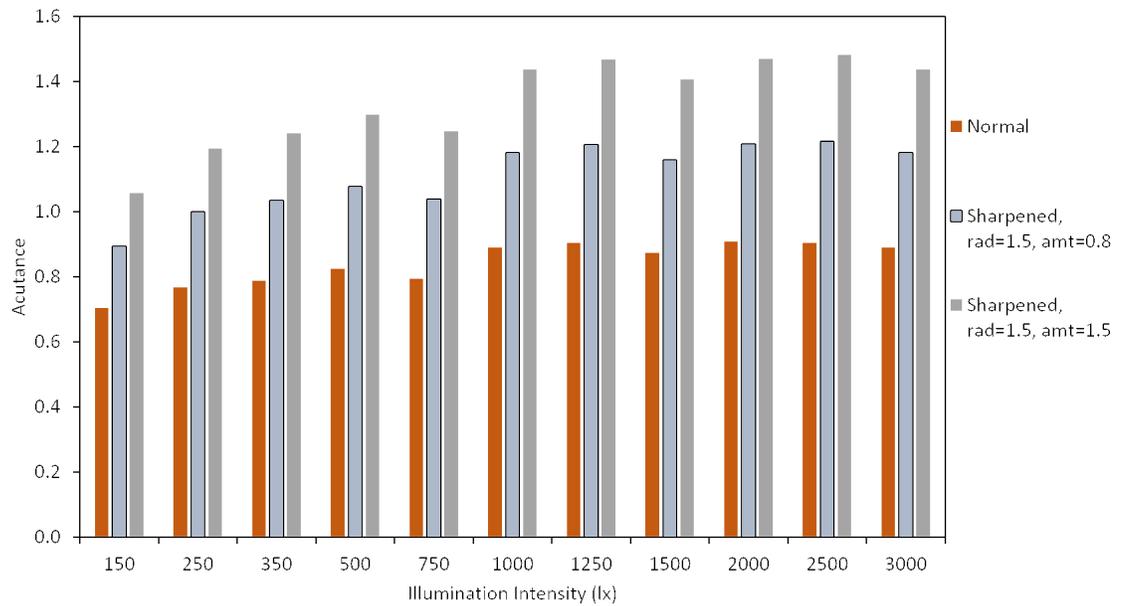


Figure 4.6: Results for texture MTF calculation for images sharpened varying the radius and amount parameters, at various illumination levels.

Chapter 5: Distance from Camera to Target

Setting up the target at a proper distance from the imaging system is important in order to limit the size of the target in the cameras field of view (FOV). This step is taken in order to limit the effect of light and color shading, which can be due to slightly non-uniform illumination of the target, or due to particular characteristics of the lens or sensor [1]. By limiting the target size over the field of view, geometric distortion effects across the camera lens are also avoided. Distortion effects are observed when straight lines in the object space are rendered instead as curved lines on the sensor, leading to the name curvilinear distortion [9]. The two main types of distortion commonly occurring in camera systems are barrel and pincushion distortion, as observed in Figures 5.1(a) and 5.1(b). Distortion effects mainly appear at the outer extremity of the scene, and have negligible effects at the center of the lens FOV.

Most camera systems have aperture stops specifically designed to reduce spherical aberrations or astigmatism, which can be a cause of distortion [10] [11].

The position of the obtained image from the object follows the laws of optics, and forms at the intersection of the three principal rays, which are in turn affected by the location of the stop. For the location of the stop at the lens, there is no

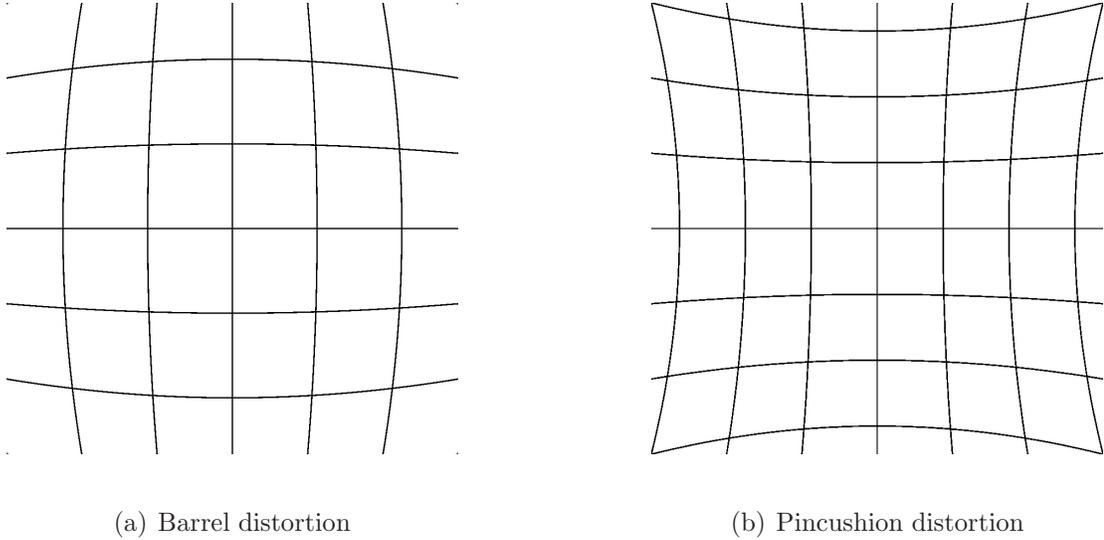


Figure 5.1: Types of distortion

distortion, and the system is called orthoscopic. Barrel distortion occurs when the image magnification $\frac{h}{y}$ is smaller than in the orthoscopic case, with the value of the magnification progressively decreasing toward the image edges. On the other hand, pincushion distortion occurs when the magnification is greater than the orthoscopic case, and increases towards the image edges. Since the center of the image is relatively unaffected by such distortion effects, the dead leaves target is limited to lying in this region in the field of view. Several images of the target at varying distances are obtained using several camera systems, in order to understand the effect of distance on the calculated texture MTF.

5.1 Results

The Canon digital single-lens reflex (DSLR) camera is included in this experiment along with two other mobile cameras - an iOS system and an Android system. The spatial frequency readings are converted from *cycles/pixel* to *cycles/mm* for

Table 5.1: % coverage of HFOV at different distances, Canon DSLR

| Distance (cm) | HFOV (%) |
|---------------|----------|
| 60 | 60 |
| 66 | 52 |
| 72 | 47 |
| 78 | 43 |
| 85 | 39 |

all the results, in order to generate accurate representation of the results, since the same cropped target image occupies a different number of pixels at different distances. Table 5.1 gives the percentage coverage of horizontal field of view (HFOV) by the dead leaves target, at different distances.

Figure 5.2 graphs the results of texture MTF obtained for the DSLR camera. There is an observed difference in the observed MTF readings at low spatial frequencies and high spatial frequencies. At low spatial frequencies, the target occupying the smallest field of view is observed to have the highest texture MTF. At higher spatial frequencies, beyond approximately 20 cy/mm, the texture MTF is observed to decrease in inverse proportion to the distance from the target. A reason for this may be the limiting resolution of the camera, which may be unable to distinguish fine texture details in the image, as the distance from the target increases. At larger camera-target distances, the absence of distortion effects result in an accurate MTF reading. The obtained acutance values are observed in Figure 5.3. Due to the effect

of the contrast sensitivity function, the effect of the dip at higher frequencies is overshadowed by the improved MTF at the lower spatial frequencies. Sensor noise can also be a factor in the lower MTF reading at the higher spatial frequencies.

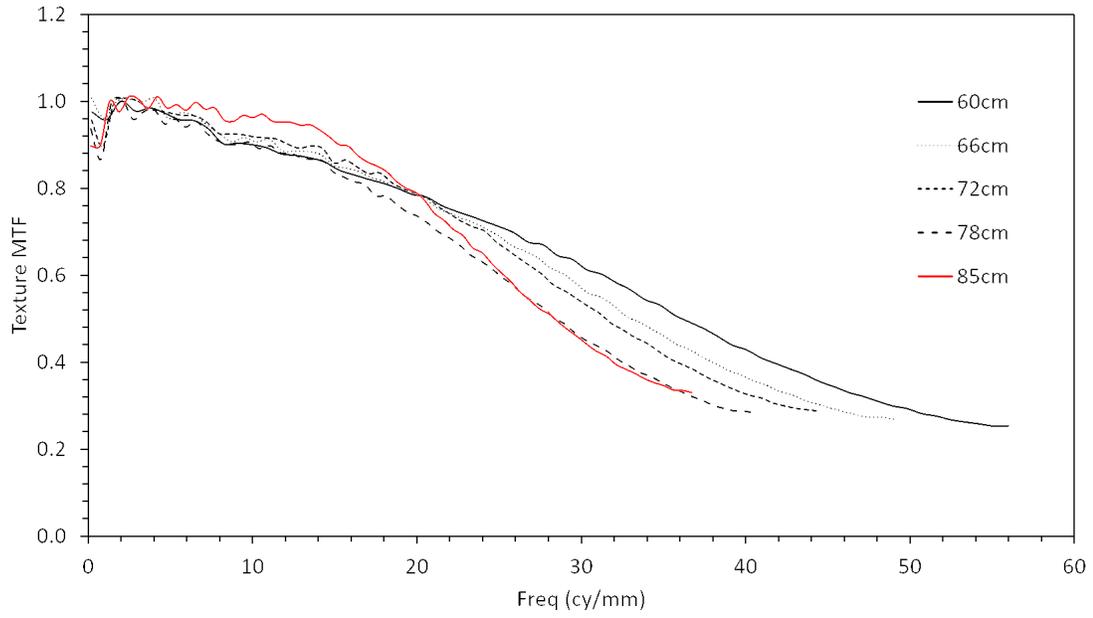


Figure 5.2: Variation in texture MTF with distance for the Canon DSLR camera.

The mobile cameras considered are the iPhone 5s and the Nexus 5, running the iOS operating system and the Android operating system respectively. Table 5.2 and Table 5.3 give the percentage coverage of horizontal field of view (HFOV) by the dead leaves target, at different distances, for the iPhone 5s and the Nexus 5.

Figure 5.4 and Figure 5.5 display the texture MTF and the acutance values obtained for the iPhone 5s, for images taken at various distances. A similar trend to the DSLR camera is observed, with the difference that the threshold point is at a much higher spatial frequency. Also the difference in the MTF at different distances is not as high as seen in the case of the DSLR. In the case of the Nexus 5, at the

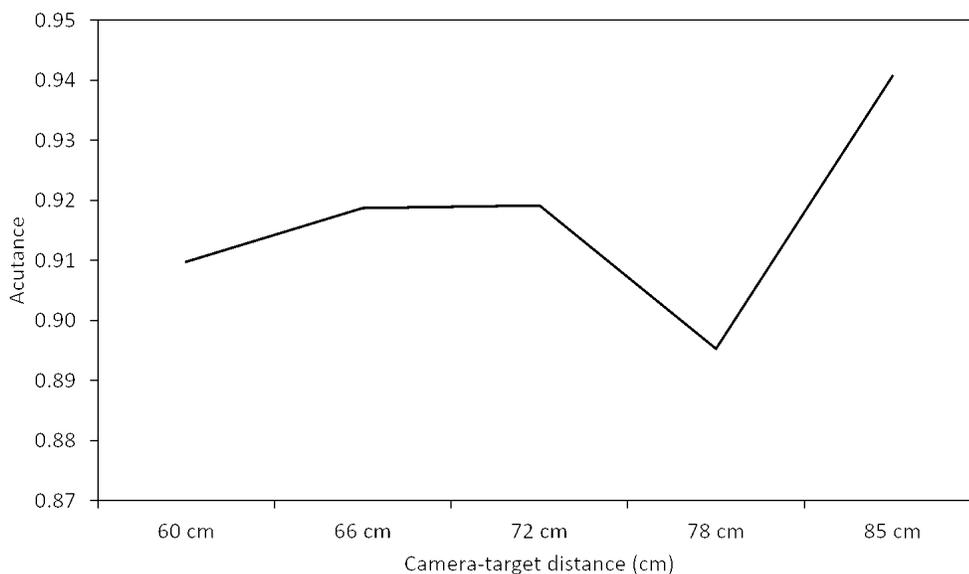


Figure 5.3: Variation of acutance with distance of camera to target - Canon DSLR

lower spatial frequencies, the obtained texture MTF is almost constant, with the variations at higher spatial frequencies being due to the limiting resolution of the device, as observed in Figure 5.6 and Figure 5.7.

5.2 Conclusion

This chapter examined the effect of varying the camera-target distance on the obtained texture MTF values for different camera systems. A trend of obtaining initially higher and progressively decreasing texture MTF was identified with increasing distances. This lower MTF values obtained at the higher spatial frequencies may be due to the effects of noise at larger distances, limiting camera resolution, and non-uniformity in illumination.

Table 5.2: % Coverage of HFOV at different distances, iPhone 5s

| Distance (cm) | HFOV (%) |
|---------------|----------|
| 55 | 43 |
| 60 | 40 |
| 65 | 36 |
| 70 | 34 |
| 75 | 31 |

Table 5.3: % Coverage of HFOV at different distances, Nexus 5

| Distance (cm) | HFOV (%) |
|---------------|----------|
| 55 | 41 |
| 60 | 38 |
| 65 | 36 |
| 70 | 33 |
| 75 | 30 |

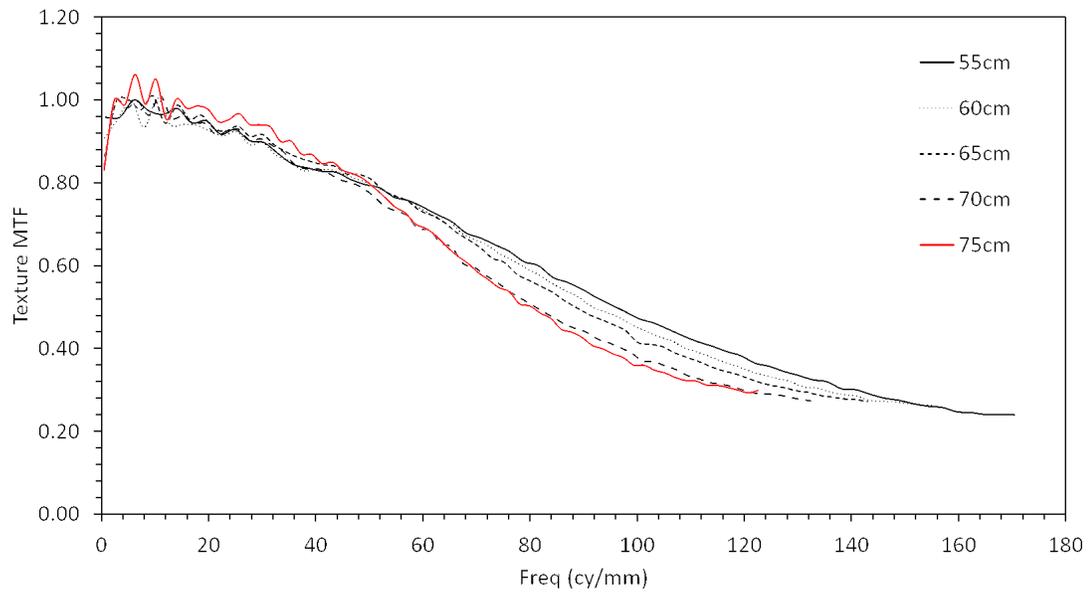


Figure 5.4: Variation in texture MTF with distance for the iPhone 5s.

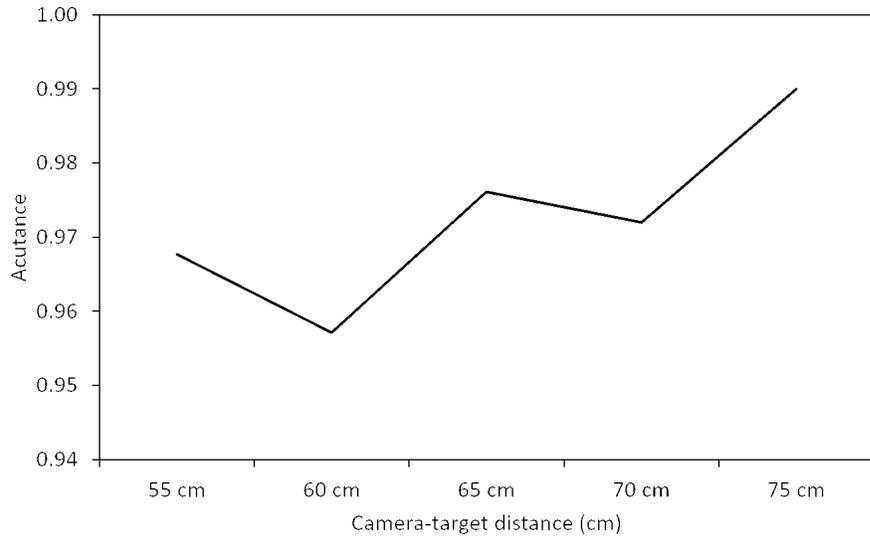


Figure 5.5: Variation of acutance with distance of camera to target, iPhone 5s.

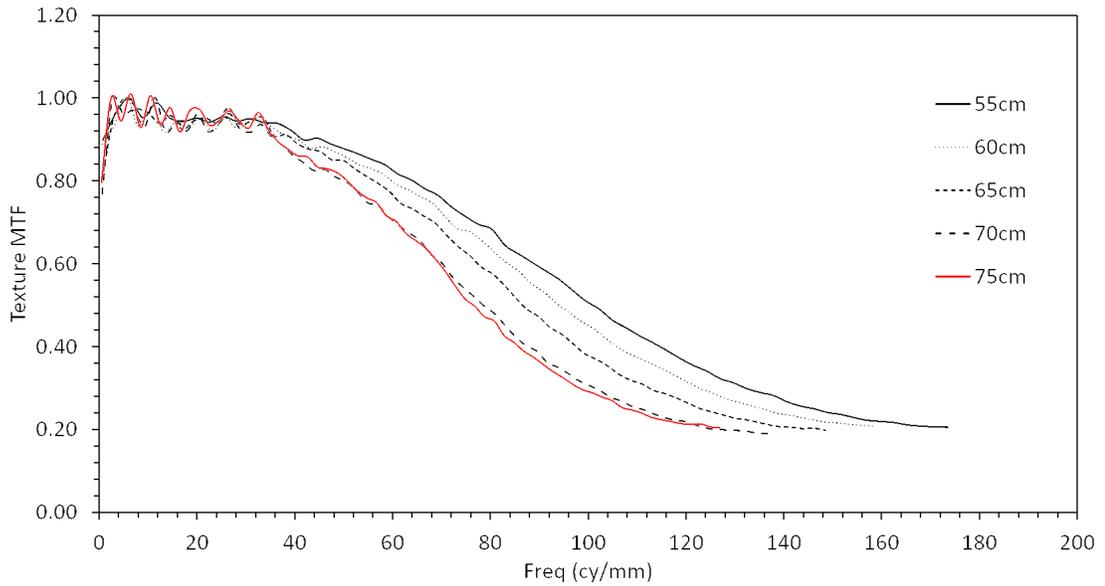


Figure 5.6: Variation in texture MTF with distance for the Nexus 5.

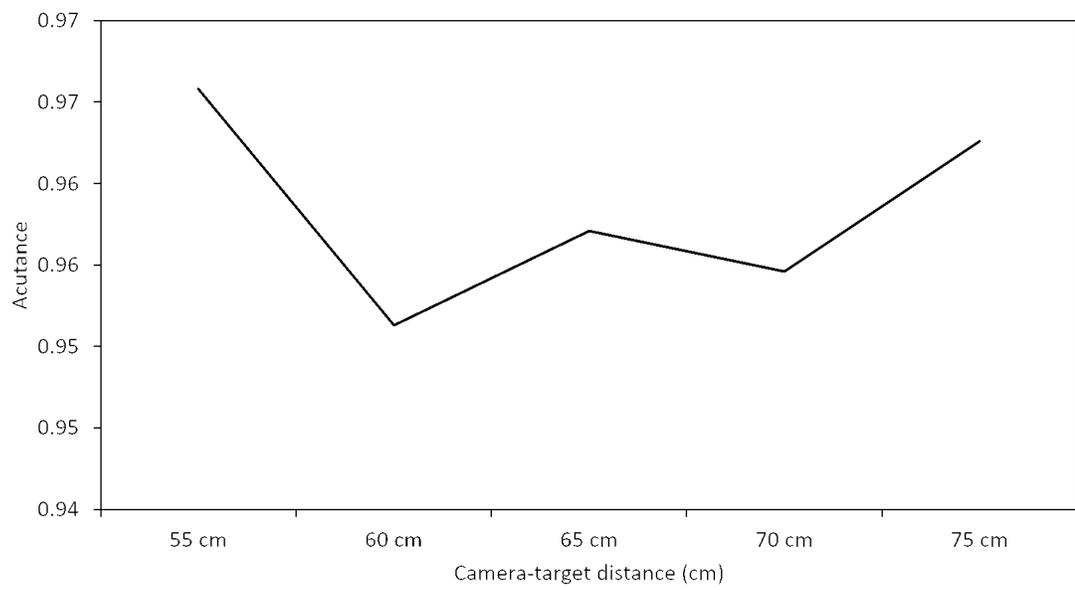


Figure 5.7: Variation of acutance with distance of camera to target, Nexus 5.

