AUTOMATIC TESTS FOR CAMERA PERFORMANCE ANALYSIS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

ALİCAN HASARPA

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
ELECTRICAL AND ELECTRONICS ENGINEERING

JUNE 2018

Approval of the thesis:

AUTOMATIC TESTS FOR CAMERA PERFORMANCE ANALYSIS

submitted by **ALİCAN HASARPA** in partial fulfillment of the requirements for the degree of **Master of Science in Electrical and Electronics Engineering Department, Middle East Technical University** by,

Prof. Dr. Halil Kalıpçılar
Dean, Graduate School of **Natural and Applied Sciences** _____

Prof. Dr. Tolga Çiloğlu
Head of Department, **Electrical and Electronics Engineering** _____

Prof. Dr. Gözde Bozdağı Akar
Supervisor, **Electrical and Electronics Engineering, METU** _____

Examining Committee Members:

Prof. Dr. Aydın Alatan
Electrical and Electronics Eng. Dept., METU _____

Prof. Dr. Gözde Bozdağı Akar
Electrical and Electronics Eng. Dept., METU _____

Assoc. Prof. Dr. Fatih Kamışlı
Electrical and Electronics Eng. Dept., METU _____

Assoc. Prof. Dr. Ahmet Oğuz Akyüz
Computer Engineering Department, METU _____

Assoc. Prof. Dr. Aykut Erdem
Computer Engineering Department, Hacettepe University _____

Date: _____

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: ALİCAN HASARPA

Signature   :

ABSTRACT

AUTOMATIC TESTS FOR CAMERA PERFORMANCE ANALYSIS

HASARPA, ALİCAN

M.S., Department of Electrical and Electronics Engineering

Supervisor   : Prof. Dr. Gözde Bozdağı Akar

JUNE 2018, 170 pages

The camera technology is consistently improving, with high definition, smart cameras being utilized all around the world. Because of the different quality of such cameras, the camera performance analysis plays a critical role for the end users in order to determine the real difference between the available alternatives.

The image quality of a camera may be assessed visually using digitally generated test patterns in a controlled environment. The main purpose of this thesis is to automate this assessment and reduce the interaction between the user and the test software in order to prevent any operator-related errors. Although there is a various number of performance criteria about the camera performance, the most important and distinctive properties are chosen in order to evaluate both image and video quality of the test cameras. The studied tests are separated into 2 mainstreams with respect to the main focuses.

In the first set of tests, dynamic range, color accuracy and the sharpness of the cameras are assessed automatically. In the second set of tests, the video performances of the cameras are studied. In order to provide a controlled environment, a dark room

has been prepared in Multimedia Research Laboratory of Middle East Technical University. In the dark room, only two LG LED 5000K – 4000 lux flat light sources have been used. For the video performance tests, the 1D motion platform in Laboratory of Robotics and Autonomous Systems is utilized.

The test results are also compared with the results obtained by a commercially available software, Imatest.

Keywords: Image Quality, Camera Performance, Dynamic Range, Color, Sharpness, Motion

ÖZ

KAMERA BAŞARIM ANALİZİ İÇİN OTOMATİK TESTLER

HASARPA, ALİCAN

Yüksek Lisans, Elektrik ve Elektronik Mühendisliği Bölümü

Tez Yöneticisi   : Prof. Dr. Gözde Bozdağı Akar

Haziran 2018 , 170 sayfa

Kamera teknolojisi, dünyanın her yerinde yüksek çözünürlüklü akıllı kameraların kullanımıyla sürekli olarak gelişmektedir. Kamera performans analizleri, kameralar arasında kalite farkının bulunduğu dikkate alındığında, mevcut kamera seçenekleri arasındaki gerçek farklılıkların belirlenmesinde son kullanıcılar için kritik bir rol oynamaktadır.

Bir kameranın görüntü kalitesi, kontrollü bir ortamda dijital olarak oluşturulmuş test örüntüleri kullanılarak görsel olarak değerlendirilmektedir. Bu tezin ana amacı, operatör kaynaklı herhangi bir hatayı önlemek için bu değerlendirmeyi otomatikleştirmek ve kullanıcı ile test yazılımı arasındaki etkileşimi azaltmaktır. Kamera performansının belirlenmesiyle ilgili çeşitli performans ölçütleri bulunmakta olsa da test kameralarının hem görüntü hem de video kalitesini değerlendirebilmek için kamera performansını etkileyen en önemli ve ayırt edici özellikler seçilmiştir. Üzerine çalışılan testler, gerçek amaçları dikkate alınarak 2 ana gruba ayrılmıştır.