Chapter 6: Illumination Variation

The illumination of the target plays an important role in the accuracy of the observed MTF results. The primary requirement is of uniform illumination covering the area of the chart which is utilized for the modulation calculation. The presence of non-uniform illumination can be a cause of shading, which gives erroneous light intensity readings in parts of the image [1]. In this chapter we take a look at how the experimental setup ensures illumination uniformity, followed by the results showing the effect of using different levels of illumination for the dead leaves texture MTF calculation.

Illumination uniformity is ensured by the use of diffusers in front of the light sources, which diffuses or scatters the light into a wider angle. The target is then divided into 12 grids and the light intensity in lux is measured in each grid, using a handheld illuminance meter. The illuminance uniformity over the target is then measured using Equation 6.1.

$$\text{Illuminance Uniformity} = \frac{I_{min}}{I_{avg}} \quad (6.1)$$

where I_{min} is the minimum measured illumination level over the grid, and I_{avg} is the area weighted-average measured illumination level over all the grids. Illumination uniformity $\sim > 0.9$ is required to be set for accurate MTF measurements.

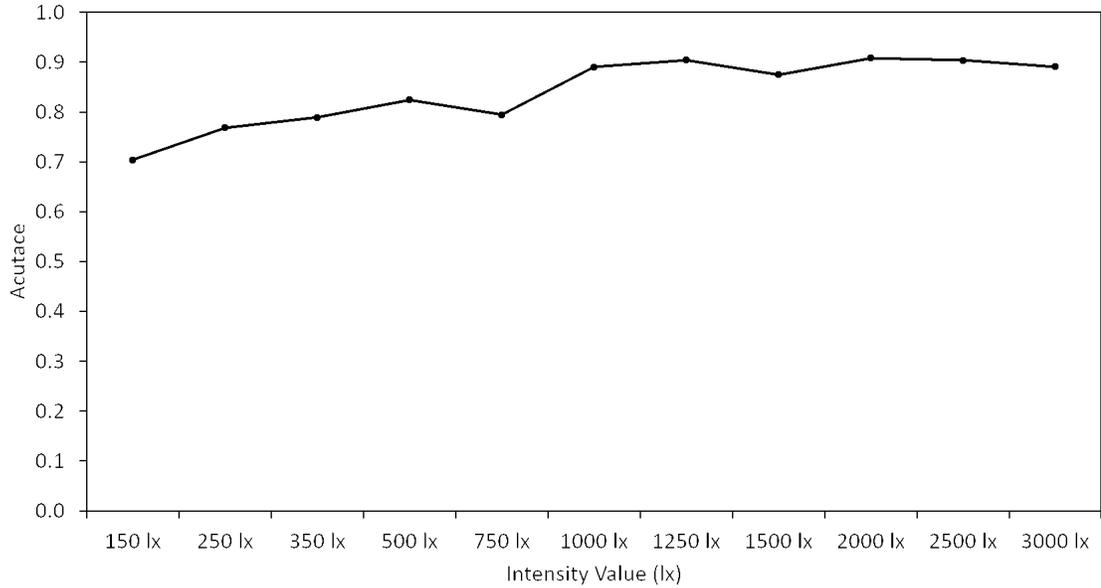


Figure 6.1: Variation of texture acutance with illumination levels for the iPhone 5s, using the Eiko bulbs.

For the experiment, we vary the illumination levels and the light sources used, and identify the texture MTF and acutance values for two camera phones. The spectral power distribution of the light sources used are also obtained to identify if there is any correlation with the observed texture MTF.

Figure 6.1 shows the acutance values obtained for the iPhone 5s, at various illumination levels. The default camera app of the phone is used for capturing the images. The illuminating source used is the Eiko Supreme Photoflood 500W bulb. The texture acutance is low at low illumination levels below 1000 lux. This is due to the light intensity being insufficient to accurately capture fine details in the scene. At ~ 1250 lux, the texture acutance stabilizes and plateaus for further higher illumination intensities.

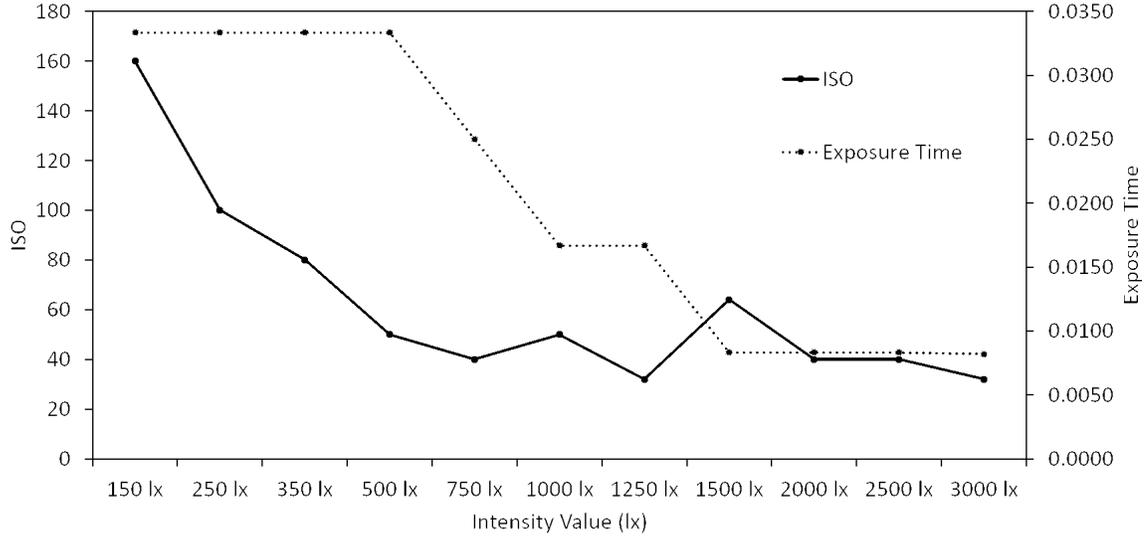
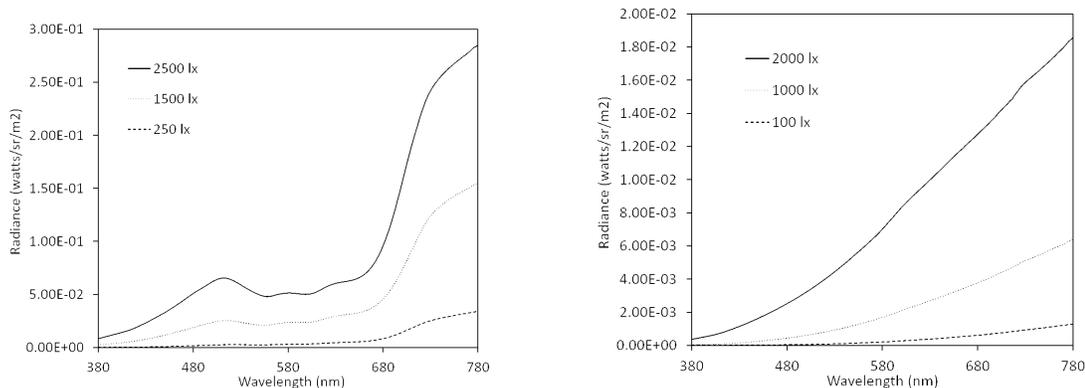


Figure 6.2: Exposure time and ISO sensitivity values for the images obtained at the different intensity levels, for iPhone 5s automatic metering mode, using Eiko bulbs.

Figure 6.2 displays the exposure time values and the ISO sensitivity numbers for the images taken at the different light intensity levels. These are the settings that are automatically metered by the camera phone at the time of capturing the image. A correlation between the two is observed, for intensity levels above 1000 lux. For lower light intensity values the exposure time reading is less sensitive and remains at a high value, relative to the ISO number, which varies sharply with increasing intensity. This points to the fact that the variation in ISO number is preferable to the variation in exposure time, which can be due to the increased sensor sensitivity. One of the main factors affecting the sensor sensitivity is change in the ISO number.

The second light bulb considered is the Philipps Halogena BR-30 floodlight.



(a) PSD for the Eiko Supreme lightbulb

(b) PSD for the Philipps Halogena lightbulb

Figure 6.3: Power Spectral Densities for the light sources utilized

The two light sources are used for their effectiveness in providing true-color uniform bright light for coverage of the target. Figures 6.3(a) and 6.3(b) plot the Power Spectral Distribution (PSD) for the two light sources utilized. Figure 6.4 gives a comparison of the texture acutance obtained for the two different light sources utilized. A similar curve is obtained in both cases, with the acutance values plateauing at above 1100 lux.

Figure 6.5 shows the ISO numbers and exposure time values obtained for the images taken using the iPhone 5s with the Philipps Halogena lightbulb. In this case, the exposure time remains stable for a different range of light intensities as compared to the Eiko bulb. At the higher intensities, the exposure time and ISO numbers both decrease as observed in the case of the Eiko bulb.

Figure 6.6 plots the texture acutance for the Nexus 5 phone, obtained at different illumination intensities. This plot correlates with the earlier observation that the texture acutance increases with increasing illumination intensity up until

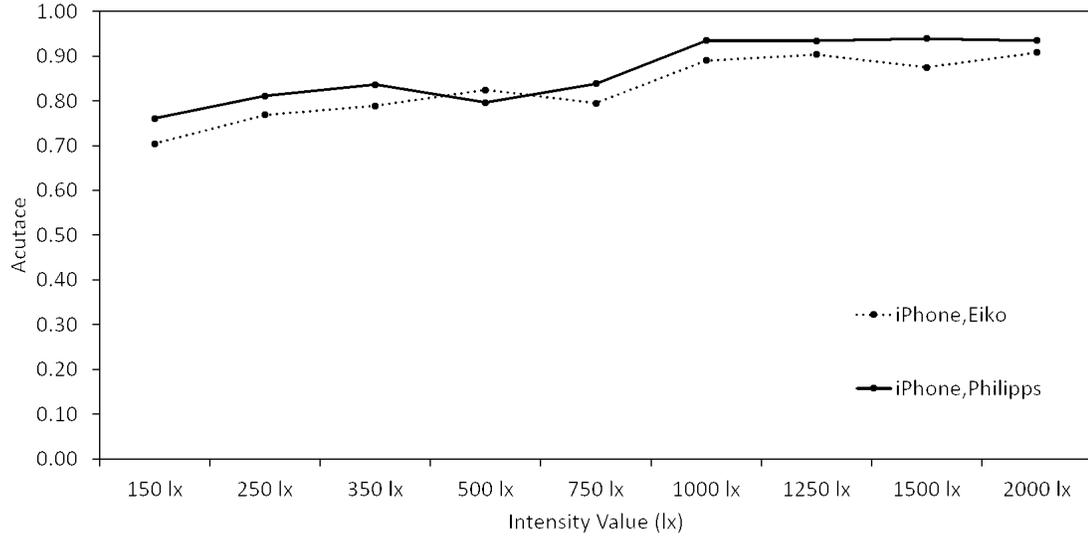


Figure 6.4: Comparison of the texture acutance obtained for the iPhone 5s for two light sources, the Eiko Supreme and the Philipps Halogena, at different light intensities.

1000 lx after which it stabilizes and remains constant. Figure 6.7 displays the overall values of texture acutance for the different camera phones and light sources utilized. While the Nexus 5 gives lower texture acutance values as compared to the iPhone 5s, the illumination intensity level after which the acutance values stabilize are approximately the same. There is also no observable correlation between the observed MTF using the different light sources, and their power spectral densities.

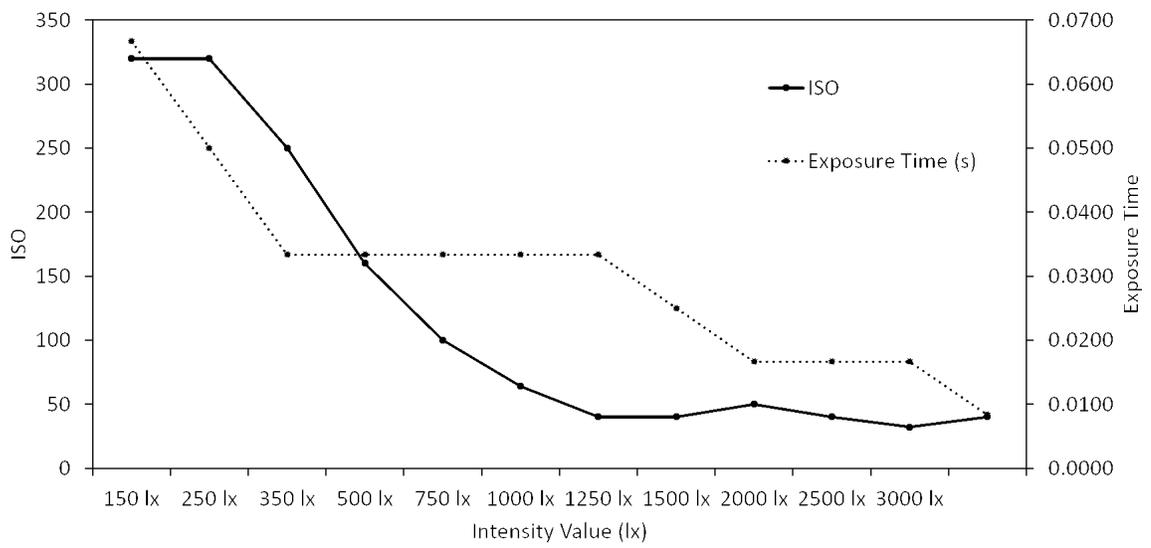


Figure 6.5: Exposure time and ISO sensitivity values for the images obtained at the different intensity levels, for the iPhone 5s, using the Philipps Halogena bulbs.

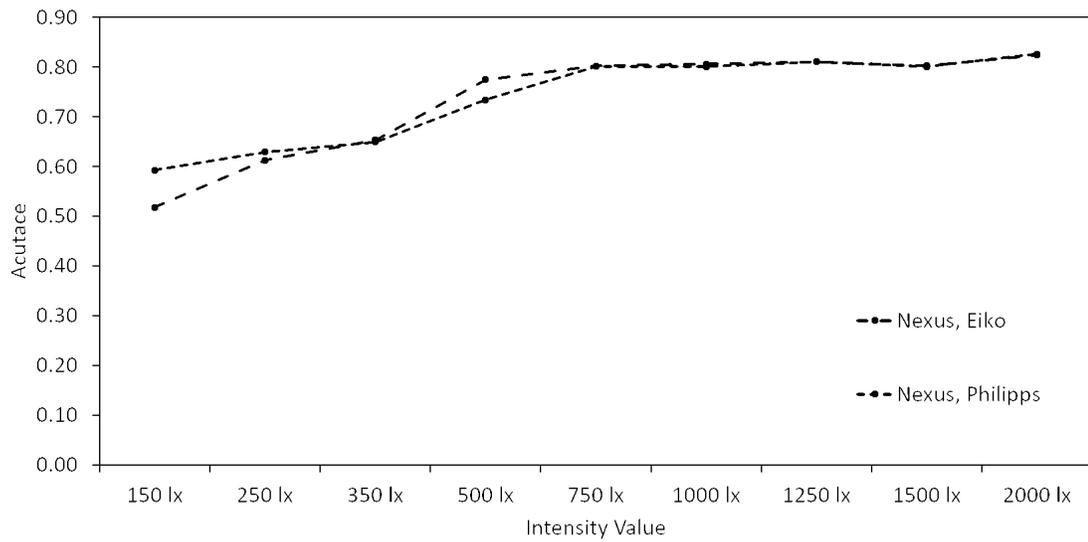


Figure 6.6: Comparison of the texture acutance obtained for the Nexus 5 for two light sources, the Eiko Supreme and the Philipps Halogena, at different light intensities.

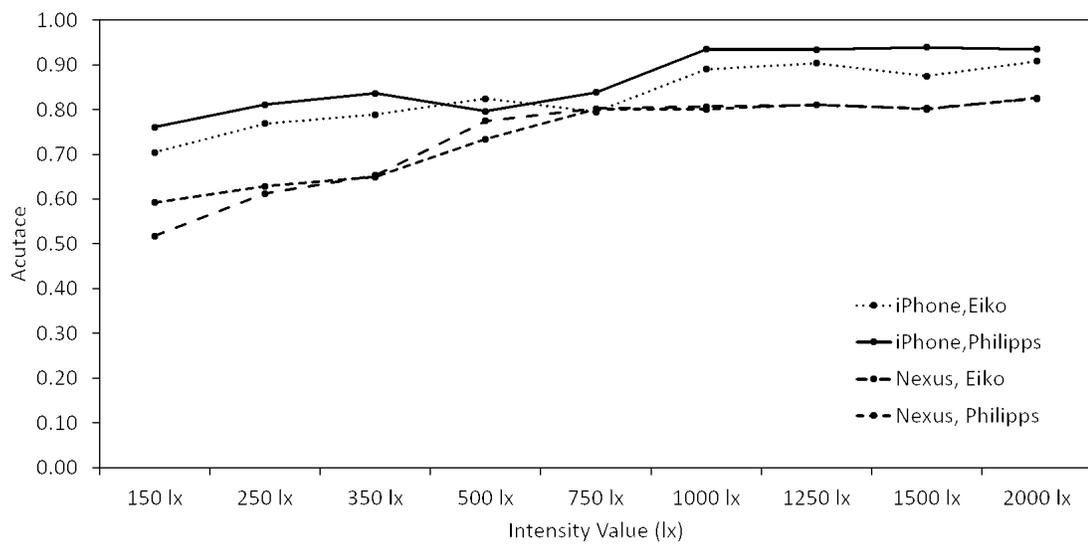


Figure 6.7: Overall texture acutance for the iPhone 5s and the Nexus 5.

Chapter 7: Exposure Time

The exposure time/shutter speed setting of the camera determines the amount of light which the digital sensor is exposed to, at the time of obtaining the picture. The amount of light reaching the sensor is directly proportional to the exposure time. This setting is important mainly in the capture of moving subjects, where the shutter speed has to be high enough to freeze the motion in the particular frame. The image obtained can be under-exposed or over-exposed, in case there is less light or more light than required, respectively. The values for the shutter speed are generally specified in terms of fractions of a second, such as $1/15$ s, $1/30$ s, $1/60$ s etc.

With reference to the Dead Leaves target, the parameters of ISO sensitivity and exposure time are more relevant in the aspect of reliable capture of details, since the target represents a stationary scene in nature. A sufficiently high exposure time has to be present in order to allow sufficient light on the sensor, so as to capture the texture details faithfully. The two parameters of exposure time and ISO number are interlinked and the study of any one of the parameters involves strictly specifying the level maintained for the other. For the experiment, in order to vary the parameters in the camera, we use the VSCO Cam app on the iOS device.

In the experiment on variation of texture MTF with exposure time, the ISO number setting of the camera is held at a constant value, and the exposure time is varied. We isolate the two parameters while performing the experiment. The ISO number determines the sensitivity to light of the digital sensor of the camera. A high ISO number is usually required when obtaining images in low-light conditions. Another aspect of changing the ISO sensitivity is in the noise level obtained in the final image. A high ISO number increases the sensitivity of the sensor to such an extent that the noise level in the final image also correspondingly increases. Further details are presented in the chapter regarding ISO sensitivity.

7.1 ISO32

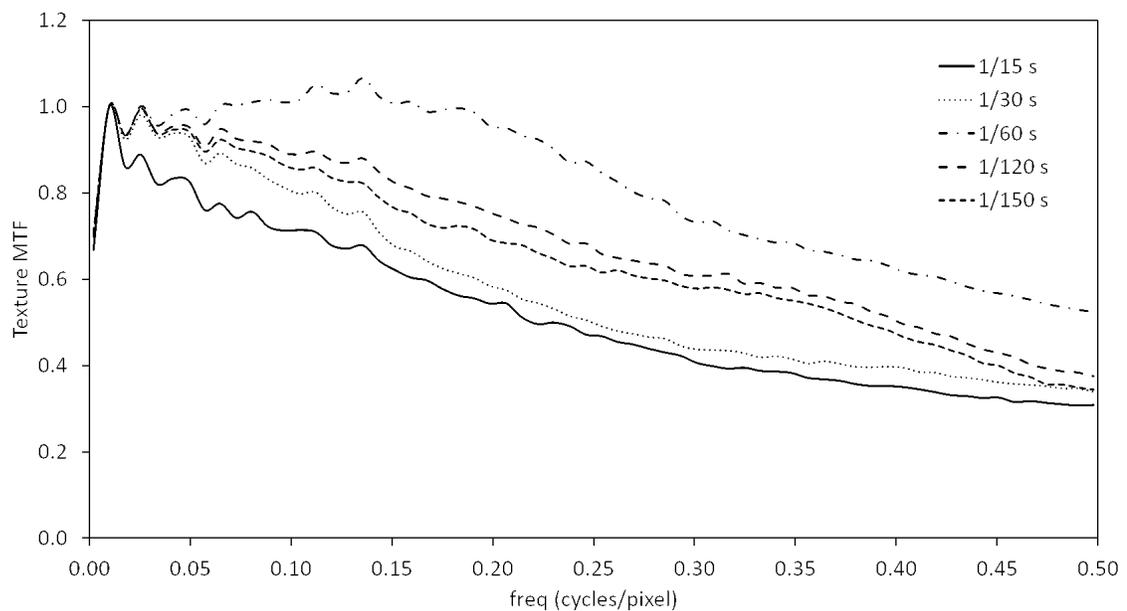


Figure 7.1: Variation in texture MTF with exposure time, at ISO32.

The ISO sensitivity is held constant at ISO32, and images are obtained at different exposure time settings. The illumination level is set at 1200 lx.

Figure 7.1 shows the texture MTF plots for the different images. The observation is that the calculated texture MTF attains a maximum value for a particular exposure time.

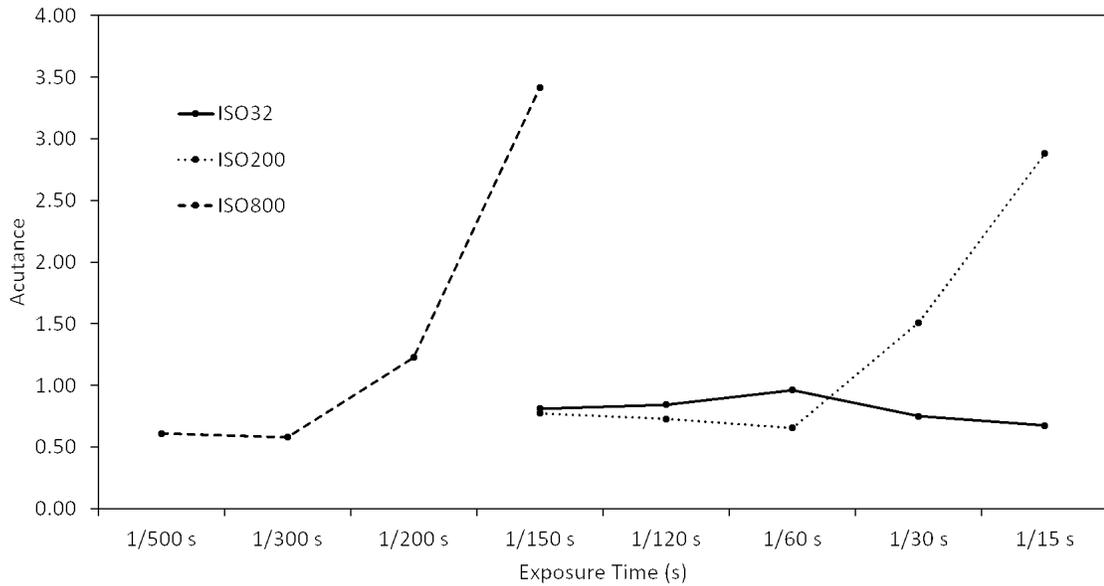


Figure 7.2: Acutance values for different exposure times, at ISO32, ISO200 and ISO800.

With reference to Figure 7.2, we observe the acutance values at ISO32 peaking at the exposure time of 1/60 s. The images do not show any saturation, which may be due to the relatively low ISO speed at which the images are obtained.

7.2 ISO200

The same experimental conditions are utilized as earlier, and the ISO sensitivity of the camera is set to ISO200.

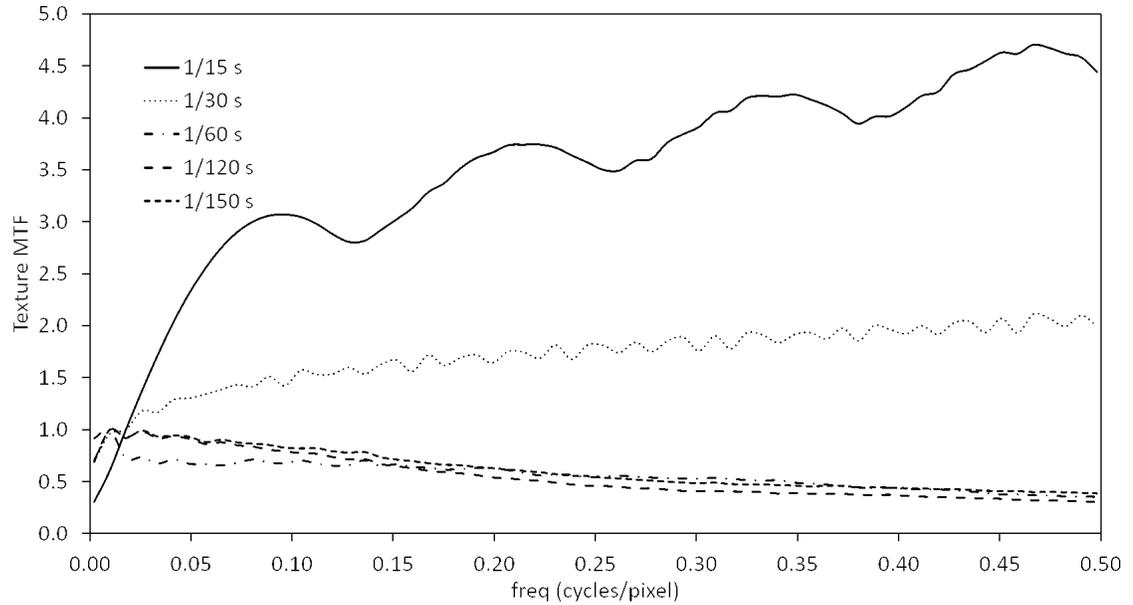


Figure 7.3: Variation in texture MTF with exposure time, at ISO200.

Figure 7.3 shows the variation at ISO200. The same exposure times are tested as in the case of ISO32. Using a higher ISO number, and the longer exposure time of 1/15 s gives a highly saturated and over-exposed image. As such, the image cannot be used for the purpose of texture MTF calculation, and gives an abnormal MTF reading. For the higher value of ISO number, a shorter exposure time of 1/150 s gives an unsaturated image.

7.3 ISO800

The ISO number is set to the higher setting of ISO800. The exposure times are different from the previous experiments, so as to permit images with proper exposure. Even using the adjusted smaller exposure times leads to over-saturated images as seen in the case of 1/150 s and 1/200 s.

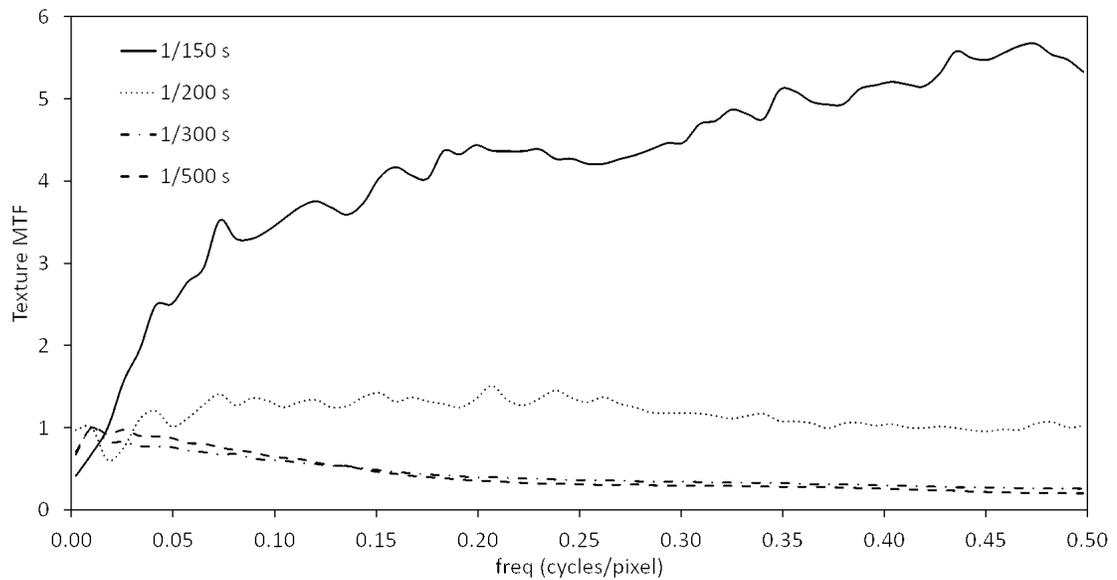


Figure 7.4: Variation of texture MTF with exposure time, at ISO800.

With reference to Figure 7.4, overall a similar trend is seen as in the case of ISO200, and is slightly different from the trend observed in the case of the lower ISO number, ISO32. Figure 7.2 contains the plots of the acutance values obtained at ISO800.

Chapter 8: ISO Speed

The camera's ISO setting determines the sensitivity of the sensor to the light that reaches it. A higher ISO number implies higher sensitivity to light. Along with aperture size, and exposure time, it is one of the factors that determines the image quality and exposure of the photograph. For getting the same exposure, doubling the ISO number would require the exposure time to be halved - this is usually the condition observed in automatic exposure metering in cameras. A side-effect of this higher sensitivity of the sensor is image noise, which becomes clearly visible at higher ISO settings.

In the context of the dead leaves target, a high ISO number can cause some of the fine texture detail present in the target to be lost. It becomes unclear as to whether the graininess present in the final image is due to noise, or due to the occluded circles with varied radii. As in the chapter on exposure time, the images with variation in ISO number are obtained while keeping the exposure time constant.

8.1 Exposure Time Setting 1

The exposure time is set to $1/50$ s, and images are obtained from the iOS device at various ISO speed settings.

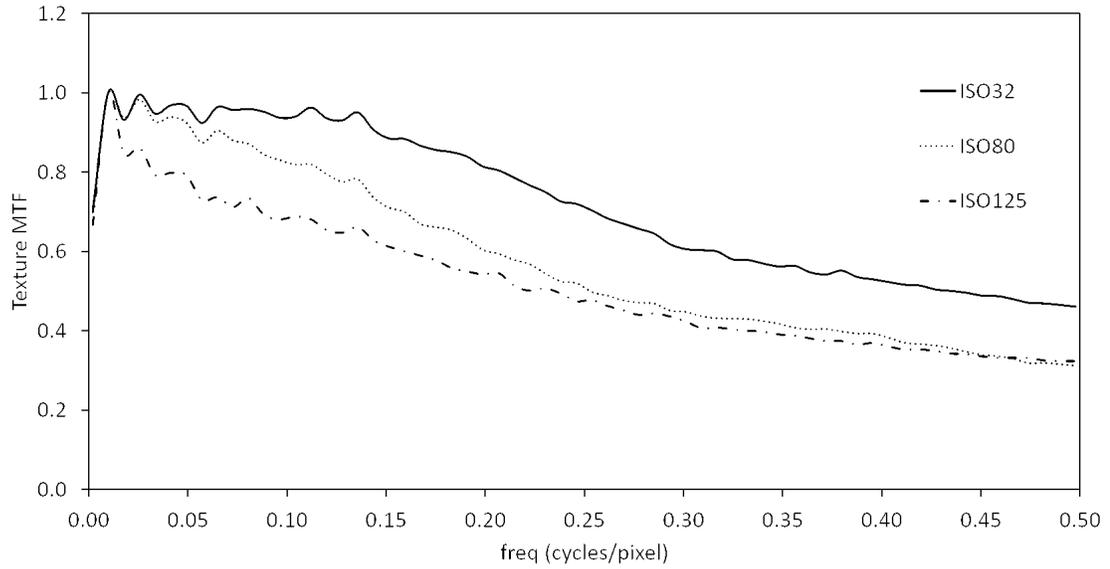


Figure 8.1: Variation of texture MTF with ISO Speed, at 1/50 s exposure time.

The acutance values at different ISO speeds, from Figure 8.2, show a monotonic decrease in acutance with an increase in ISO number. The correlation between ISO number and sensitivity to light is linear – enhancing the ISO number by a particular factor enhances the light sensitivity of the digital sensor by the same factor. From the acutance plot in Figure 8.2, when the ISO number is increased by a factor of approximately 4 from ISO32 to ISO125, it is observed that the acutance decreases from 0.88 to 0.65 (a factor of 0.73). It is also observed that the image gets saturated at around the ISO125 setting.

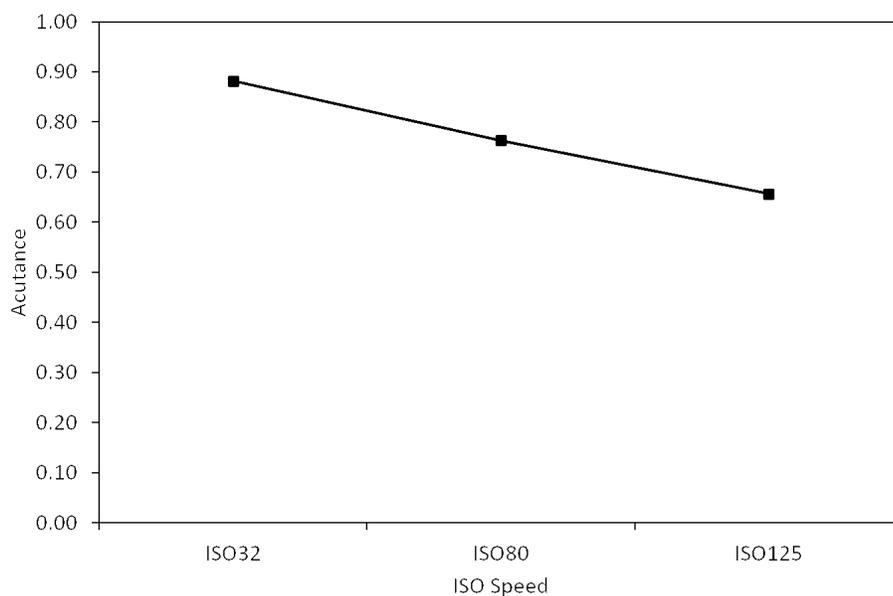


Figure 8.2: Acutance values at different ISO speeds, at 1/50 s exposure time.

8.2 Exposure Time Setting 2

The exposure time is set to 1/160 s, and the texture MTF at various ISO speeds is obtained.

For the exposure time of 1/160 s, the same pattern of decreasing acutance values is observed, for increasing ISO sensitivities. On increasing the ISO by a factor of 2 from ISO200 to ISO400, the acutance is observed to drop from 0.77 to 0.63 (a factor of 0.81). This factor is more than double the case of 1/50 s exposure time.

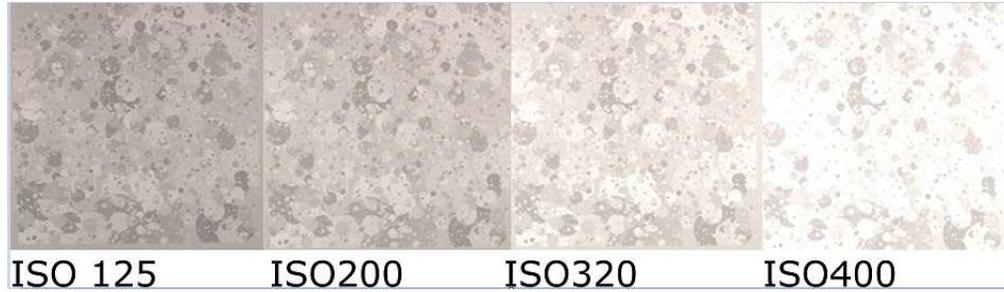


Figure 8.3: Sample images obtained at different ISO speeds, at 1/160 s exposure time.

8.3 Exposure Time Setting 3

The exposure time is decreased further to 1/550 s and the ISO numbers are taken accordingly to obtain images with reasonable levels of exposure.

The acutance values for the images taken at 1/550 s exposure time show a decreasing trend with increasing ISO number. This fits with the earlier observation in the cases of 1/50 s and 1/160 s exposure time. In the case of 1/50 s exposure time, the acutance values followed a linearly decreasing trend. At shorter exposure time values, the acutance values decrease at a much higher rate. For a change from ISO640 to ISO1250, the acutance decreases by a factor of 0.85.

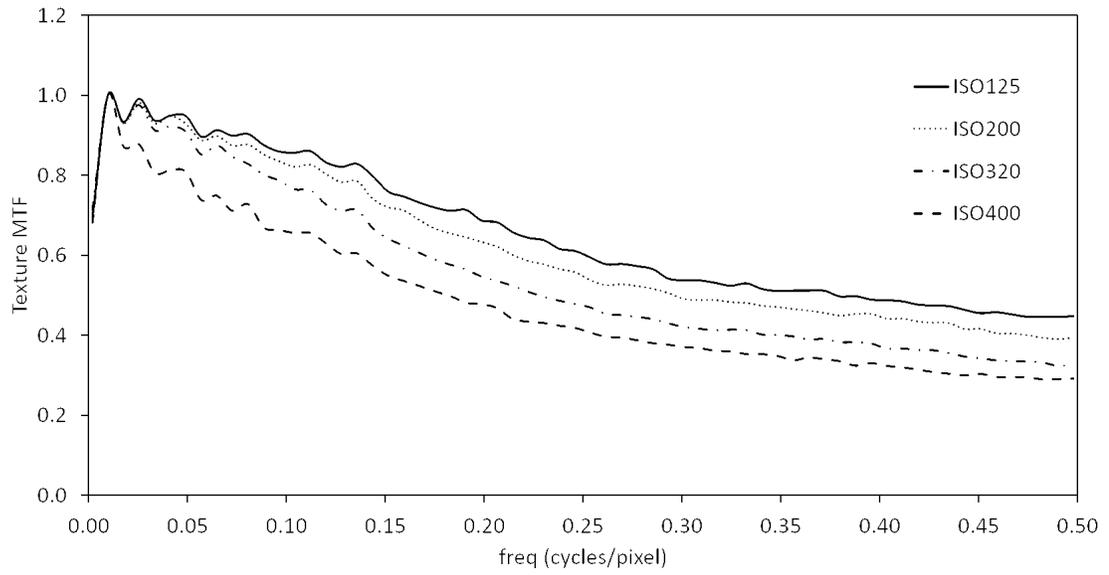


Figure 8.4: Variation in texture MTF with ISO speed, at 1/160 s exposure time.

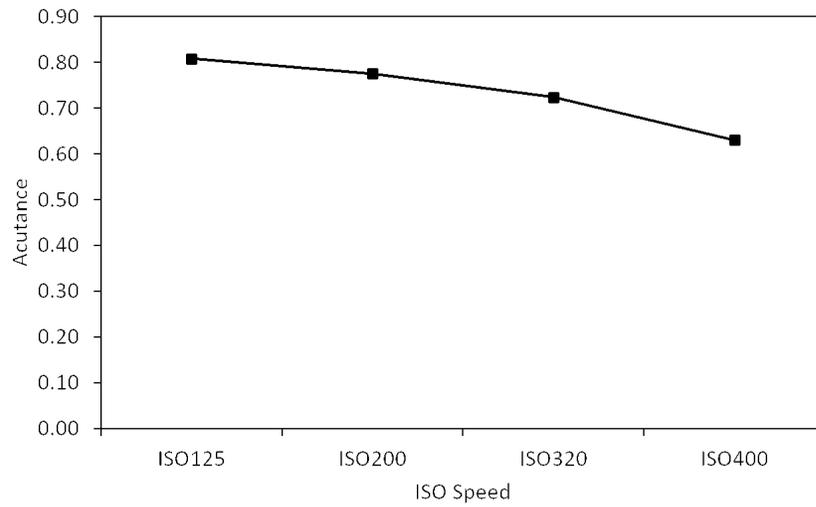


Figure 8.5: Acutance values for different ISO speeds, at 1/160 s exposure time.