İlk test dizisinde, dinamik aralık, renk doğruluğu ve kameraların netliği otomatik olarak değerlendirilmektedir. İkinci test dizisinde ise kameraların video performansları

incelenmiştir. Kontrollü bir ortam sağlamak için Orta Doğu Teknik Üniversitesi Multimedya Araştırma Laboratuvarı'nda karanlık bir oda hazırlanmış olup bu karanlık odada sadece iki adet LG LED 5000K - 4000 lüks düz ışık kaynağı yer almıştır. Video performans testleri için ise Robot ve Otonom Sistemler Laboratuarında 1D hareket ortamı kullanılmıştır.

Ayrıca test sonuçları, piyasada bulunan Imatest yazılımı kullanılarak elde edilen sonuçlarla da karşılaştırılmaktadır.

Anahtar Kelimeler: Görüntü Kalitesi, Kamera Performansı, Dinamik Aralık, Renk, Keskinlik, Hareket

To my family

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

OECF	Opto-Electronic Conversion Function
MTF	Modulation Transfer Function
CCD	Charge-Coupled Device
CMOS	Complementary Metal Oxide Semiconductor
SFR	Spatial Frequency Response
ADC	Analog to Digital Converter
DR	Dynamic Range
SNR	Signal to Noise Ratio
LED	Light Emitting Diode
CHT	Circular Hough Transform
CIE	International Commission on Illumination
DC	Direct Current
ISO	International Organization of Standardization
SQF	Subjective Quality Factor
CPIQ	Camera Phone Image Quality Group
CSF	Contrast Sensitivity Function
USAF	United States Air Force
CIPA	Standard of Camera & Imaging Products Association
PS	Power Spectrum
PWM	Pulse-Width Modulation
CDF	Cumulative Distribution Function

CHAPTER 1

INTRODUCTION

1.1   Introduction to Camera Performance Tests

The visual system is one of the most interesting systems of the human body and thus, it has been researched in different areas of the science like biology, medicine, physics and engineering since the first days of the humankind. Even though the studies on this system have been conducted for such a long time, the whole process performed in the human eye and brain still could not be fully understood. On the other hand, all the information obtained about the human visual system provided the necessary ideas for producing a replica of the system which is called camera.

The analogies between the cornea of the human eye and the lens of the camera, the pupil of the eye and the aperture of the camera, the retina of the eye and the film of the camera were mainly used in order to create the first prototypes of the cameras in the history. Until the 1950s, image acquisition was realized using stationary ancient devices of analog capturing technology. With the gradual shift from analog to digital and the revealing of the working principles of sensor cells in the human eye, the manufacture of captures with smaller dimensions became possible and this resulted in increasingly handy and mobile cameras. The dramatic increase in their production, purchase, and usage coincides with the worldwide utilization of the low-cost CCD and CMOS technologies, which play an essential role in the conversion of incident light to electrical data during capture. From then on, the affordable camera manufacture made the cameras indispensable for our daily lives. Due to the decrease in the prices of the high performance microprocessors and the understanding of the processing functions of human brain in the visual system contributed thoroughly, to the

blossoming of brand-new application areas that require digital camera usage such as security and surveillance systems, automation and quality control systems, medical applications, transportation and traffic control. Although today, it is accepted that the acquired images or videos by using an average camera are much closer to human vision than the ones years ago, the perfect performance matching is not possible because of different reasons. As mentioned above, since all the details of the visual system have not been covered yet, the progress in the camera industry has not reached to the end and the development will continue till the cameras provide a perfect match with the human visual system.

The increase in the production, purchase, and usage of the cameras results with the increase in both the number of the camera manufacturers and the number of the different models in the market. Hence the development of the industry is fast, the properties and the capabilities of the cameras are changing very quickly and the different models become outdated rapidly. In such case, it is almost impossible to choose the camera which satisfies both the requirements of the system and the budget price. Although the human visual system is one of the most complicated systems of the human body, the datasheets of the cameras do not provide too much information about the performance of the camera. The main reason for this is the lack of the totally accepted standards about the true indicators of the performance of the camera. Hence, in the existence of too many alternatives from different manufacturers, choosing the right camera becomes harder and harder according to the application. Thus, the researchers from all around the world work on the different characteristics of the cameras in order to provide metrics to evaluate the camera performances. Although some of these test procedures are accepted like standards such as ISO 12233, ISO 14524, ISO 15739; most of them are still arguable. The most frequently studied properties of the digital cameras can be listed as: [51]:

- Dynamic Range
- White Balancing
- Signal to Noise Ratio
- Sharpness
- Lens Distortion
- Shading

- Color Accuracy
- Shutter Lag
- Chromatic Aberration
- Uniformity
- Veiling Glare

It is crucial to note that for each of the properties of the camera performance mentioned above requires different test setups, test charts and measuring algorithms or procedures. Even for the same property, different evaluation methods have been proposed and these methods might show some varieties. The researchers, working groups and camera performance measuring companies prefer to use their own test setups, test charts, and test procedures in order to grade the cameras because there is no clearly defined standard to measure most of them. Although the differences are quite small for the most common performance tests like color accuracy or sharpness; for some criteria like chromatic aberration and shutter lag, different test setups and charts are used by the different experimenters. Because of the indicated differences, it is hard to compare the results of the different groups even if they are aimed to measure the same performance parameter of the camera. The key idea which is followed by people concerned is to compare the performance of the different cameras objectively. Thus, the same test setup and the test chart in fixed conditions are provided for all the alternative cameras and evaluate the same property in the outputs in order to compare them. Due to the fact that the only change in the test setup is the camera under test, the difference in the image or the video indicates the camera performance. Here, the main purpose is to show the abilities of the cameras for different conditions in order to lead the end-user or the customer to choose the right product for their requirements. Thus, the test setup, test chart, and the test procedure can be changed according to the application type because the performance of the cameras can differ for the different environments. Due to this fact, the custom test conditions should be prepared most of the time and obviously, these test results are more realistic about the camera performance rather than the theoretical information provided in the datasheets of the cameras.

Designing the application for specific test setups, deciding the right test charts in the market or designing them if necessary and preparing the test procedure with evaluation algorithm and the software require lots of background information and great effort. Thus, camera performance measuring companies provide services for both the end users and the camera manufacturers. Imatest GmbH and Image-Engineering GmbH are the biggest companies in this sector which can provide a packet solution with the test setup, test chart, and the evaluation software. Although the best way is to perform the camera performance tests in the application specific conditions, a classical test setup has been used for most of the measurements by these companies when the details of the application are not well defined. The test setup which is used by Imatest GmbH is given in Figure 1.1.
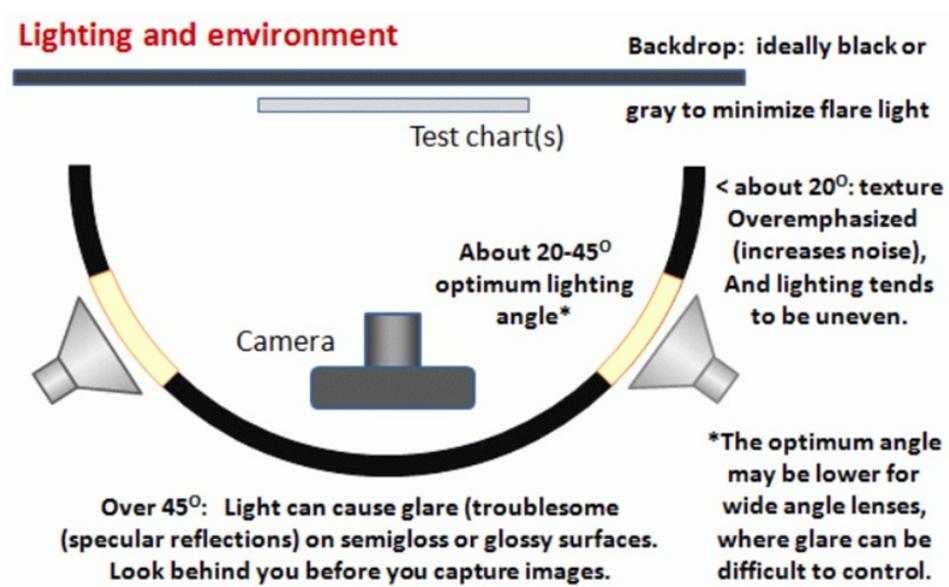
Figure 1.1: The Test Setup Proposed by Imatest GmbH [1]