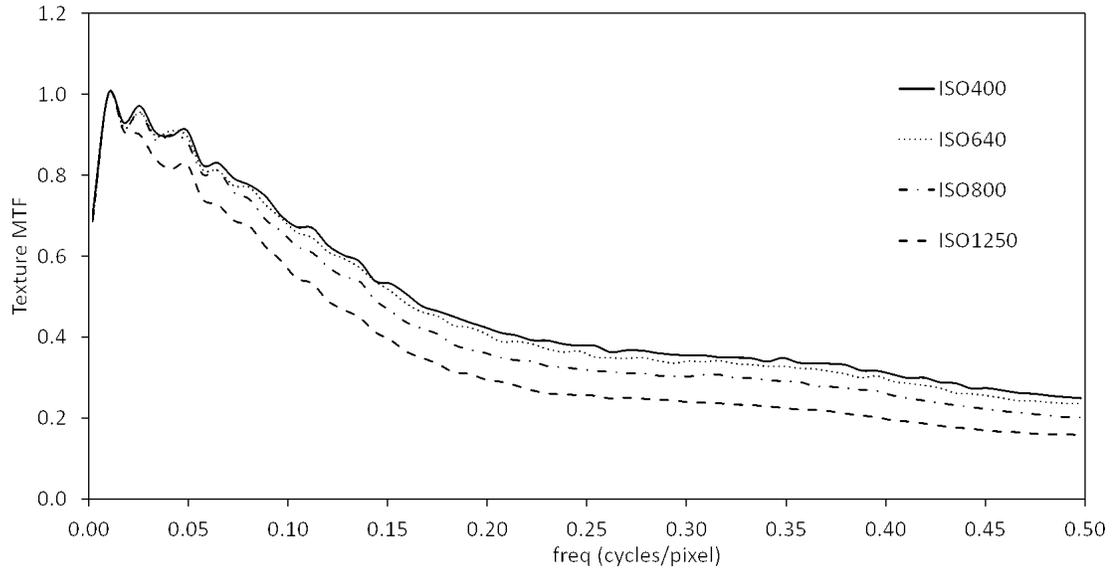


Figure 8.6: Variation in texture MTF with ISO speed, at 1/550 s exposure time.

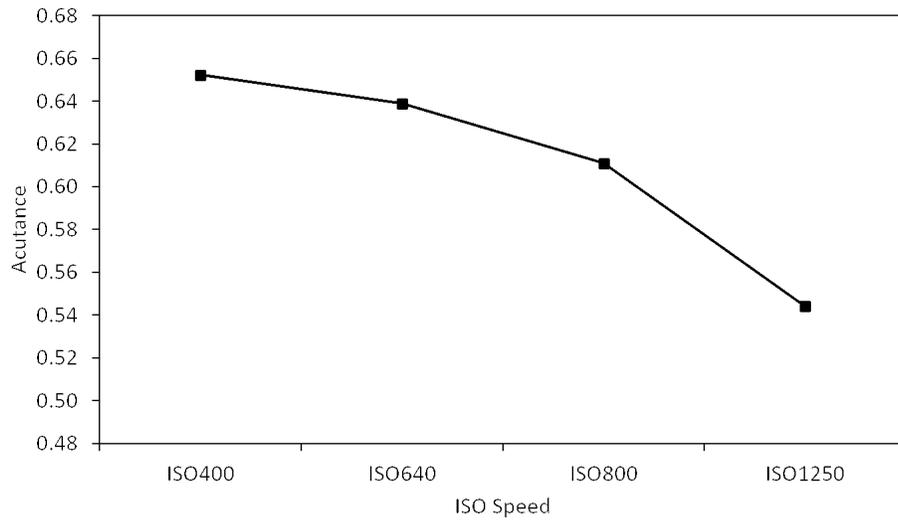


Figure 8.7: Acutance values for different ISO speeds, at 1/550 s exposure time.

Chapter 9: Noise Addition

Noise in a digital photograph is a random variation in the image brightness or in the colour values present in the image. It is more noticeable when the images are taken in low-illumination conditions. Alternatively, high ISO speed settings can also lead to more visible noise present in the image. The signal-to-noise ratio of an image gives an idea about the proportion of noise present in the image.

Noise models that are usually observed in digital images include Gaussian noise, Poisson noise, Salt & Pepper noise and combinations of these. The high-frequency random noise present in images share a similar frequency characteristic to fine-scaled texture details and edges, and camera processing for noise removal can result in the removal of these fine details as well.

9.1 Gaussian Noise Model

Gaussian noise is random noise following the normal probability distribution, where the probability density function p of a Gaussian random variable z is given by Equation 9.1.

$$p(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(z-\mu)^2}{2\sigma^2}} \quad (9.1)$$

In Equation 9.1, z is the gray level, μ the mean value, and σ the standard

deviation. Gaussian noise usually arises in digital images at the time of acquisition and/or transmission [12]. Examples of sources of Gaussian noise include sensor noise caused by insufficient illumination and high temperature. During transmission, Gaussian noise can manifest as an effect of electronic circuit noise.

Sensor noise can either be fixed pattern noise, or thermal noise. Fixed pattern noise consist of spatially fixed variations of the output signal, and can be perceived more easily as compared to random variations. The thermal noise in a CMOS image sensor is mainly Johnson-Nyquist noise, which is caused due to the thermal fluctuations or agitations of charge carriers in the circuit. The Johnson noise in electrical resistors has a power spectrum which is flat over a very large bandwidth. Due to this reason, white noise can be used as a model for the thermal noise without any significant loss in accuracy [13]. Thermal noise is also modelled as a Gaussian random process, from the central limit theorem, due to being comprised of the superposition of many independent events (i.e. thermal agitations of the charge carriers). Thus the overall model for thermal noise would be an additive zero-mean Gaussian process, with independent samples, which are also independent of the intensity.

Johnson Nyquist noise also includes the noise caused due to the reset operation [14]. On resetting the accumulated charge through a reset transistor, the thermal noise $4k_b \cdot T \cdot R_{on} \cdot \delta f$ is sampled in the accumulation node. Here δf is the frequency bandwidth, R_{on} is the ON-resistance of the reset transistor, k_b is Boltzmann's constant and T is the absolute temperature in kelvin. The configuration of the accumulation node depends on the sensor utilized in the camera.

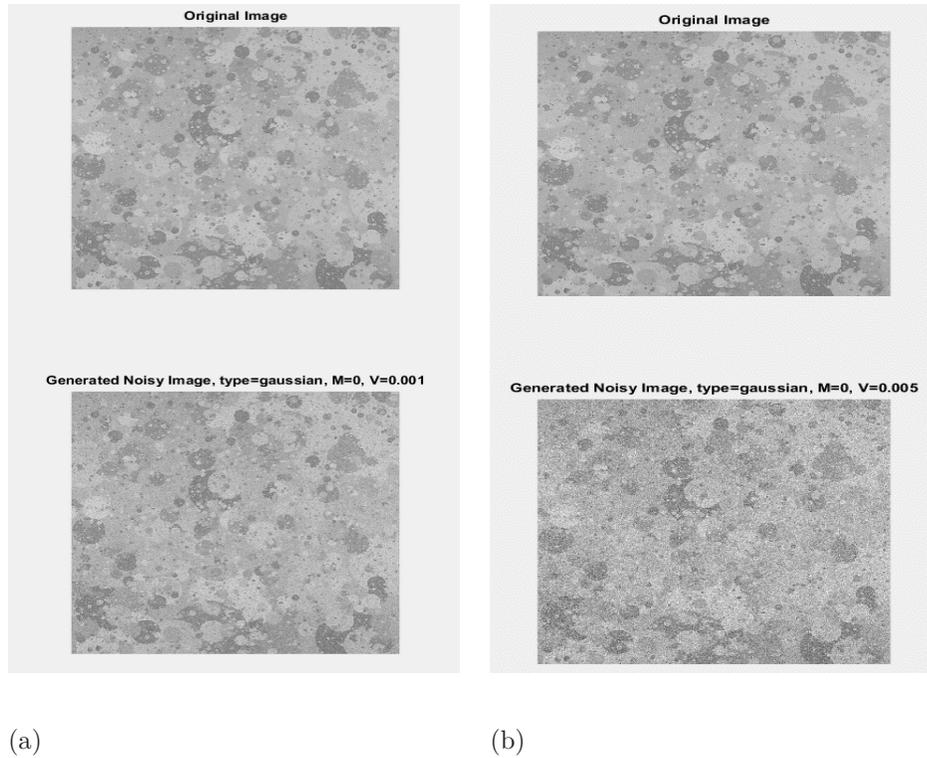


Figure 9.1: The figure shows a comparison between the original image and image with Gaussian noise added. The mean and variance levels are stated with reference to image having intensity between 0 and 1.

For our experiment, Gaussian noise of varied mean and variance levels are added to the image obtained from the camera. The texture MTF and acutance values obtained from the various images are then compared. An iPhone 5s is utilized for obtaining the images, using the dead leaves target. The illumination level is fixed at 1400 lx.

Figure 9.2 graphs the texture MTF values obtained for images with added Gaussian noise at various mean and variance levels. For the addition of low level variances of noise, there is an observed improvement in the texture MTF values at high frequencies, with a corresponding increase in acutance, as observed in Figure

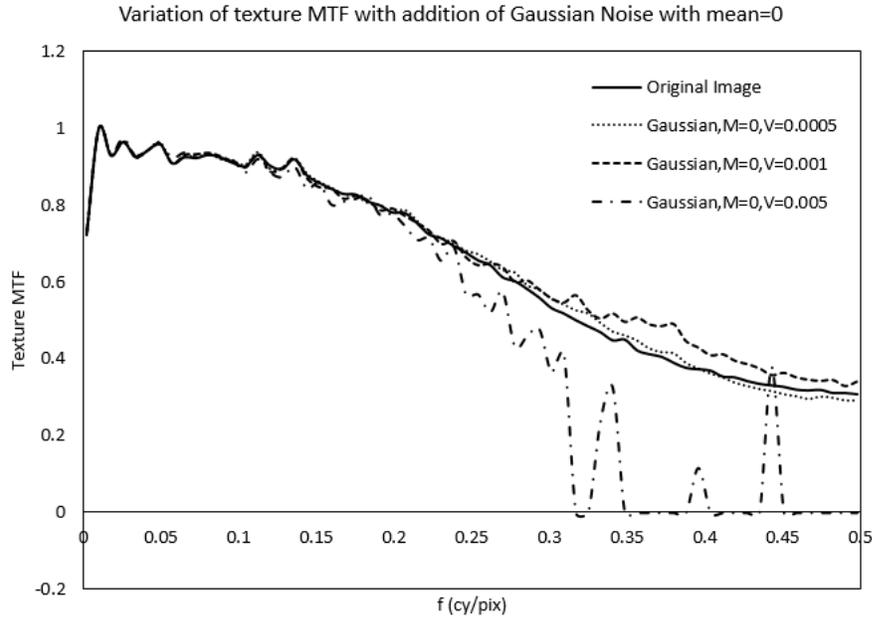


Figure 9.2: Variation in texture MTF, Gaussian noise images.

9.3. This improvement in texture MTF may be due to noise appearing as high-frequency detail within the image. At high levels of noise, the noise power spectrum subtraction gives rise to low texture MTF, which correlates with the observed visual image quality.

Adding Gaussian noise of non-zero mean shows results having approximately the same texture MTF and acutance values, as observed from Figure 9.4 and Figure 9.5.

9.2 Poisson Noise Model

Poisson or photon noise is an uncertainty in the measurement of light, which arises due to the quantized nature of light, and the independence of photon detec-

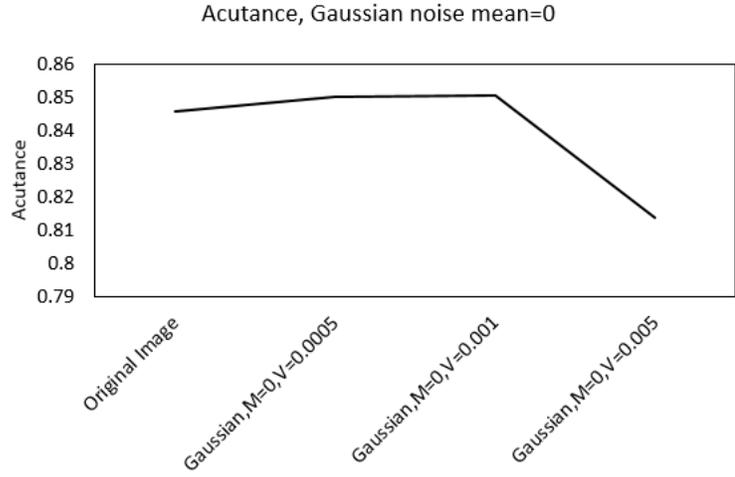


Figure 9.3: Acutance values for images with added Gaussian noise.

tions [15]. The number of discrete photons incident on the camera sensor determine the final intensity observed in the image. The independence of random arrivals of the photons lead to noise, which depends on the signal level at that sensor location. This noise model is signal dependent, contrary to the Gaussian model which is signal independent. In a realistic scenario, there would be a combination of different noise models which would be signal dependent in most cases.

Taking the photon count presents itself as a Poisson model, since the individual photon detections can be considered as independent events that follow a random distribution in time [16]. The number of photons N measured at a certain pixel location, over a time period t can be modelled by a Poisson distribution, as given in Equation 9.2.

$$p(N = k) = \frac{e^{-\lambda t} (\lambda t)^k}{k!} \quad (9.2)$$

With reference to Equation 9.2, k takes values (0,1,2, . . .), and λ is the ex-

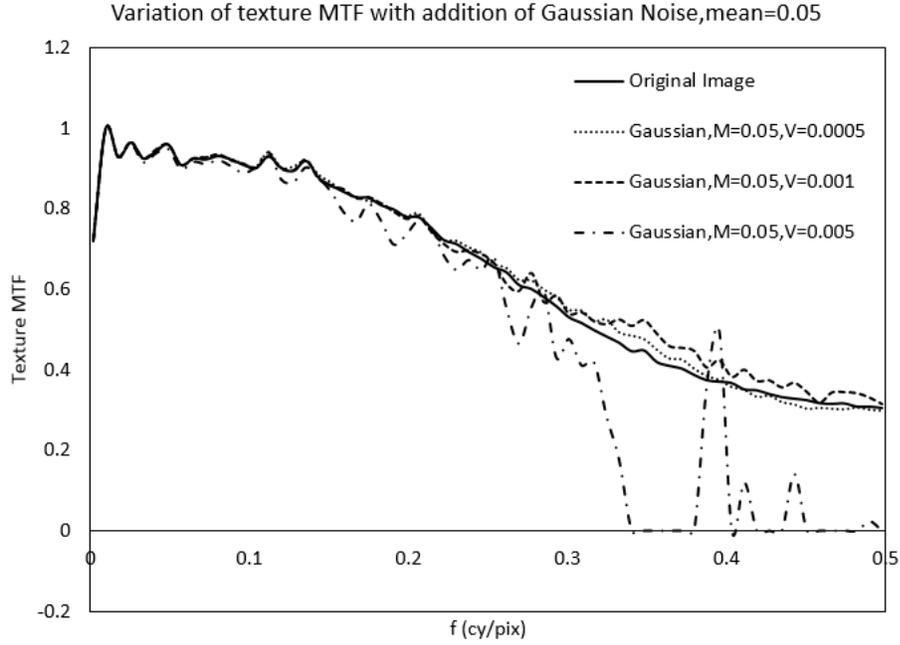


Figure 9.4: Variation in texture MTF on addition of Gaussian noise of non-zero mean.

pected number of incident photons per unit time interval. The uncertainty described by this distribution is the photon or shot noise. By definition λt is the expected number if incident photons for a time period t , which is proportional to the final image intensity. The effect of photon noise is negligible for images taken at high SNR.

For modelling Poisson noise, the value of the pixel at each location in the image is considered as the mean of the Poisson distribution, since it would be proportional to the number of captured photons. Noise is thus not artificially added, but is obtained from the signal level at each location.

For the experiment, Poisson noise generated from the image data is added to the image obtained from the camera. The texture MTF and acutance values

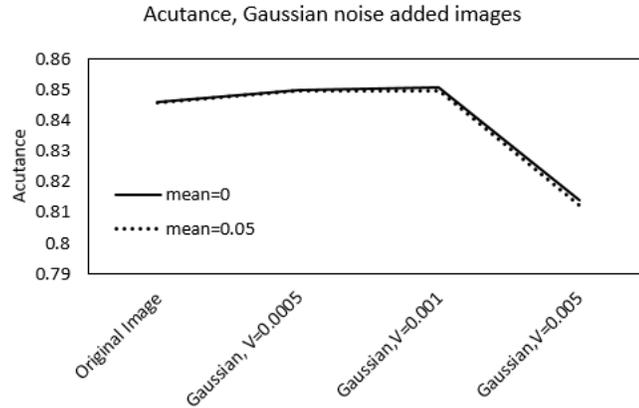


Figure 9.5: Acutance value comparison for images with added Gaussian noise.

obtained from the various images are then compared. An iPhone 5s is utilized for obtaining the images, using the dead leaves target. The illumination level is fixed at 1400 lx.

Figure 9.6 plots the variation in texture MTF between the original image, and the image with Poisson noise added. The texture MTF at the higher spatial frequencies are observed to dip, due to the subtraction of the noise power spectral density (PSD) from the image PSD. This observation of texture MTF is in line with the visual subjective image quality evaluation of the original and noisy images.

9.3 Salt and Pepper Noise Model

The salt and pepper noise model mainly characterizes pixel defects, and appears as a variation of impulsive noise. The two types of impulse noise are salt and pepper noise, and random-valued noise [17]. Impulse noise occurs in images due to bit errors in transmission, errors in signal acquisition, malfunctioning pixels in

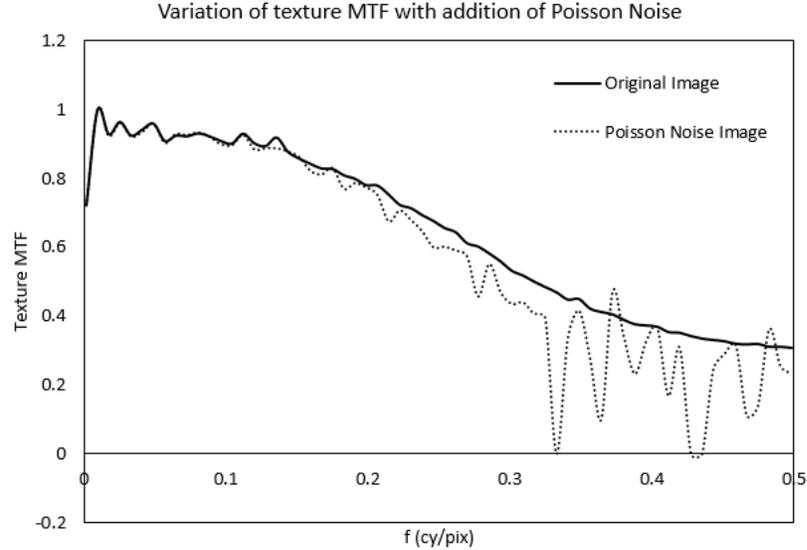


Figure 9.6: Variation of texture MTF with addition of Poisson noise.

camera sensors, faulty memory locations in hardware, or noisy channel transmissions [17] [18]. For images corrupted by salt and pepper noise, the noisy pixels take the maximum and minimum value in the sensors dynamic range. The standard median filter and its variations are generally used for the removal of salt and pepper noise [19].

For the experiment we specify the number of pixels that are affected by the salt and pepper noise, as a fraction of the total number of pixels in the image. This fraction can be called the noise density d . Several images of the dead leaves target are obtained varying the noise density, and the texture MTF for each are calculated. An example of noisy image generation is shown in Figure 9.8.

Figure 9.9 and Figure 9.10 plot the texture MTF and acutance results obtained using salt and pepper images. There is a high irregularity observed in the texture

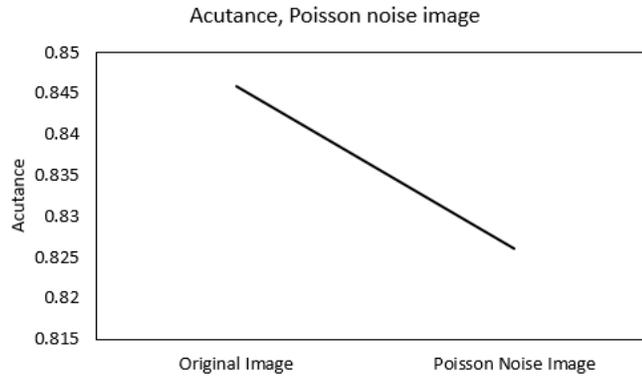


Figure 9.7: Change in acutance observed on addition of Poisson noise.

MTF plots obtained for salt and pepper noise images with a large fraction of affected pixels. The acutance values also do not show a monotonic trend of decreasing with increasing noise density. One reason for this behavior of the texture MTF can be that the salt and pepper noise mimics the appearance of extremely small discs that can be present in the ideal dead leaves target. This would artificially enhance the MTF at certain high spatial frequencies, as observed in the graph. The variation between the images at different noise densities starts at approximately the 0.25 cy/pixel mark.

9.4 Conclusions

This chapter presented the variations observed in texture MTF on artificially adding different types and intensities of noise to the dead leaves target. The different noise models considered were the Gaussian, Poisson and the Salt and Pepper noise models. It is observed that as the added noise level increases, there is a slight

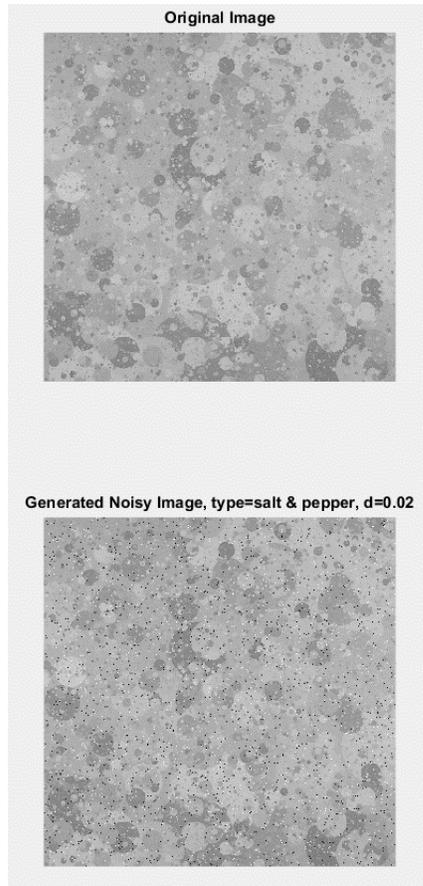


Figure 9.8: Noisy image generated with salt and pepper noise, with noise density $d=0.02$.

improvement in the texture acutance up until a certain noise threshold, with a corresponding increase in the texture MTF at high spatial frequencies. Further increasing the noise intensity above this threshold causes a decrease in the measured texture acutance. This low amount of added noise increases the apparent sharpness of the image by inflating the high frequency detail present in the image. Above the threshold intensity, the algorithm identifies the noise PSD and corrects the texture MTF according to Equation [2.4](#).

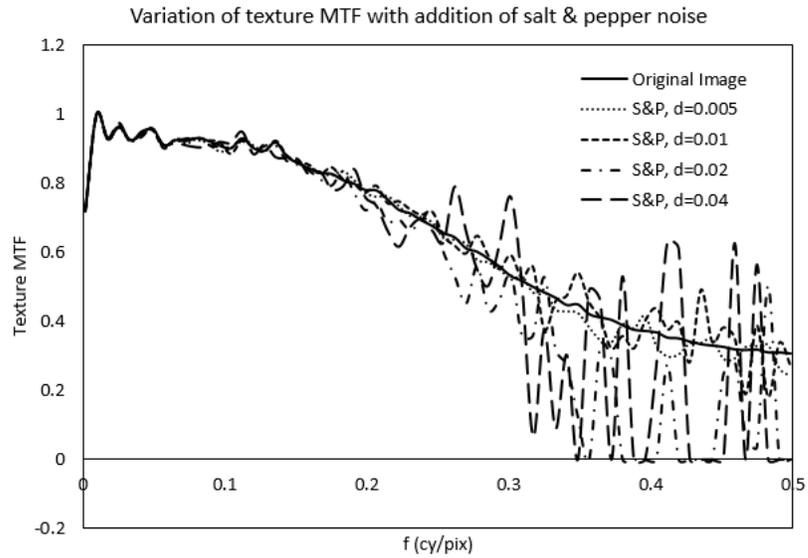


Figure 9.9: Variation in texture MTF with addition of salt and pepper noise with various noise densities.

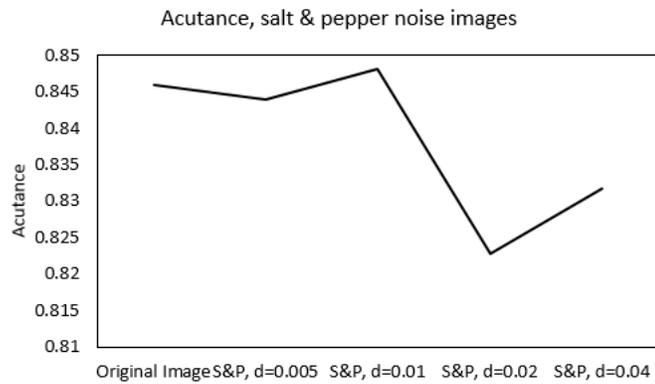


Figure 9.10: Acutance values for images with added salt and pepper noise.

Chapter 10: Implementation of the Wavelet Thresholding Approach

For everyday camera systems, captured images are corrupted by sensor noise and artifacts introduced due to various compression steps. The PSD that is generated by the target in itself falls off at higher frequencies, but the noise detail present in the image can artificially improve the measured MTF at these higher frequencies. This is the reason why the noise PSD is computed and removed from the measured target PSD, while computing the texture MTF, according to Equation 10.1.

$$SFR_{DeadLeaves}(f) = \sqrt{\frac{PS_{image}(f) - PS_{noise}(f)}{PS_{target}(f)}} \quad (10.1)$$

The current approach that has been utilized for calculating the dead leaves texture MTF makes use of a uniform 50% gray patch, with the image taken at the same time as that of the target, in order to calculate the noise power spectral density (PSD) [3]. Figure 10.1 shows an example of the dead leaves chart, with the texture region and the uniform gray regions. The noise level present in the gray image is an approximation to the true noise level present in the texture region, i.e. the dead leaves target. This approximation has been utilized in the algorithm since it is difficult to calculate the noise level present in the actual texture region. The first question to be answered is whether the noise variance in the gray region is actually a good approximation to the noise variance present in the texture region.

In the case that the noise variance for the texture region is different from the noise variance in the gray region, the next step is to identify a technique to accurately estimate the noise variance in the texture region.

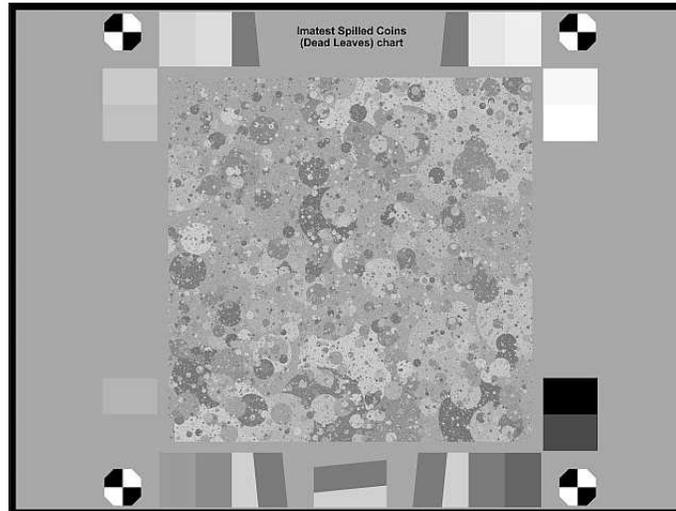


Figure 10.1: Example Dead Leaves chart.

10.1 Region for estimating the noise variance

The dead leaves model consists of a region of low-contrast high-frequency detail with components present at several scales and locations. Texture as such, refers to regions with repetitive patterns and structures at various scales and orientations. The imaging target consists of both the highly detailed dead leaves texture region, and a uniform 50% luminance gray region. This image passes through the image processor in the camera, and undergoes various enhancements, denoising and compression processes before the final observed image is produced. Since the image processing in the camera might be region- dependent, the homogeneous gray region

in the target image might not be processed in the same manner as the texture region. It is a simple matter to denoise the homogeneous region, as even a simple low-pass filter can remove the noise present in this region. On the other hand, such a simplistic technique would negatively affect the reconstruction of textured regions in the image [20]. Liu et al. [21] describe a technique to estimate the noise variance in images by first determining the weakly textured patches in the image, and then using these patches to estimate the noise variance. However since the dead leaves texture model almost completely comprises of texture detail, this method would not be suitable for the noise variance estimation.

To identify the noise variance present in both the gray region, and the texture region, we first obtain a denoised estimate of the target image using multiple target images. A similar averaging approach has been utilized in [22]. A tripod is important to ensure stabilization so that the images can be properly aligned and registered. These multiple images are then combined using weighted pixel-level image averaging in order to get the denoised estimate. In general it can be assumed that the noise model present in the image is the additive zero mean Gaussian noise. Suppose the multiple images obtained from the original image x , with the addition of noise z_n are denoted by y_n . The weighted average is estimated from these copies by Equation 10.2.

$$A = \sum_{i=1}^n \alpha_n y_n \tag{10.2}$$

where, the α_n are inversely proportional to the noise variance in each image. As a

consequence, the α_i are calculated using Equation 10.3.

$$\alpha_i = \frac{1}{\sigma^2} \bigg/ \sum_{i=1}^n \frac{1}{\sigma^2} \quad (10.3)$$

The weightage based on an inverse proportion to the noise variance ensures that higher weightage is given to those images which contain a lower proportion of noise with respect to the signal. This method of averaging results in enhanced SNR as a result of improving the signal content, while the noise deteriorates by a factor \sqrt{n} , where n is the number of images.

After the weighted estimate of the denoised image is calculated, the pixel-wise difference between this estimate and each of the multiple input images gives a model of the noise present in the image, according to Equation 10.4. The standard deviations from the gray level region and the texture regions are then obtained from this noise model in order to estimate the noise levels in the different regions.

$$Image_{noise} = Image - Image_{average} \quad (10.4)$$

Along with the weighted mean, another method considered for obtaining the denoised estimate of the image was by using the pixel-level median images from among the multiple images obtained.

Table 10.1 shows the mean values and the standard deviations obtained from the Image noise obtained from 10 images, using the weighted average denoised estimate. The average value of the standard deviation observed in the gray region is 0.857, while the average value of the standard deviation observed in the texture region is 2.31. These values are obtained for pixel values ranging from 0-255.

Table 10.1: Noise standard deviations obtained for the gray region and texture regions from the $Image_{noise}$, obtained for 10 images, using the weighted average denoised estimate image

| Image No. | Gray, μ | Gray, σ | Texture, μ | Texture, σ |
|-----------|--------------|----------------|----------------|-------------------|
| 1 | 0.20194421 | 0.895047 | 0.166602608 | 2.233122 |
| 2 | 0.005228837 | 0.840611 | 0.070238372 | 2.211464 |
| 3 | 0.131421326 | 0.869204 | 0.115405404 | 2.215862 |
| 4 | 0.102922352 | 0.885005 | 0.11344647 | 2.219322 |
| 5 | 0.113790605 | 0.849941 | 0.114552582 | 2.225755 |
| 6 | 0.019116049 | 0.858227 | 0.026628263 | 2.252515 |
| 7 | -0.047301051 | 0.842842 | -0.032295462 | 2.424686 |
| 8 | -0.086185243 | 0.838951 | -0.108970432 | 2.305418 |
| 9 | -0.159606328 | 0.863088 | -0.183842161 | 2.302975 |
| 10 | -0.281330757 | 0.827521 | -0.281765644 | 2.732183 |

Table 10.2: Noise standard deviations obtained for the gray region and texture regions from the $Image_{noise}$, obtained for 10 images, using the median-sorted denoised estimate image.

| Image No. | Gray, μ | Gray, σ | Texture, μ | Texture, σ |
|-----------|--------------|----------------|----------------|-------------------|
| 1 | 0.21682164 | 0.938086 | 0.17383742 | 2.281616 |
| 2 | 0.020106267 | 0.886725 | 0.077473184 | 2.262992 |
| 3 | 0.146298756 | 0.909126 | 0.122640216 | 2.269184 |
| 4 | 0.117799783 | 0.922468 | 0.120681282 | 2.266803 |
| 5 | 0.128668035 | 0.894269 | 0.121787394 | 2.275389 |
| 6 | 0.033993479 | 0.901433 | 0.033863074 | 2.299265 |
| 7 | -0.03242362 | 0.877321 | -0.02506065 | 2.478237 |
| 8 | -0.071307813 | 0.879493 | -0.101735621 | 2.355171 |
| 9 | -0.144728897 | 0.900167 | -0.176607349 | 2.372189 |
| 10 | -0.266453327 | 0.871712 | -0.274530832 | 2.833274 |

Table 10.2 shows the mean values and the standard deviations obtained from the Image noise obtained from 10 images, using the median-sorted denoised estimate. The average value of the standard deviation observed in the gray region is 0.89, while the average value of the standard deviation observed in the texture region is 2.37. These values are obtained for pixel values ranging from 0-255.

The difference in the standard deviation values obtained for the gray region and the textured regions shows that the approximation in the noise level is not completely accurate. A better technique to evaluate the noise present in the texture region is required in order to calculate the texture MTF. Weighted averaging to obtain the denoised model goes in this direction to some extent, as it provides a noise model which can be used for the noise PSD calculation. Furthermore the noise levels obtained on using weighted- averaging are lower compared to median-sorting, which make it the preferred method. However, an improvement to this denoising method is considered later by the addition of further denoising using wavelet thresholding in the pipeline.

The number of multiple images required for the weighted averaging step is also a parameter that is considered. There is to be found some kind of a trade-off between experimental efficiency, and the SNR levels required, since it is not practical to take greater than 10 or 15 images for a target at a particular condition. Another parameter considered is the burst-mode in image capture which is becoming prevalent in the current generation of smart-phones. The burst-mode enables the capture of multiple images in quick succession by the camera, by continuously holding the capture button. It is to be seen whether there is any difference in the

obtained standard deviation values while using the burst mode, and while taking the images manually with brief time gaps in between. For the results that follow, the pixel values are normalized to lie between 0 and 1.

The sets of images are specified as follows

- Experiment 1 - A set of 10 images taken in burst mode
- Experiment 2 - A set of 10 manually taken images
- Experiment 3 - A different set of 10 manually taken images
- Experiment 4 - A set of 15 manually taken images

The mean and standard deviation values for the Image noise obtained in each case are specified in the plots that follow.

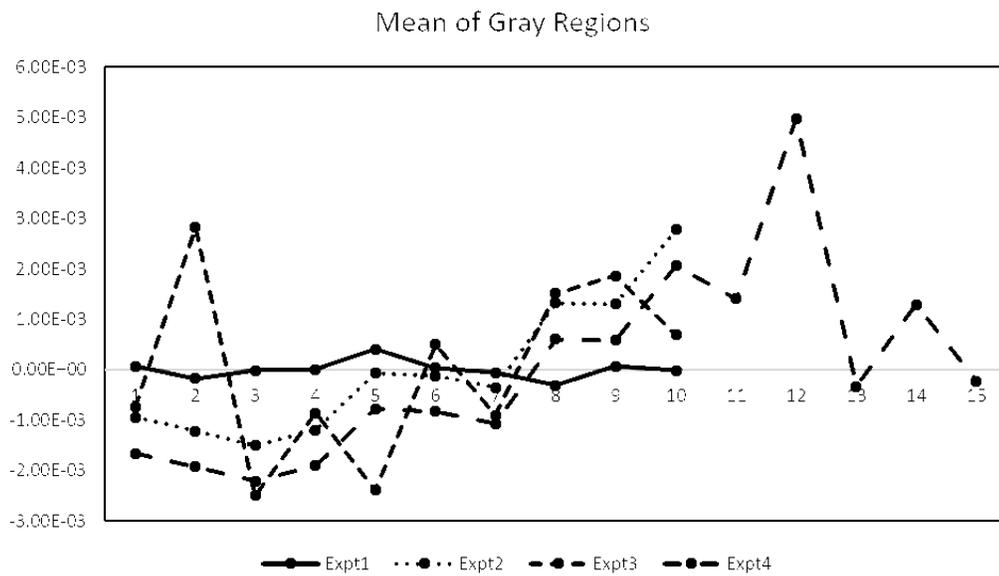


Figure 10.2: The mean noise values obtained for the gray regions, for the different sets of images.

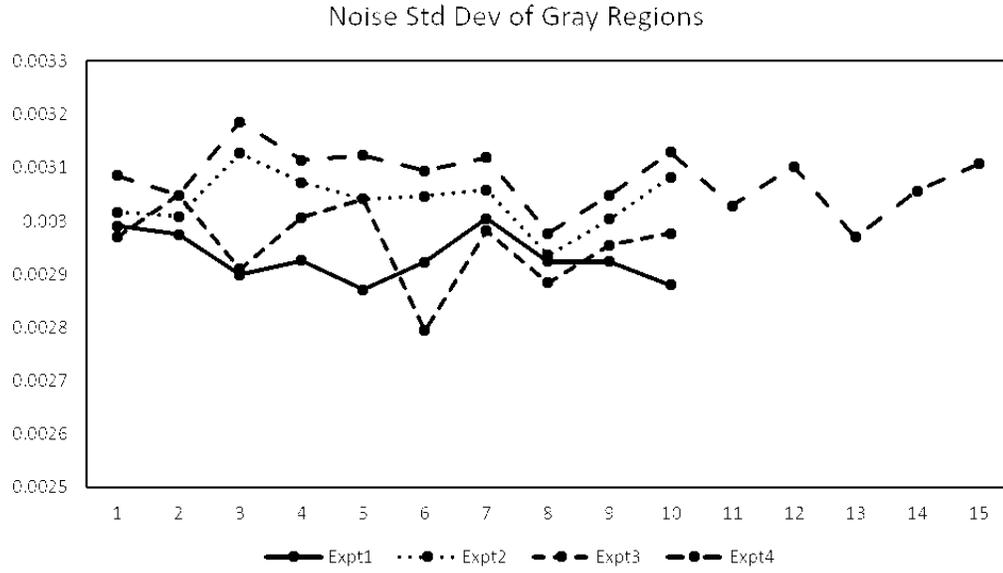


Figure 10.3: The noise standard deviations obtained for the gray regions, for the different sets of images.

Figures 10.2-10.5 graph the mean and standard deviations obtained for the $Image_{noise}$, from the different sets of images. In every case, the mean values for both the gray region and the texture region are close to zero, which validates the denoising method. Also, the values for the images taken using the burst mode and the manual modes are approximately the same. Thus we use the burst mode when it is available, and use the manual mode in the case that the feature is not present. Considering the noise standard deviations in the different sets of images, it is consistently seen that the noise levels in the gray regions are lower than those observed in the texture regions. This strengthens the earlier assumption that the mobile camera processes different regions of the image differently i.e. the homogeneous and textured regions of the image undergo different processing. Moreover, there is not much variation in

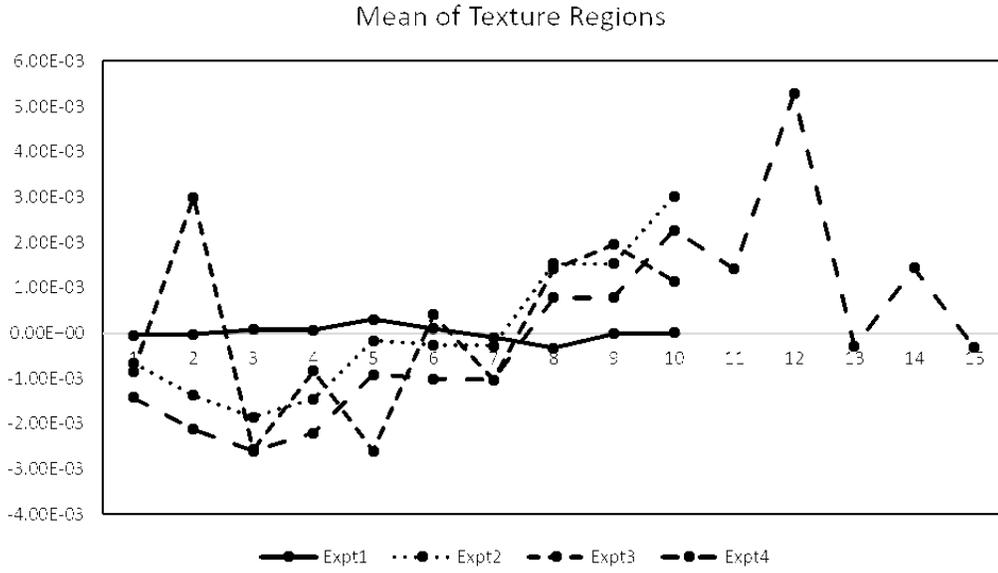


Figure 10.4: The mean values obtained for the texture regions, for the different sets of images.

the noise levels on taking a higher number of images, so we consider a minimum of 8 images, which is necessary in order to obtain an appreciable improvement in SNR.