As can be seen, the camera is placed perpendicularly in front of the test chart. Two light sources are placed nearly on the diagonals such that they can illuminate the test chart homogenously. The test setup which is used by Image-Engineering GmbH is also given in Figure 1.2. The same ideas are accepted and thus, the test setup has too many similarities.
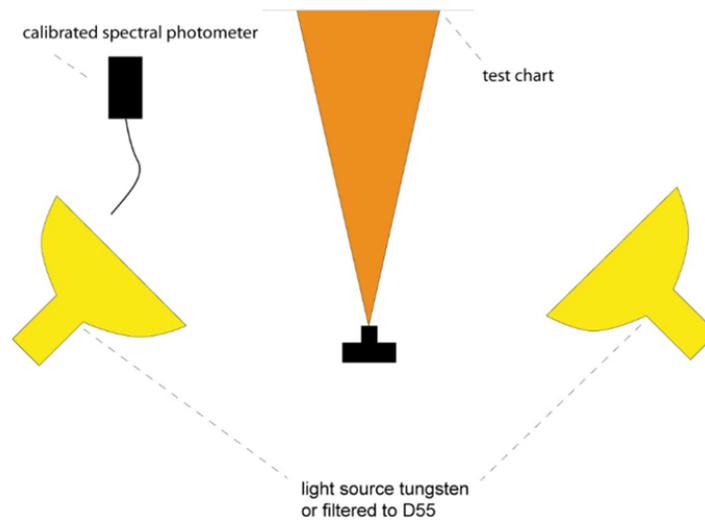
Figure 1.2: The Test Setup Proposed by Image-Engineering GmbH [2]

1.2 Scope and Outline

In the scope of this thesis, the most critical performance criteria of the digital cameras such as dynamic range, color accuracy, color constancy, sharpness, change sensitivity and motion blur are discussed. In order to measure the performances of the cameras, well-defined test procedures based on the standards are proposed. Throughout the work, the main concern is to minimize the user interface in order to reduce the operator related errors while designing and preparing the test setups and the test charts if it is necessary. Also, the test procedures and their implemented form do not require any parameter updates for different cameras under test. In other words, the methods are prepared such that they are applicable for all the cameras in the market. In order to reveal the generalization achieved, the eight different cameras are tested in the darkroom which prepared in Multimedia Research Laboratory at METU and their results are compared for each test. Due to the fact that the test setups are constant and the evaluation algorithms are camera independent, the differences between the output images and videos indicate the camera performances. Using directly the output of the cameras instead of the raw data provides to include all the effects of post-processing application like demosaicing algorithm, gamma correction, unsharp masking, noise reduction, data conversion, compression and codec performance into the test result. Thus, it is consistent to say that the measurement tests evaluate the

cameras as a system from the beginning to the end, with all the imperfections.

In the second chapter, the dynamic range property of the digital camera is discussed and using a reflective gray ramp chart is explained in order to determine the high dynamic range of the camera performance. In the third chapter, the color performance of the camera is considered. The color accuracy measurement is one of the most commonly used properties and its measurement procedure is well defined when the test chart is chosen. In our work, the procedure is automated for color accuracy such that no user interaction is needed. Also, by using the results of the color accuracy test, the color correction matrices are prepared and results are discussed. Lastly, the importance of the color constancy is emphasized and the related test results are given. In the fourth chapter, the sharpness performances of the cameras are covered. Although there are different methods to measure the sharpness of the camera, there exists a consensus about each measurement method. The slanted edge method is chosen and like Chapter 3, the measurement of sharpness is automated. Although the still image performances of the camera are examined in the previous three chapters, video performances of the camera are studied in the next two chapters. A test method is proposed in Chapter Five for comparison of the change detection performance of the cameras. For this purpose, a test setup which performs the same scenario for each camera is prepared and a change detection algorithm is used in order to measure the difference. The change in this test setup is created by altering the illumination conditions such that the motion blur does not affect the test results. On the other hand, in the last test, motion blur measurement is performed. A test setup is designed such that the test chart can be moved from left to the right in the scene of the camera. The gradient-based motion blur measurement method is explained in Chapter Six. All the results of the tests are given at the end of each chapter. Lastly, all the test results and the future work are gathered in the last chapter.