10.2 Wavelet Thresholding for Image Denoising

The wavelet transform can be used to represent a signal or an image with a high degree of sparsity. The main signal components would be concentrated in a few coefficients, and the information contained in the rest of the coefficients would be mainly noise. The wavelet transform is similar to the windowed Fourier transform, and was developed as a better alternative to the Short-Time Fourier Transform. The Fourier transform is a form of frequency analysis which transforms time-based signals to frequency-based signals. The drawback of the Fourier transform is that the

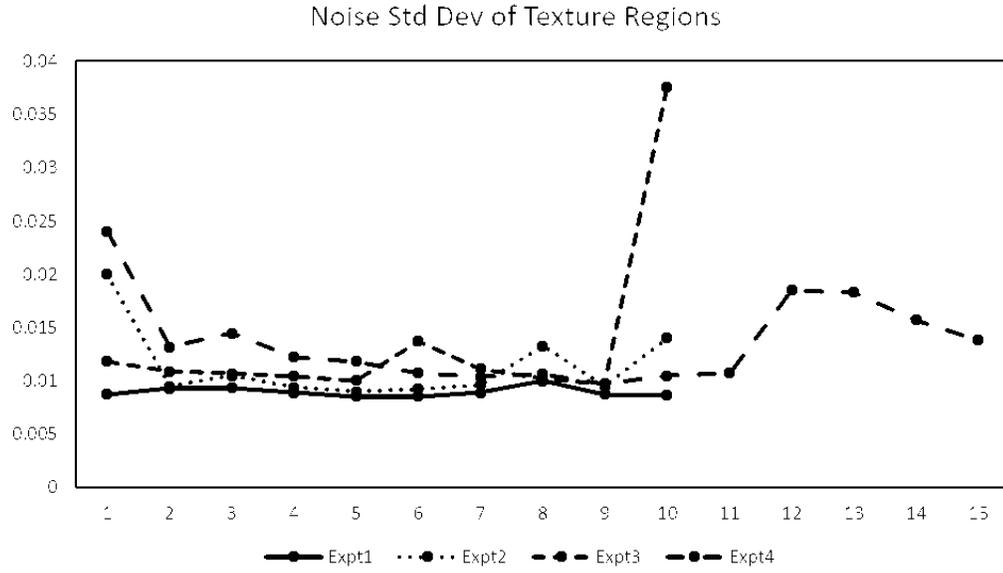


Figure 10.5: The noise standard deviations obtained for the texture regions, for the different sets of images.

time information in the signal is lost in its Fourier transform. The main advantage of the wavelet transform is the different perspective it provides on data analysis, which is according to scale [23]. The wavelet transform provides both frequency and time information.

10.2.1 The Discrete Wavelet Transform

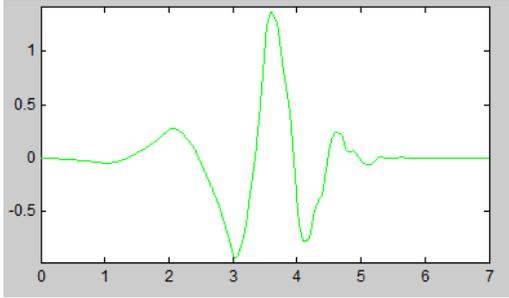
The wavelet analysis procedure starts with the selection of a mother wavelet or an analyzing wavelet $\psi(t)$, which is then scaled and translated to get the family of associated wavelets, according to Equation 10.5, where s is the scale factor and τ is the translation factor [24]. The wavelet transform then represents any arbitrary function as a superposition of these generated wavelets. For the discrete wavelet

transform, which applies in the case of 2-D images, we obtain a set of coefficients which represent the weights for each $\psi_{s,\tau}(t)$.

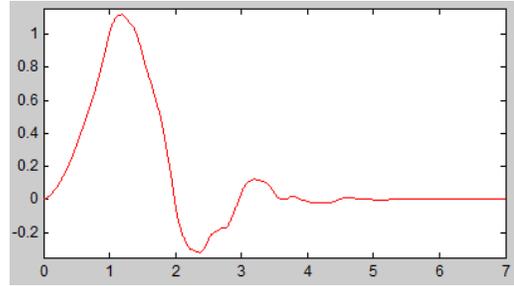
$$\psi_{s,\tau}(t) = \frac{1}{\sqrt{s}}\psi\left[\frac{t-\tau}{s}\right] \quad (10.5)$$

Several different wavelet families have been proposed which include Daubechies wavelets [25], Coiflets [26] [27], Symmlets [27], Haar wavelets [28], and so on. The variations between the families depend on the initial mother wavelet used and signal scaling and translation definitions. The selection of the wavelet used may affect the performance of the denoising technique, and this is one of the parameters which we have inspected. Wavelets with compact support, which are smooth and have vanishing moments are preferred [29]. Some examples of these analyzing wavelets are displayed in Figures 10.6(a)-10.8(b). A wavelet has p vanishing moments if and only if the wavelet scaling function can generate polynomials up to degree $p-1$. The numbers beside the wavelet type in the figure captions represent the number of vanishing moments. The Daubechies wavelets are a family of orthogonal wavelets, whose associated wavelet transform is orthogonal. The inverse wavelet transform of an orthogonal wavelet transform is its adjoint. For a given support width, Daubechies wavelets have the maximum number of vanishing moments. Symmlets have also been used frequently in literature for denoising purposes.

The discrete wavelet transform (DWT) is observed to provide a fast computation procedure for obtaining the wavelet transform, when the wavelets are discretely sampled. For a 1-D signal or a 2-D image, its DWT is calculated by passing the signal/image through a series of filters. This series consists of several filters with

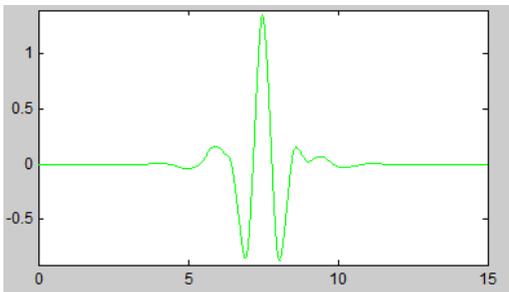


(a) Daubechies-4 wavelet function.

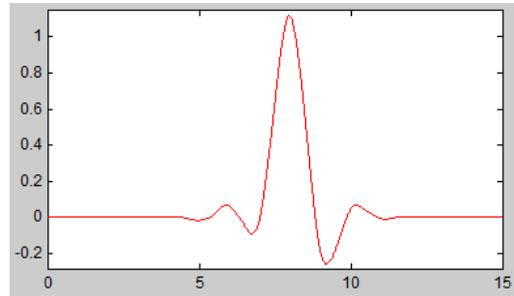


(b) Daubechies-4 scaling function.

Figure 10.6: Daubechies-4 wavelet and scaling functions



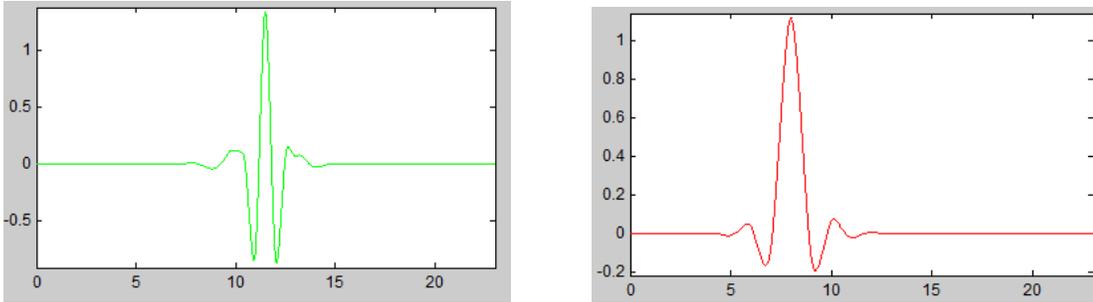
(a) Symmlet-8 wavelet function.



(b) Symmlet-8 scaling function.

Figure 10.7: Symmlet-8 wavelet and scaling functions

different cutoff frequencies at different scales. The signal in question is decomposed simultaneously through related low-pass and high-pass filters that together form the analysis filter bank. After the signal passes through a set of signals, it is further subsampled, which determines the scale. Figure 10.9 gives a representation of the DWT decomposition of a 1-D signal. The high pass filter is denoted as $h(n)$ and the low pass filter as $g(n)$ at each level. The detail coefficients are obtained from the high pass filters, and approximation coefficients are obtained from the low-pass filters. Due to half of the frequencies being removed at each filter operation, by the Nyquist theorem, the signal can be down-sampled by 2. A sufficient number of levels of decomposition can be specified for time and frequency resolution. Reconstruction



(a) Coiflet-4 wavelet function.

(b) Coiflet-4 scaling function.

Figure 10.8: Coiflet-4 wavelet and scaling functions

of the signal is carried out by the reverse of the decomposition, i.e. by up-sampling the detail and approximation coefficients, passing them through the respective high and low filters respectively and adding them, at each level. Carrying out the process for the same number of levels as the decomposition would result in obtaining the original signal.

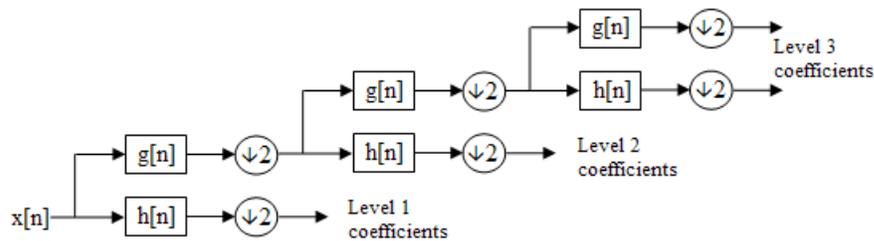


Figure 10.9: The Mallat algorithm for the DWT decomposition.

In the case of a 2-D image the wavelet decomposition is performed for the rows and then columns, to obtain 4 sub-bands which give the diagonal detail coefficients (HH or high-high), the horizontal detail coefficients (HL or high-low), the vertical detail coefficients (LH or low-high), and the approximation coefficients (LL or low-low), as shown in Figure 10.10. The LL band contains the low frequency components

which are further split at higher levels of the decomposition.

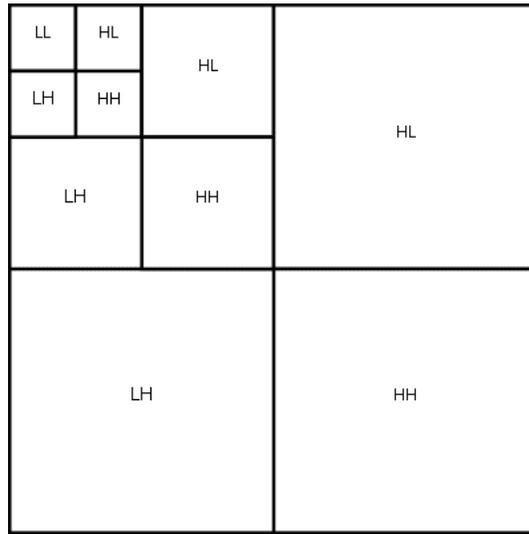


Figure 10.10: Description of the 2-D DWT decomposition in images showing the HH, HL, LH sub-bands, and further decomposed LL sub-band.

10.2.2 Denoising via Wavelet Thresholding

One of the earliest proposals of the application of wavelet thresholding for denoising was by Donoho et al. [30]. They proposed an algorithm for the shrinkage of the wavelet coefficients, which would result in a denoised version of the signal. This seminal work has been followed by several extensions dealing with the same technique of wavelet thresholding. Chang et al. [31] describe that compression of the signal, using coefficient quantization, is an approximation to wavelet thresholding, and results in denoising of the signal. Chang et al. [32] provide a discussion on the estimation of a data-driven threshold for use in wavelet thresholding, which is derived in a Bayesian framework. Kaur et al. [33] propose another adaptive

threshold estimation method for image denoising, based on generalized Gaussian distribution (GGD) modelling of the sub-band coefficients, which has become a prevalent model in use. It is observed in a large class of natural images that the sub-band coefficients can be modelled by a GGD, which can be further simplified to a Laplacian distribution [34]. Since the dead leaves target models these kind of natural images, there is further motivation for utilizing wavelet thresholding for the denoising of the captured dead leaves image. Wavelet thresholding has also been used for signal compression, and coding [35]. Luisier et al. [36] propose an approach to wavelet thresholding which does not model the sub-band coefficients and instead uses the Steins unbiased risk estimate as a mean-squared error (MSE) estimate and minimizes this for threshold estimation.

The main motivation for using wavelet thresholding for the purpose of denoising is because of the sparse representation which it provides. The noise is spread uniformly across all the coefficients and can be removed, since most of the image information is concentrated in the largest coefficients. One important parameter which affects the accuracy of wavelet thresholding is the threshold selection. This determines the amount of information which is lost, and ultimately the final quality of the denoised image. Another parameter which affect denoising performance is the type of thresholding that is used soft thresholding or hard thresholding. The theoretical justifications for the performance of soft-thresholding have been proposed in detail [37]. Figure 10.11 shows the operations of soft and hard thresholding.

The soft thresholding function provides a smoother transition, with wavelet shrinkage, and is usually preferred in the case of image denoising as it provides more

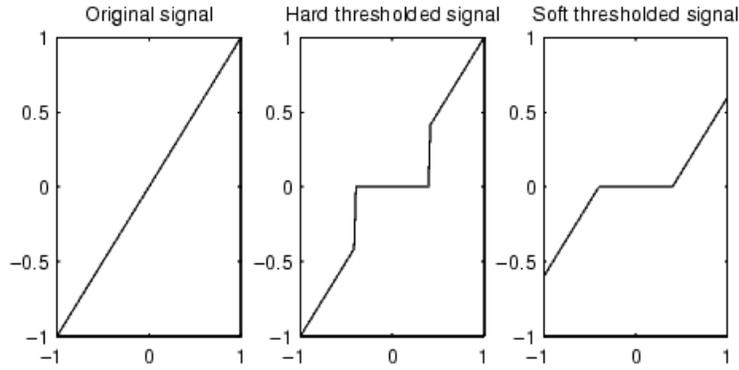


Figure 10.11: Functions for soft and hard thresholding.

visually pleasing images. The soft thresholding function is defined as in Equation 10.6, where λ is the threshold.

$$\eta_{\lambda,s}(x) = \text{sgn}(x)(|x| - \lambda)_+ \quad (10.6)$$

Hard thresholding, defined as in Equation 10.7, is a cruder approach which completely turns to zero those coefficients which are below the threshold. As a consequence of this, the images denoised using hard thresholding also have a slightly cruder quality.

$$\eta_{\lambda,h}(x) = x \cdot 1_{(|x|>\lambda)} \quad (10.7)$$

Denoising of the signal is performed after the stage of DWT decomposition by thresholding appropriate wavelet coefficients, before carrying out the inverse wavelet transform. Assuming that the noise model generally present in captured images is the i.i.d. Gaussian model, $\nu = N(O, \sigma^2)$ we have our model for the noisy signal y as shown in Equation 10.8.

$$y = x + \nu \quad (10.8)$$

Let Y be the vector of wavelet transformed coefficients of the noisy signal. An estimate \hat{x} of the original signal is obtained by thresholding the wavelet coefficients Y and then performing the inverse transform. The threshold is selected so as to minimize the mean squared error between the original image and the estimate.

The quality of the denoising is measured by calculating the PSNR of the denoised image according to Equation 10.9. We assume that we have the pixel-intensity details of the original image in this case, since the noisy images in the experiments are obtained by manual addition to this original image.

$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX_I^2}{MSE} \right) \quad (10.9)$$

Where MAX_I is the maximum possible pixel value of the image, which is generally 255 for 8-bit images. The MSE is the mean squared error between the noisy image y and the original noise-free image x , given by Equation 10.10. The original image and the noisy image have to be the same size $m \times n$, and are grayscale.

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (y(i, j) - x(i, j))^2 \quad (10.10)$$

Choosing an appropriate threshold is important to obtain good performance for the denoising procedure. Using a larger than necessary threshold would result in the loss of texture details, and lead to an unnatural blurring of the image. At the same time, too small of a threshold would result in no observable denoising effect. We consider both a universal threshold, and a sub-band adaptive threshold for the purposes of this study. Donoho and Johnstone [30] proposed the universal threshold of $\sigma\sqrt{2\log(N)}$, where N is the number of pixels in the image, and σ is the noise

variance which is estimated by the robust median estimator in the highest sub-band HH_1 according to Equation 10.11, with all $Y_{ij} \in HH_1$.

$$\hat{\sigma} = \frac{\text{Median}(|Y_{ij}|)}{0.6745} \quad (10.11)$$

The second threshold considered is one using a level-wise threshold estimator method proposed by Birge and Massart [38]. Suppose Y_0 denotes the decomposition level, m is the length of the coarsest approximation (> 2), and $\alpha (> 1)$ is the sparsity level, the thresholding strategy is as follows -

1. At the coarser level $Y_0 + 1$, all coefficients are retained.
2. For any level Y from 1 to Y_0 , the K_Y larger coefficients are retained according to Equation 10.12.

$$K_Y = \frac{m}{(Y_0 + 1 - Y)\alpha} \quad (10.12)$$

The parameter α defines the sparsity level of the retained number of thresholds, and usually varies from 1 to 3.

10.2.3 Results

We consider the effects of the following parameters on the denoising performance utilizing wavelet thresholding -

1. Added noise model (Gaussian, Poisson and Salt & Pepper noise) and noise value.
2. The wavelet used for the transform - Daubechies, Symmlet and Coiflet wavelets are considered.

3. The threshold estimation technique utilized - universal or the level-wise thresholding.
4. The type of thresholding utilized - soft or hard thresholding.
5. The number of levels of decomposition utilized.

The performance of the de-noiser based on each parameter is determined by the PSNR of the final denoised image, as according to Equation 10.9. Figures 10.12-10.14 give examples of some of the obtained denoised images. For these examples, the denoising is performed using a level-5 Symmlet 8 wavelet, with soft thresholding. The mean and variance of the Gaussian noise are with respect to pixel values that have been normalized between 0 and 1.

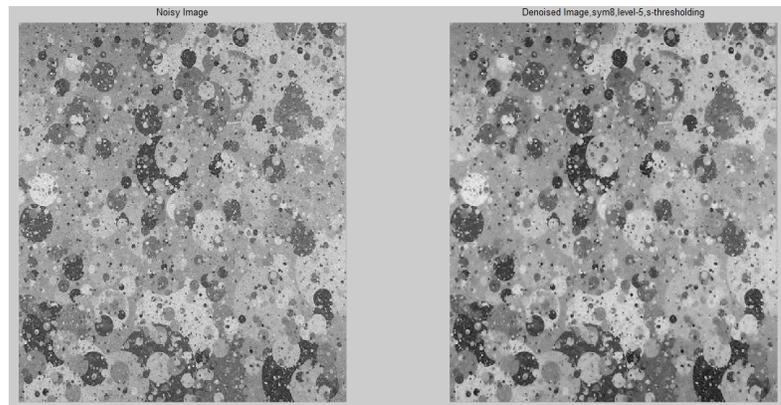


Figure 10.12: Denoising performance for an image having Gaussian noise with mean=0 and variance=0.0005.

Table 10.3 displays the PSNR values obtained on denoising images using the Symmlet-8 wavelet, at various levels of decomposition.

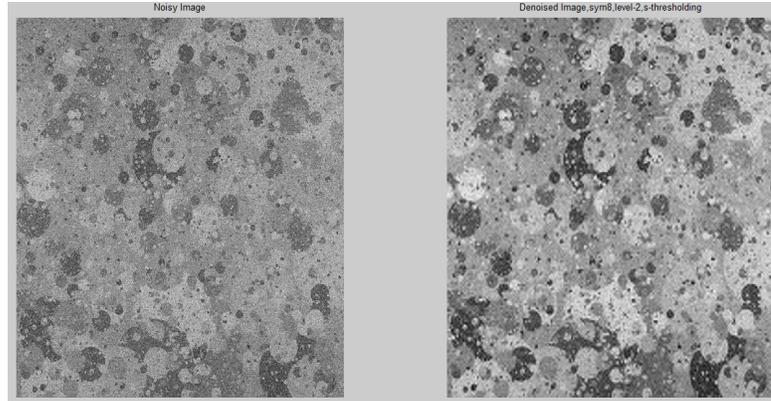


Figure 10.13: Denoising performance for an image having Gaussian noise with mean=0 and variance=0.005.

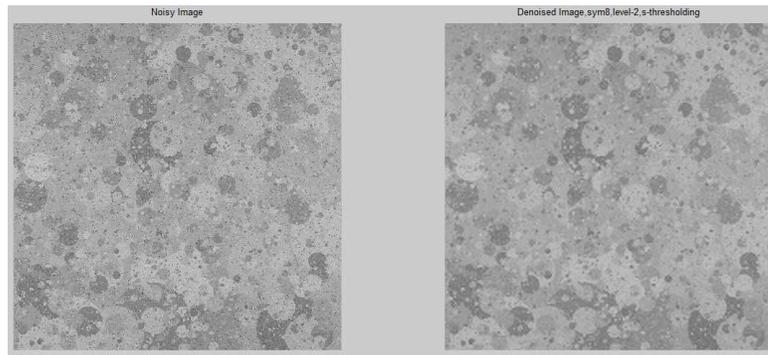


Figure 10.14: Denoising performance for an image having salt & pepper noise with noise density=0.02.