CHAPTER 2

CAMERA TESTS ON DYNAMIC RANGE AND RESULTS

2.1 Background Information about Dynamic Range

The dynamic range is one of the most common criterions to be used when comparing the performance of the cameras is considered. Basically, dynamic range shows the scale that the camera can cover from its whitest white to its blackest black. In order to understand the concepts of the whitest white and the blackest black, the image acquisition system of the digital cameras should be revised.
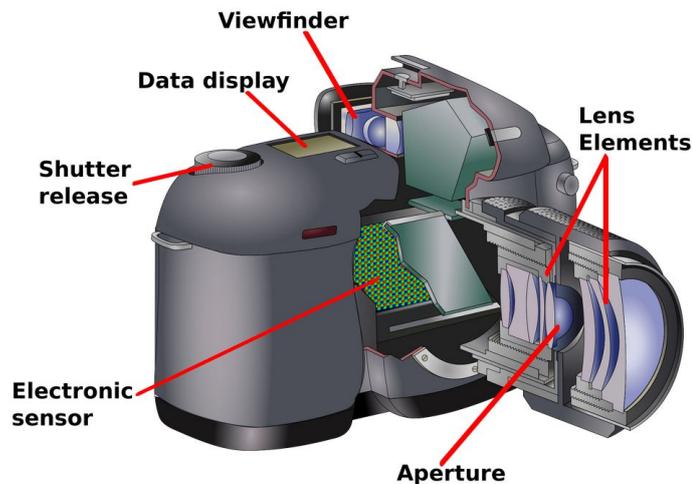
Figure 2.1: The Common Parts of Digital Cameras [3]

Even there exist vast diversity among the manufactured cameras around the world as mentioned in the previous section; all digital cameras are composed of the parts which are shown in Figure 2.1. The image acquisition process starts on the lens elements. First, the light passes through a lens and then the aperture before reaching to the electronic sensor. The amount of the light which passes through the aperture

depends on the aperture diameter and shutter speed. The most vital part of the camera is the electronic sensor because it is the place where light is converted into electrical data. Most of the cameras in the market use charge-coupled devices (CCD) while obtaining the image whereas the rest utilizes complementary metal oxide semiconductor (CMOS) technology. Although there are differences between these two technologies, the basic idea is the same.

For both CCD and CMOS technologies, the light is converted into electricity by using a cavity or well which is called photosite. When the light hits a photosite, the electrons in the sensor are energized. The number of the energized electrons depends on the energy of the light and the potential of the photosite increases accordingly. As mentioned in [4], the photosites can be thought as "buckets of electrons" in the very basic level. If the number of electrons in the photosite is few, as shown in the left of Figure 2.2, then the corresponding electrical potential becomes less and the corresponding pixel is represented as close to black level. If the photosite has more energized electrons, as shown in the middle of Figure 2.2, then the potential increases to a higher value and the corresponding pixel becomes whiter.
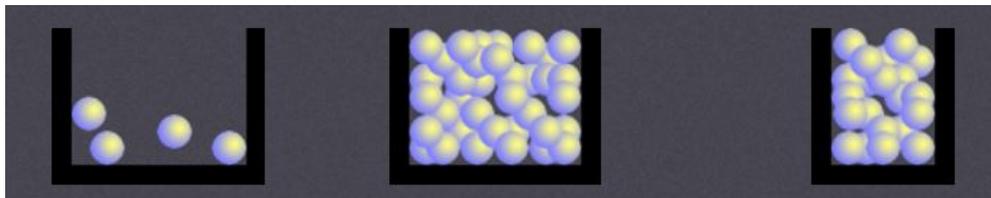
Figure 2.2: The Representation of Photosites for Different Cases [4]

Although theoretically it is easy to linearly map the number of photosites into the pixel values, in reality, some physical problems show their effects for two different cases. The first case defines the concept of the blackest black. Due to the thermal vibrations, electrical leakage and noise, neither the number of energized electrons in the photosite nor the electrical potential can be zero. Hence, the blackest level that the image has becomes limited according to these effects. As mentioned in [52], the blackest black level can be found by acquiring the images from the camera in a dark room at 25 degrees by covering the camera aperture with black cloth. As expected, the image will be dominated by the noise and the mean value of the pixels can be assumed as blackest black value.

The second case defines the concept of the whitest white. In the case that the incident light has just enough energy to energize all the electrons in the photosite, the camera reaches its maximum potential. Thus, this value becomes the whitest white level of the camera. It is important to notice that for all the light beams have more energy than this level will saturate the photosite and would be indistinguishable from the whitest white value.

As explained above, if the color information ignored and it is assumed that the camera is acquiring the grayscale images, then the meaningful areas in the image are being represented with the range between blackest black and the whitest white. For the dynamic range, one of the most important factors is the capacity of the photosite. As shown in Figure 2.2, the whitest white values of the photosites in the middle and on the right are different. If the size of the photosite increases, then the maximum number of electrons that it can measure also increases. However, there is a trade-off. If the manufacturer increases the width or/and length of the photosites, the number of the photosites in the sensor decreases in order to fit the same size. Because of the less number of photosites, the camera would have less pixel value which also deteriorates the image quality. On the other hand, increasing the depth of the photosite is the possibility but this solution is limited by the current semiconductor technology we have at the present time. Moreover, this information is generally confidential and the camera manufacturers generally do not share them with the end-users.

The second factor which determines the dynamic range of the camera is the analog to digital converter which is used in the electronic sensor. The electrical potentials of the photosites are being measured by using analog to digital converters. After clarifying the blackest black and the whitest white concepts, it is crucial to state that dynamic range does not only depends on these two values. Interpreting all the data between these two values is also important. Thus, the number of bits that the ADC can measure affects the dynamic range of the camera. According to the studies of Ching Man [53], the dynamic range of the ADC can be calculated with the formula,

$$DR = 6.02 * N + 1.76 \tag{2.1}$$

Although the formula above shows that mathematically dynamic range depends on only the number of bits in the ADC, explaining the details would bring more insight

about it. For example, theoretically, a photosite can has 5 volts when it is saturated and 0 volts when there is no light. If the camera has 12-bit ADC, then it has a resolution of 4095 by the formula of $(2^n - 1)$. Thus, the voltage step size would be equal to 5V/4095 which is equal to 1.22 millivolts. The dynamic range can be calculated as 20*log10(5V/1.22mV) which is equal to 72.25 dB . The same result can be also calculated by the formula above as (6.02*12) +1.76 which is equal to 74 dB. It is crucial to note that 1.22 millivolts are the limit that the camera can distinguish between two lights which hit the photosite. In other words, if the energy difference between two lights is not greater than the calculated value, then the pixel values in the image would be the same even if they were different in real life. Hence, having more bits in the ADC would result in the higher dynamic range. Actually, this is the only value about the dynamic range which the end users can find the datasheet of the cameras. Nowadays, most of the manufacturers use 12 or 16 bit ADC in their cameras. Thus, in the catalogs, consumers can only compare the numbers of 74 dB or 96 dB which are simply the number of bits of the ADCs in the camera.

As explained above, just checking the number of bits in the ADC cannot be enough to compare the wide dynamic range of the cameras. Due to the lack of the consensus about how to calculate the dynamic range of the camera, a test is proposed and implemented in this thesis. The proposed test chart, the test setup and the implementation details with the results of 8 sample cameras will be given in the next parts. In this test, 21-step reflective gray ramp test chart is used. The main purpose of this test is determining the distinguishability of the different sections on a test chart as discussed before in the studies of Wong and Lu [54]. Although using a reflective test chart results with some differences between the work of Wong and Lu, the idea behind is the same. Calculating the difference between the theoretical values and the mean values in the image and the noise levels in each segment give a good idea of the dynamic range of the camera. With this test, we aimed to create a system that can be used in order to measure objectively the dynamic range of the camera so that it can be used for comparing the different alternatives in the market.