Table 10.3: PSNR values of the denoised image obtained after denoising using wavelet thresholding, under variation of several parameters

| Gaussian noise model, $m=0$, $v=0.0005$, initial PSNR=35.7048 | | | | | | | | | |
|---|--|-----------------|-----------------|------------------------|------------------------|------------------------|------------------------|---------------------------|---------------------------|
| Threshold used | | Universal, soft | Universal, hard | B-M, $\alpha=3$, soft | B-M, $\alpha=3$, hard | B-M, $\alpha=2$, soft | B-M, $\alpha=2$, hard | B-M, $\alpha=1.25$, soft | B-M, $\alpha=1.25$, hard |

| | | | | | | | | | |
|-------|---|---------|---------|---------|---------|---------|---------|----------------|---------|
| Level | 1 | 36.4715 | 36.1039 | 36.5386 | 36.0543 | 36.6581 | 35.8764 | 36.6477 | 35.8275 |
| | 2 | 35.8083 | 35.9045 | 36.2034 | 35.9130 | 36.8936 | 35.8216 | 36.8991 | 35.7745 |
| | 3 | 35.2971 | 35.8236 | 35.8094 | 35.8581 | 36.7992 | 35.8068 | 36.8405 | 35.7606 |
| | 4 | 35.1500 | 35.8161 | 35.6822 | 35.8524 | 36.7622 | 35.8046 | 36.8155 | 35.7577 |
| | 5 | 35.1163 | 35.8158 | 35.6507 | 35.8525 | 36.7539 | 35.8033 | 36.8091 | 35.7566 |

Gaussian noise model, m=0, v=0.001, initial PSNR= 33.6143

| Threshold used | | Universal, soft | Universal, hard | B-M, $\alpha=3$, soft | B-M, $\alpha=3$, hard | B-M, $\alpha=2$, soft | B-M, $\alpha=2$, hard | B-M, $\alpha=1.25$, soft | B-M, $\alpha=1.25$, hard |
|----------------|---|-----------------|-----------------|------------------------|------------------------|------------------------|------------------------|---------------------------|---------------------------|
| Level | 1 | 35.3405 | 34.7538 | 35.3626 | 34.5913 | 35.3299 | 34.0733 | 35.2519 | 33.9312 |
| | 2 | 34.7631 | 34.6067 | 35.0951 | 34.4420 | 35.7028 | 33.9500 | 35.5996 | 33.7981 |
| | 3 | 34.0718 | 34.4669 | 34.5328 | 34.3260 | 35.5461 | 33.9075 | 35.4905 | 33.7629 |
| | 4 | 33.8659 | 34.4481 | 34.3477 | 34.3091 | 35.4866 | 33.8976 | 35.4474 | 33.7537 |
| | 5 | 33.8120 | 34.4460 | 34.2945 | 34.3061 | 35.4679 | 33.8946 | 35.4326 | 33.7508 |

Gaussian noise model, m=0, v=0.005, initial PSNR= 25.5448

| Threshold used | | Universal, soft | Universal, hard | B-M, $\alpha=3$, soft | B-M, $\alpha=3$, hard | B-M, $\alpha=2$, soft | B-M, $\alpha=2$, hard | B-M, $\alpha=1.25$, soft | B-M, $\alpha=1.25$, hard |
|----------------|---|-----------------|-----------------|------------------------|------------------------|------------------------|------------------------|---------------------------|---------------------------|
| Level | 1 | 31.0131 | 30.9271 | 31.0135 | 30.9874 | 30.9990 | 30.3036 | 30.9703 | 29.7139 |
| | 2 | 31.6893 | 31.7444 | 31.6551 | 31.6982 | 31.9863 | 30.5734 | 32.1144 | 29.4625 |
| | 3 | 30.0818 | 30.7254 | 30.0082 | 30.5792 | 30.9734 | 29.7334 | 31.3581 | 28.7104 |
| | 4 | 29.2888 | 30.4056 | 29.1738 | 30.2352 | 30.4988 | 29.4644 | 31.0005 | 28.4510 |
| | 5 | 29.0471 | 30.3579 | 28.9032 | 30.1918 | 30.3611 | 29.3871 | 30.9004 | 28.3662 |

Gaussian noise model, m=0.05, v=0.0005, initial PSNR= 25.6054

| Threshold used | | Universal, soft | Universal, hard | B-M, $\alpha=3$, soft | B-M, $\alpha=3$, hard | B-M, $\alpha=2$, soft | B-M, $\alpha=2$, hard | B-M, $\alpha=1.25$, soft | B-M, $\alpha=1.25$, hard |
|----------------|---|-----------------|-----------------|------------------------|------------------------|------------------------|------------------------|---------------------------|---------------------------|
| Level | 1 | 25.6772 | 25.6441 | 25.6827 | 25.6393 | 25.6919 | 25.6224 | 25.6909 | 25.6178 |
| | 2 | 25.6158 | 25.6247 | 25.6523 | 25.6259 | 25.7094 | 25.6172 | 25.7099 | 25.6125 |
| | 3 | 25.5640 | 25.6171 | 25.6161 | 25.6209 | 25.7023 | 25.6156 | 25.7055 | 25.6114 |
| | 4 | 25.5483 | 25.6160 | 25.6040 | 25.6204 | 25.6995 | 25.6152 | 25.7037 | 25.6111 |
| | 5 | 25.5439 | 25.6159 | 25.6002 | 25.6202 | 25.6985 | 25.6151 | 25.7029 | 25.6110 |

Gaussian noise model, $m=0.05$, $v=0.001$, initial PSNR= 25.3671

| Threshold used | | Universal, soft | Universal, hard | B-M, $\alpha=3$, soft | B-M, $\alpha=3$, hard | B-M, $\alpha=2$, soft | B-M, $\alpha=2$, hard | B-M, $\alpha=1.25$, soft | B-M, $\alpha=1.25$, hard |
|----------------|---|-----------------|-----------------|------------------------|------------------------|------------------------|------------------------|---------------------------|---------------------------|
| Level | 1 | 25.5760 | 25.5143 | 25.5785 | 25.4961 | 25.5767 | 25.4299 | 25.5690 | 25.4114 |
| | 2 | 25.5139 | 25.4963 | 25.5520 | 25.4776 | 25.6155 | 25.4133 | 25.6058 | 25.3932 |
| | 3 | 25.4282 | 25.4798 | 25.4882 | 25.4642 | 25.6009 | 25.4076 | 25.5955 | 25.3884 |
| | 4 | 25.4001 | 25.4776 | 25.4652 | 25.4620 | 25.5950 | 25.4062 | 25.5911 | 25.3871 |
| | 5 | 25.3918 | 25.4772 | 25.4577 | 25.4615 | 25.5927 | 25.4058 | 25.5893 | 25.3866 |

Gaussian noise model, $m=0.05$, $v=0.005$, initial PSNR=22.7789

| Threshold used | | Universal, soft | Universal, hard | B-M, $\alpha=3$, soft | B-M, $\alpha=3$, hard | B-M, $\alpha=2$, soft | B-M, $\alpha=2$, hard | B-M, $\alpha=1.25$, soft | B-M, $\alpha=1.25$, hard |
|----------------|---|-----------------|-----------------|------------------------|------------------------|------------------------|------------------------|---------------------------|---------------------------|
| Level | 1 | 24.8654 | 24.8457 | 24.8655 | 24.8599 | 24.8624 | 24.6987 | 24.8559 | 24.5227 |
| | 2 | 25.0168 | 25.0274 | 25.0096 | 25.0185 | 25.0778 | 24.7647 | 25.1050 | 24.4564 |
| | 3 | 24.6124 | 24.7796 | 24.5909 | 24.7414 | 24.8449 | 24.5288 | 24.9368 | 24.1939 |
| | 4 | 24.3725 | 24.6905 | 24.3341 | 24.6446 | 24.7223 | 24.4473 | 24.8516 | 24.0971 |
| | 5 | 24.2895 | 24.6769 | 24.2391 | 24.6308 | 24.6824 | 24.4221 | 24.8241 | 24.0661 |

| Poisson noise model, initial PSNR=29.4151 | | | | | | | | | |
|---|---|-----------------|-----------------|------------------------|------------------------|------------------------|------------------------|---------------------------|---------------------------|
| Threshold used | | Universal, soft | Universal, hard | B-M, $\alpha=3$, soft | B-M, $\alpha=3$, hard | B-M, $\alpha=2$, soft | B-M, $\alpha=2$, hard | B-M, $\alpha=1.25$, soft | B-M, $\alpha=1.25$, hard |
| Level | 1 | 33.1812 | 32.7498 | 33.1795 | 32.6515 | 33.0649 | 31.2658 | 32.9350 | 30.7634 |
| | 2 | 32.8624 | 32.8258 | 32.9797 | 32.6147 | 33.6819 | 30.9611 | 33.5950 | 30.3381 |
| | 3 | 31.6563 | 32.3581 | 31.8679 | 32.1406 | 33.2360 | 30.7400 | 33.2973 | 30.1548 |
| | 4 | 31.2088 | 32.2665 | 31.4296 | 32.0369 | 33.0419 | 30.6822 | 33.1625 | 30.1052 |
| | 5 | 31.0872 | 32.2565 | 31.3013 | 32.0202 | 32.9862 | 30.6632 | 33.1214 | 30.0886 |
| Salt and Pepper noise model, density=0.02, initial PSNR= 25.0577 | | | | | | | | | |
| Threshold used | | Universal, soft | Universal, hard | B-M, $\alpha=3$, soft | B-M, $\alpha=3$, hard | B-M, $\alpha=2$, soft | B-M, $\alpha=2$, hard | B-M, $\alpha=1.25$, soft | B-M, $\alpha=1.25$, hard |
| Level | 1 | 30.4534 | 28.2361 | 30.3087 | 27.7564 | 29.4062 | 26.2942 | 28.9839 | 25.9308 |
| | 2 | 31.4131 | 28.6697 | 31.2644 | 28.0398 | 30.2739 | 26.2379 | 29.6016 | 25.7792 |
| | 3 | 30.2951 | 28.3674 | 30.2288 | 27.7271 | 29.9109 | 26.1083 | 29.3263 | 25.6649 |
| | 4 | 29.7145 | 28.2648 | 29.6550 | 27.6045 | 29.6867 | 26.0680 | 29.1674 | 25.6273 |
| | 5 | 29.5297 | 28.2518 | 29.4563 | 27.5771 | 29.6067 | 26.0543 | 29.1103 | 25.6168 |
| Salt and Pepper noise model, density=0.04, initial PSNR= 21.8779 | | | | | | | | | |
| Threshold used | | Universal, soft | Universal, hard | B-M, $\alpha=3$, soft | B-M, $\alpha=3$, hard | B-M, $\alpha=2$, soft | B-M, $\alpha=2$, hard | B-M, $\alpha=1.25$, soft | B-M, $\alpha=1.25$, hard |
| Level | 1 | 28.1704 | 27.3641 | 28.1636 | 27.2295 | 27.8178 | 24.9414 | 27.5445 | 24.2342 |
| | 2 | 30.4617 | 29.1201 | 30.4443 | 28.8776 | 29.9080 | 25.4077 | 29.3728 | 24.3489 |
| | 3 | 29.3457 | 28.5274 | 29.2793 | 28.1183 | 29.2798 | 25.0948 | 28.9180 | 24.0641 |
| | 4 | 28.4131 | 28.1609 | 28.2697 | 27.6020 | 28.7264 | 24.9221 | 28.5142 | 23.9231 |

| | | | | | | | | | |
|--|---|---------|---------|---------|---------|---------|---------|---------|---------|
| | 5 | 28.0595 | 28.0921 | 27.8571 | 27.4757 | 28.5120 | 24.8618 | 28.3567 | 23.8730 |
|--|---|---------|---------|---------|---------|---------|---------|---------|---------|

It is observed from the denoised images that soft thresholding generally yields visually more pleasing images than hard thresholding. In the case of the obtained PSNR values as well, the soft thresholding performs better. This may be because the shrinkage of the wavelet coefficients tend to better preserve the high frequency texture detail in the dead leaves target, which may otherwise be totally removed in the case of hard thresholding.

The level of decomposition used also has an effect on the obtained PSNR of the final denoised image. It is consistently seen that the PSNR attains a maximum at a level-2 decomposition, and then decreases for further higher levels. This implies that a 2-level decomposition would be ideally suited for the purpose of denoising. This characteristic is followed for all the sample images that have been processed.

In the case of the images with the added Gaussian noise, the level-dependent threshold estimation technique based on the Birge-Massart strategy gives the best PSNR values, for α with values ranging from 1.25 to 2. This property also holds in the case of the Poisson noise model, with a change in the sparsity value. In the case of the salt and pepper noise corrupted image, the universal threshold provides a better performance, and visually the denoised image has a smoothing effect with the removal of the error pixels that characterize this noise model. As real-life captured images generally follow the Gaussian noise model, higher importance is given to the results obtained in that case.

Table 10.4 shows the PSNR values for the same image that is processed with

different wavelets. The image has Gaussian noise with $mean = 0$ and $variance = 0.001$. The obtained PSNR values give an idea about the variation in denoising performance with the wavelet used.

Table 10.4: Variation in the denoised image PSNR with the wavelet used for the wavelet transform

| | | Symmlet-6 | | | | | | | |
|----------------|---|------------------|-----------------|------------------------|------------------------|------------------------|------------------------|---------------------------|---------------------------|
| Threshold used | | Universal, soft | Universal, hard | B-M, $\alpha=3$, soft | B-M, $\alpha=3$, hard | B-M, $\alpha=2$, soft | B-M, $\alpha=2$, hard | B-M, $\alpha=1.25$, soft | B-M, $\alpha=1.25$, hard |
| Level | 1 | 35.2708 | 34.7454 | 35.2995 | 34.6019 | 35.3072 | 34.0935 | 35.2451 | 33.9488 |
| | 2 | 34.6746 | 34.5881 | 35.0067 | 34.4405 | 35.6834 | 33.9666 | 35.6108 | 33.8059 |
| | 3 | 33.9865 | 34.4522 | 34.4403 | 34.3307 | 35.5273 | 33.9219 | 35.5079 | 33.7729 |
| | 4 | 33.7759 | 34.4304 | 34.2448 | 34.3100 | 35.4620 | 33.9109 | 35.4628 | 33.7644 |
| | 5 | 33.7234 | 34.4277 | 34.1909 | 34.3068 | 35.4432 | 33.9075 | 35.4486 | 33.7610 |
| | | Symmlet-8 | | | | | | | |
| Threshold used | | Universal, soft | Universal, hard | B-M, $\alpha=3$, soft | B-M, $\alpha=3$, hard | B-M, $\alpha=2$, soft | B-M, $\alpha=2$, hard | B-M, $\alpha=1.25$, soft | B-M, $\alpha=1.25$, hard |
| Level | 1 | 35.3405 | 34.7538 | 35.3626 | 34.5913 | 35.3299 | 34.0733 | 35.2519 | 33.9312 |
| | 2 | 34.7631 | 34.6067 | 35.0951 | 34.442 | 35.7028 | 33.95 | 35.5996 | 33.7981 |
| | 3 | 34.0718 | 34.4669 | 34.5328 | 34.326 | 35.5461 | 33.9075 | 35.4905 | 33.7629 |
| | 4 | 33.8659 | 34.4481 | 34.3477 | 34.3091 | 35.4866 | 33.8976 | 35.4474 | 33.7537 |
| | 5 | 33.812 | 34.446 | 34.2945 | 34.3061 | 35.4679 | 33.8946 | 35.4326 | 33.7508 |

| db-4 | | | | | | | | | |
|----------------|---|-----------------|-----------------|------------------------|------------------------|------------------------|------------------------|---------------------------|---------------------------|
| Threshold used | | Universal, soft | Universal, hard | B-M, $\alpha=3$, soft | B-M, $\alpha=3$, hard | B-M, $\alpha=2$, soft | B-M, $\alpha=2$, hard | B-M, $\alpha=1.25$, soft | B-M, $\alpha=1.25$, hard |
| Level | 1 | 35.0811 | 34.7022 | 35.1242 | 34.5853 | 35.2153 | 34.1134 | 35.1886 | 33.9798 |
| | 2 | 34.3809 | 34.4817 | 34.7161 | 34.3862 | 35.5426 | 33.9894 | 35.5598 | 33.8382 |
| | 3 | 33.6660 | 34.3219 | 34.1138 | 34.2581 | 35.3689 | 33.9419 | 35.4563 | 33.7967 |
| | 4 | 33.4564 | 34.2999 | 33.9146 | 34.2369 | 35.3032 | 33.9308 | 35.4145 | 33.7850 |
| | 5 | 33.4051 | 34.2984 | 33.8589 | 34.2333 | 35.2841 | 33.9277 | 35.4019 | 33.7823 |
| db-5 | | | | | | | | | |
| Threshold used | | Universal, soft | Universal, hard | B-M, $\alpha=3$, soft | B-M, $\alpha=3$, hard | B-M, $\alpha=2$, soft | B-M, $\alpha=2$, hard | B-M, $\alpha=1.25$, soft | B-M, $\alpha=1.25$, hard |
| Level | 1 | 35.1742 | 34.7105 | 35.2106 | 34.5825 | 35.2559 | 34.0919 | 35.2110 | 33.9586 |
| | 2 | 34.4850 | 34.4877 | 34.8233 | 34.3737 | 35.5845 | 33.9618 | 35.5658 | 33.8248 |
| | 3 | 33.7533 | 34.3282 | 34.2110 | 34.2450 | 35.3980 | 33.9167 | 35.4483 | 33.7874 |
| | 4 | 33.5486 | 34.3074 | 34.0208 | 34.2235 | 35.3362 | 33.9064 | 35.4075 | 33.7773 |
| | 5 | 33.4982 | 34.3056 | 33.9684 | 34.2218 | 35.3193 | 33.9032 | 35.3958 | 33.7748 |
| db-10 | | | | | | | | | |
| Threshold used | | Universal, soft | Universal, hard | B-M, $\alpha=3$, soft | B-M, $\alpha=3$, hard | B-M, $\alpha=2$, soft | B-M, $\alpha=2$, hard | B-M, $\alpha=1.25$, soft | B-M, $\alpha=1.25$, hard |
| Level | 1 | 35.3728 | 34.9425 | 35.3856 | 34.8332 | 35.3955 | 34.3210 | 35.3609 | 34.1610 |
| | 2 | 34.3403 | 34.5216 | 34.6744 | 34.4631 | 35.5761 | 34.1061 | 35.6423 | 33.9366 |
| | 3 | 33.4827 | 34.3222 | 33.9631 | 34.3133 | 35.3473 | 34.0414 | 35.5018 | 33.8789 |
| | 4 | 33.2510 | 34.2989 | 33.7626 | 34.2928 | 35.2825 | 34.0261 | 35.4585 | 33.8632 |

| | | | | | | | | | |
|----------------|---|------------------|-----------------|------------------------|------------------------|------------------------|------------------------|---------------------------|---------------------------|
| | 5 | 33.1876 | 34.2961 | 33.7061 | 34.2911 | 35.2644 | 34.0214 | 35.4444 | 33.8584 |
| | | Coiflet-3 | | | | | | | |
| Threshold used | | Universal, soft | Universal, hard | B-M, $\alpha=3$, soft | B-M, $\alpha=3$, hard | B-M, $\alpha=2$, soft | B-M, $\alpha=2$, hard | B-M, $\alpha=1.25$, soft | B-M, $\alpha=1.25$, hard |
| Level | 1 | 35.2750 | 34.7373 | 35.3042 | 34.5920 | 35.3067 | 34.0837 | 35.2427 | 33.9459 |
| | 2 | 34.6840 | 34.5888 | 35.0214 | 34.4446 | 35.6834 | 33.9627 | 35.6077 | 33.8145 |
| | 3 | 33.9871 | 34.4518 | 34.4569 | 34.3328 | 35.5277 | 33.9174 | 35.5015 | 33.7749 |
| | 4 | 33.7865 | 34.4338 | 34.2806 | 34.3160 | 35.4735 | 33.9071 | 35.4614 | 33.7657 |
| | 5 | 33.7322 | 34.4316 | 34.2277 | 34.3128 | 35.4552 | 33.9039 | 35.4470 | 33.7615 |
| | | Coiflet-5 | | | | | | | |
| Threshold used | | Universal, soft | Universal, hard | B-M, $\alpha=3$, soft | B-M, $\alpha=3$, hard | B-M, $\alpha=2$, soft | B-M, $\alpha=2$, hard | B-M, $\alpha=1.25$, soft | B-M, $\alpha=1.25$, hard |
| Level | 1 | 35.3841 | 34.7650 | 35.4022 | 34.6001 | 35.3463 | 34.0694 | 35.2618 | 33.9339 |
| | 2 | 34.8139 | 34.6104 | 35.1566 | 34.4335 | 35.7168 | 33.9455 | 35.5970 | 33.8010 |
| | 3 | 34.1112 | 34.4766 | 34.6082 | 34.3260 | 35.5658 | 33.9030 | 35.4819 | 33.7633 |
| | 4 | 33.9081 | 34.4567 | 34.4412 | 34.3071 | 35.5122 | 33.8904 | 35.4352 | 33.7523 |
| | 5 | 33.8537 | 34.4546 | 34.3970 | 34.3028 | 35.4969 | 33.8853 | 35.4182 | 33.7481 |

It is observed that there is not much appreciable difference in the maximum possible PSNR value on using different wavelets. Also, for the same family of wavelet, the usage of a wavelet with a higher number of vanishing moments provides a better PSNR value for the denoised image, which is observed for all 3 families of wavelets considered. In the case of the wavelets we can consider the number of vanishing moments to be an indication of the ability of the scaling function to represent more complex signals - i.e. complex functions can be represented with a sparser set

of wavelet coefficients if there are more vanishing moments. The lower PSNR may be the slight trade-off observed on reducing complexity.

Based on the obtained results in this section, the parameters for denoising of the captured dead leaves image are set. The denoised image is then used for calculating the noise PSD, and further, the dead leaves texture MTF.

10.3 Modification to the texture MTF algorithm

Utilizing the results from the previous section, we propose an algorithm for the calculation of the texture MTF using the dead leaves target. The algorithm is as given below. The power spectral density in each case is obtained by taking the Fourier transform of the image, and subsequently the power spectral density. The 2-D PSD is converted to 1-D density by radial binning averaging, which is possible due to the spherical symmetry of the dead leaves target, and its power spectral density. The wavelet thresholding parameters have been set according to the results obtained from the Section [10.2](#).

Figures [10.15\(a\)](#)-[10.15\(c\)](#) display the texture MTF results obtained for an image of the dead leaves target, with various values of added Gaussian noise. The original images have been obtained from an iOS camera system in all cases. The Figures [10.16\(a\)](#)-[10.16\(c\)](#) display the images that have been considered in this case. The mean and variance values specified are with respect to the pixel intensities of the image being normalized to within (0-1). Gaussian noise most closely models the noise model usually seen in captured real-life images. The figures display a

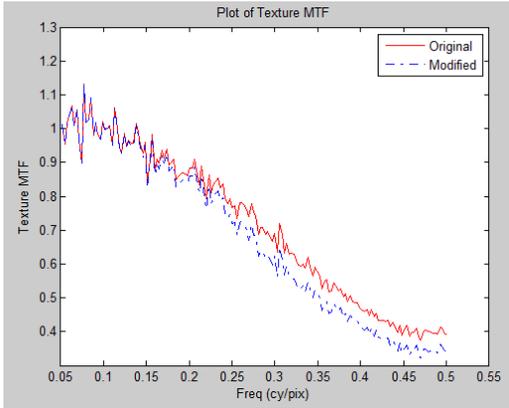
Algorithm 1: Calculation of texture MTF

- 1 Obtain multiple images of the dead leaves target. Let the number of images be n .
- 2 Perform image registration and alignment of the n copies, $I_{(1,2...n)}$.
- 3 Obtain the image average from the multiple copies, I_{avg} .
- 4 Perform wavelet thresholding on the texture region of I_{avg} to obtain the denoised version of the image, I_{th} .
- 5 Subtract the denoised image from one of the captured images (from $I_{(1,2...n)}$) to obtain the noise distribution, and obtain the corrected noise power spectral density ($PSD_{noise,corrected}$).
- 6 Calculate the spectrum of the ideal image, PSD ideal using the model equations as defined in [3].
- 7 The measured PSD can be obtained from the same image used to obtain the noise distribution.
- 8 Calculate the texture MTF according to Equation 10.13 below:

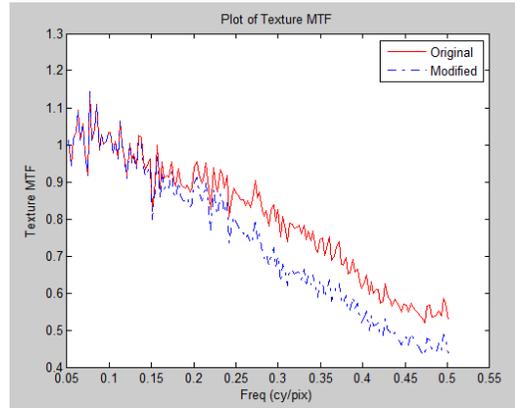
$$MTF_{texture} = \sqrt{\frac{PSD_{measured} - PSD_{noise,corrected}}{PSD_{ideal}}} \quad (10.13)$$

comparison between the modified texture MTF which includes the noise correction term, and the original MTF with no noise correction. It is observed that as the noise level increases, there is a corresponding increase in the value of the noise correction term, which results in the texture MTF being lowered. This corresponds to the visual decrease in texture reproduction quality for the noisy images. These images can be considered artificial since the noise has been added manually to the images. In the example image with the Gaussian noise of variance=0.005, it is observed that the uncorrected MTF is high due to the appearance of the noise as high frequency detail. The noise correction term thus corrects this artificial increase in texture MTF and provides a much better estimate of the actual texture MTF observed in the image.

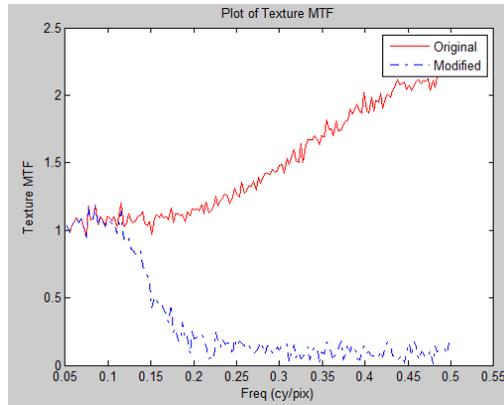
Figures 10.17(a)-10.17(d) display the texture MTF results from unaltered images of the dead leaves target, which have been obtained at various levels of exposure time and ISO speed. The texture MTF is observed to fall at mid-level frequencies for images taken at high ISO levels, which may be due to a bleaching effect which appears at high sensor sensitivities. The correction in the texture MTF due to the noise is clearly observed in each of the cases. The correction matches with the initial consideration that mostly the texture MTF at the high spatial frequencies are affected by the high-frequency noise components.



(a) Texture MTF of image with added Gaussian noise ($m=0$, $v=0.0005$).

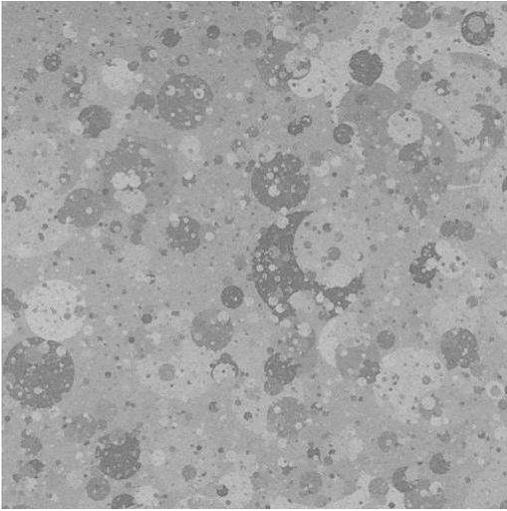


(b) Texture MTF of image with added Gaussian noise ($m=0$, $v=0.001$).

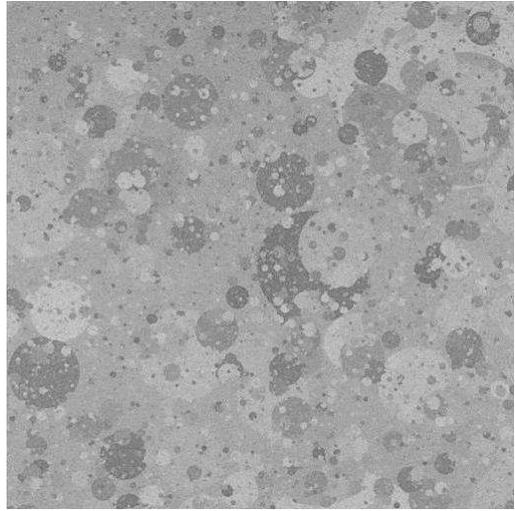


(c) Texture MTF of image with added Gaussian noise ($m=0$, $v=0.005$).

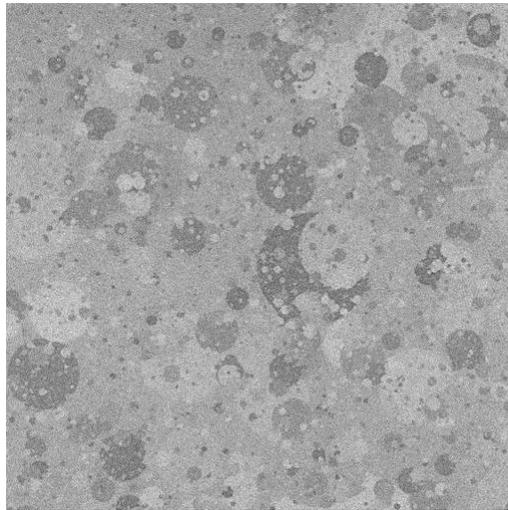
Figure 10.15: Texture MTF comparison using the proposed algorithm



(a) Dead leaves target image with added Gaussian ($m=0$, $v=0.0005$).

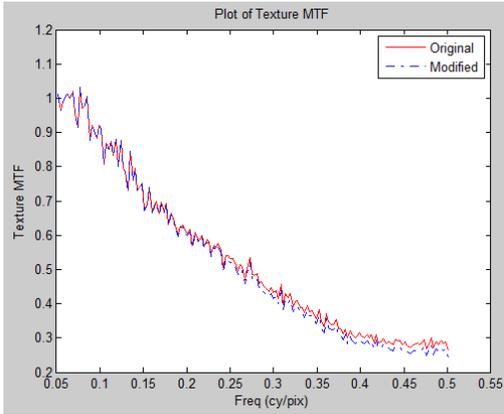


(b) Dead leaves target image with added Gaussian ($m=0$, $v=0.001$).

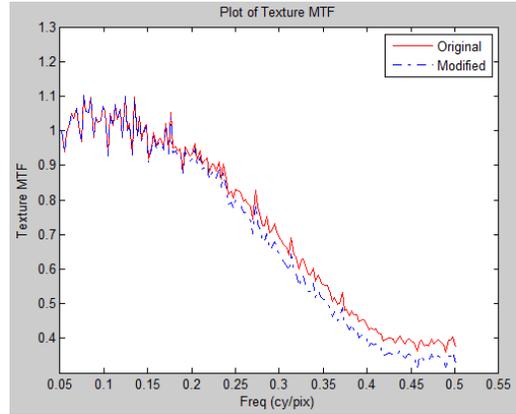


(c) Dead leaves target image with added Gaussian ($m=0$, $v=0.005$).

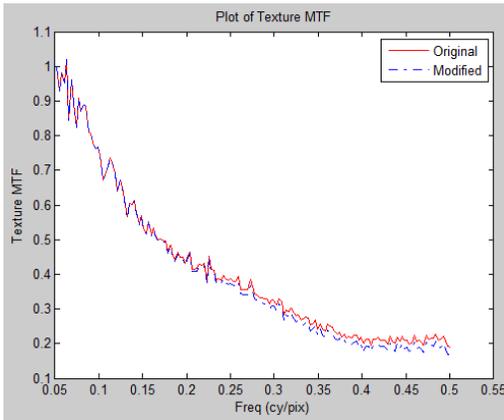
Figure 10.16: Dead leaves target image at various noise levels



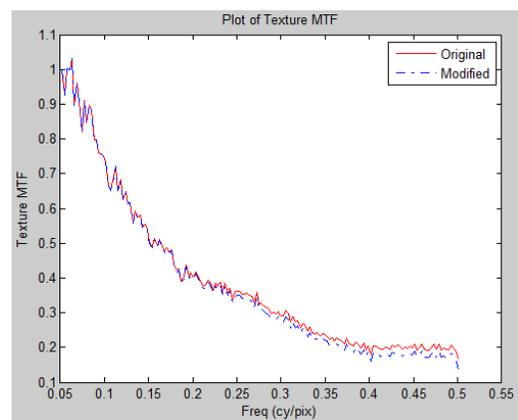
(a) Texture MTF for image captured at ISO-32, and exposure time=1/30s.



(b) Texture MTF for image captured at ISO-32, and exposure time=1/60s.



(c) Texture MTF for image captured at ISO-650, and exposure time=1/500s.



(d) Texture MTF for image captured at ISO-800, and exposure time=1/500s.

Figure 10.17: Texture MTF comparison using the proposed algorithm at various ISO levels and exposure times

Chapter 11: Conclusion

In this report we have presented information on the following topics -

11.1 Variation of the texture MTF with various parameters

The first part of the report studies the effect of several parameters that constitute the conditions under which a camera system captures an image. The target used is the dead leaves target, and the effects of these parameters on the obtained dead leaves texture are considered. The variable parameters that are considered include target-camera distance, illumination levels, exposure time during the image capture, ISO sensitivity of the camera used. The effect of the addition of noise to the dead leaves target image is also considered, along with increasing the sharpness of the image. These parameters are normally observed in the image processing cycle of a camera system, thus the results are helpful in characterizing the exact trends that become visible in these systems. Along with the $MTF_{texture}$, values of acutance are calculated in several cases as well. It is shown that the acutance values can be used as a single-value quality measure of the texture preservation quality of the camera. Appropriate conditions for the distance and illumination parameters for obtaining the target image are specified.

11.2 Modification of the calculation of the noise spectrum

The second part of the report concerns with a study of the noise spectrum calculation in the texture MTF algorithm. The estimation of the noise spectrum using the uniform gray region of the dead leaves target is shown to have some variation from the actual noise level of the texture region. This is most likely due to the camera system processing the uniform and texture regions separately. From this result the next step is in obtaining the actual noise spectrum of the texture region which is done by utilizing the wavelet thresholding denoising approach. A denoised version of the texture region is obtained, and results are provided for the PSNR results of denoised images while varying several parameters of the denoising step. Utilizing these results, the modification to the texture MTF approach are implemented, and values are obtained for images obtained under several shooting conditions.

Bibliography

- [1] Frédéric Cao, Frédéric Guichard, and Hervé Hornung. Dead leaves model for measuring texture quality on a digital camera. In *IS&T/SPIE Electronic Imaging*, pages 75370E–75370E. International Society for Optics and Photonics, 2010.
- [2] Frédéric Cao, Frederic Guichard, and Hervé Hornung. Measuring texture sharpness of a digital camera. In *IS&T/SPIE Electronic Imaging*, pages 72500H–72500H. International Society for Optics and Photonics, 2009.
- [3] Jon McElvain, Scott P Campbell, Jonathan Miller, and Elaine W Jin. Texture-based measurement of spatial frequency response using the dead leaves target: extensions, and application to real camera systems. In *IS&T/SPIE Electronic Imaging*, pages 75370D–75370D. International Society for Optics and Photonics, 2010.

- [4] Leonie Kirk, Philip Herzer, Uwe Artmann, and Dietmar Kunz. Description of texture loss using the dead leaves target: current issues and a new intrinsic approach. In *IS&T/SPIE Electronic Imaging*, pages 90230C–90230C. International Society for Optics and Photonics, 2014.
- [5] Ann B Lee, David Mumford, and Jिंगgang Huang. Occlusion models for natural images: A statistical study of a scale-invariant dead leaves model. *International Journal of Computer Vision*, 41(1-2):35–59, 2001.
- [6] P1858 - standard for camera phone image quality (cpiq), 2016.
- [7] Robert Sekuler and Randolph Blake. Perception. alfred a. *Kopf NY*, 1985.
- [8] Anil K Jain. *Fundamentals of digital image processing*. Prentice-Hall, Inc., 1989.
- [9] Paul van Walree. Distortion. *Photographic optics*. <http://toothwalker.org/optics/distortion.html>, 4, 2009.
- [10] Francis A Jenkins and Harvey E White. Fundamentals of optics 4th edition. *Fundamentals of Optics 4th edition by Francis A. Jenkins, Harvey E. White* New York, NY: McGraw-Hill Book Company, 1976, 1, 1976.
- [11] Eugene Hecht. Hecht optics. *Addison Wesley*, 997:213–214, 1998.
- [12] Dr Philippe Cattin. Image restoration: Introduction to signal and image processing. *MIAC, University of Basel*. Retrieved, 11, 2013.

- [13] John R Barry, Edward A Lee, and David G Messerschmitt. *Digital communication*. Springer Science & Business Media, 2004.
- [14] Jun Ohta and CMOS Smart. Image sensors and applications, volume 129 of optical science and engineering, 2008.
- [15] Samuel W Hasinoff. Photon, poisson noise. In *Computer Vision*, pages 608–610. Springer, 2014.
- [16] Ya M Blanter and Markus Büttiker. Shot noise in mesoscopic conductors. *Physics reports*, 336(1):1–166, 2000.
- [17] S Esakkirajan, T Veerakumar, Adabala N Subramanyam, and Prem CH Chand. Removal of high density salt and pepper noise through modified decision based unsymmetric trimmed median filter. *Signal Processing Letters, IEEE*, 18(5):287–290, 2011.
- [18] Raymond H Chan, Chung-Wa Ho, and Mila Nikolova. Salt-and-pepper noise removal by median-type noise detectors and detail-preserving regularization. *Image Processing, IEEE Transactions on*, 14(10):1479–1485, 2005.
- [19] Alan C Bovik. *Handbook of image and video processing*. Academic press, 2010.
- [20] Noura Azzabou, Nikos Paragios, and Frédéric Guichard. Uniform and textured regions separation in natural images towards mpm adaptive denoising. In *Scale Space and Variational Methods in Computer Vision*, pages 418–429. Springer, 2007.

- [21] Xinhao Liu, Mitsuru Tanaka, and Masatoshi Okutomi. Noise level estimation using weak textured patches of a single noisy image. In *Image Processing (ICIP), 2012 19th IEEE International Conference on*, pages 665–668. IEEE, 2012.
- [22] S Grace Chang, Bin Yu, and Martin Vetterli. Wavelet thresholding for multiple noisy image copies. *Image Processing, IEEE Transactions on*, 9(9):1631–1635, 2000.
- [23] Amara Graps. An introduction to wavelets. *Computational Science & Engineering, IEEE*, 2(2):50–61, 1995.
- [24] Clemens Valens. A really friendly guide to wavelets. *ed. Clemens Valens*, 1999.
- [25] Ingrid Daubechies. Orthonormal bases of compactly supported wavelets. *Communications on pure and applied mathematics*, 41(7):909–996, 1988.
- [26] Gregory Beylkin, Ronald Coifman, and Vladimir Rokhlin. Fast wavelet transforms and numerical algorithms i. *Communications on pure and applied mathematics*, 44(2):141–183, 1991.
- [27] Ingrid Daubechies et al. *Ten lectures on wavelets*, volume 61. SIAM, 1992.
- [28] Alfred Haar. Zur theorie der orthogonalen funktionensysteme. *Mathematische Annalen*, 71(1):38–53, 1911.
- [29] Wim Sweldens. Lifting scheme: a new philosophy in biorthogonal wavelet constructions. In *SPIE’s 1995 International Symposium on Optical Science, En-*

- gineering, and Instrumentation*, pages 68–79. International Society for Optics and Photonics, 1995.
- [30] David L Donoho and Jain M Johnstone. Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3):425–455, 1994.
- [31] S Grace Chang, Bin Yu, and Martin Vetterli. Image denoising via lossy compression and wavelet thresholding. In *Image Processing, 1997. Proceedings., International Conference on*, volume 1, pages 604–607. IEEE, 1997.
- [32] S Grace Chang, Bin Yu, and Martin Vetterli. Adaptive wavelet thresholding for image denoising and compression. *Image Processing, IEEE Transactions on*, 9(9):1532–1546, 2000.
- [33] Lakhwinder Kaur, Savita Gupta, and RC Chauhan. Image denoising using wavelet thresholding. In *ICVGIP*, volume 2, pages 16–18, 2002.
- [34] Stephane G Mallat. A theory for multiresolution signal decomposition: the wavelet representation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 11(7):674–693, 1989.
- [35] Adrian S Lewis and G Knowles. Image compression using the 2-d wavelet transform. *Image Processing, IEEE Transactions on*, 1(2):244–250, 1992.
- [36] Florian Luisier, Thierry Blu, and Michael Unser. A new sure approach to image denoising: Interscale orthonormal wavelet thresholding. *Image Processing, IEEE Transactions on*, 16(3):593–606, 2007.

- [37] David L Donoho and Iain M Johnstone. Adapting to unknown smoothness via wavelet shrinkage. *Journal of the american statistical association*, 90(432):1200–1224, 1995.
- [38] Lucien Birgé and Pascal Massart. Gaussian model selection. *Journal of the European Mathematical Society*, 3(3):203–268, 2001.