2.2   Test Chart

In order to measure the dynamic performance of the cameras, there are a great number of different test charts which can be found in the literature. However, they can be grouped into two mainstreams as reflective ones and transparent ones. Before explaining the differences between these two, it is crucial to note that in both groups, several test charts with the various number of test patches in different arrangements are available in the market. These differences are based on the manufacturers of the test charts and try to extend the measurements with more stable results. In general, more test patch in the test chart provides wider dynamic range measurement.

The first difference between the reflective and the transparent chart is the test setup requirements. For transparent charts, the light source should be placed in front of the camera such that light can pass through the test chart and reach to the camera. For reflective charts, the light source is placed behind the camera such that it can properly and homogeneously illuminate the test chart. The hard part of using a transparent chart is preventing the leakage of the light from the edges of the test chart which can reach to the camera directly from the light source. In order to provide the conditions for the transparent chart, special light boxes are used.

The second difference between the transparent and reflective charts can be counted as the evaluation of the dynamic range. By using a lightbox and a transparent chart, Opto-Electronic Conversion Function (OECF) of the camera can be obtained. Due to the fact the transmissivity of the transparent patches are known and the intensity of the light provided by the lightbox is controllable; it is possible to obtain a graph of luminance values versus the digital values obtained by the camera. This graph provides an idea of the dynamic range of the camera. Higher the difference between the minimum and maximum distinguishable luminance values indicates the wider dynamic range. An OECF graph of a high-dynamic range camera is given in Figure 2.3 as an example. However, it is important to note that the pixel values should be provided as raw data from the camera.

The dynamic range studies which use reflective chart can be grouped into two. In the first method found frequently in the literature, the SNR calculations are performed
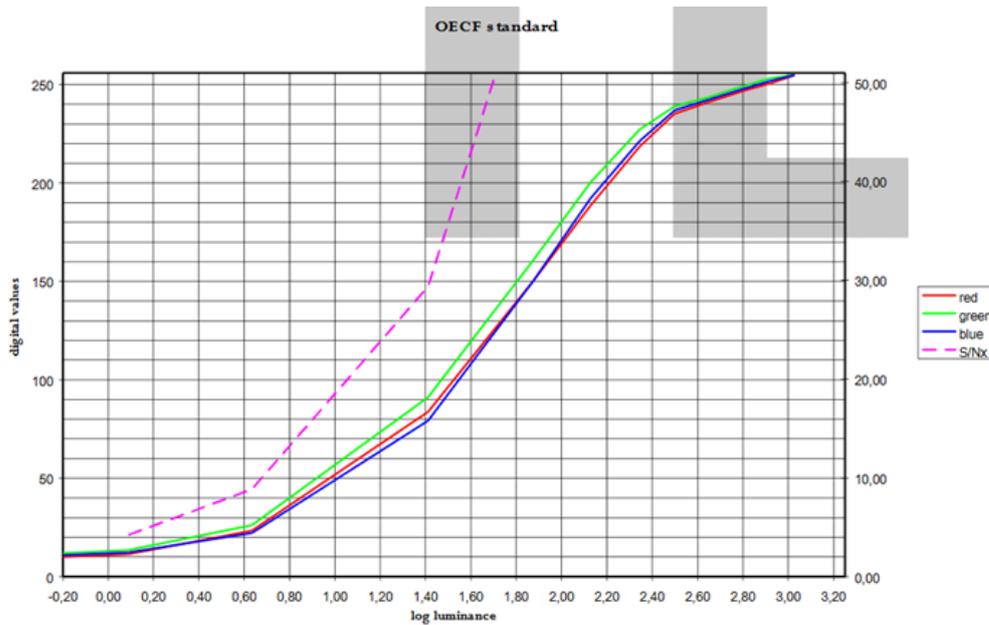
Figure 2.3: An OECF Graph Example of a High-Dynamic Range [2]

in each region till getting the darkest region with SNR=1. In the second method, the distinguishability concept is used. The human visual system is more sensitive to relative change. So, the sharp changes between the regions look clearer than the gradual passes. In the light of this information, the distinguishability of the different gray patches gives good information about the dynamic range of the camera instead of the SNR calculations.

Although using transparent chart can be useful in order to reach OECF of the camera and then observe the dynamic range; using reflective chart is adequate when the camera performances are compared with each other [55]. Also, unavailability of the raw data directly from the camera makes the dynamic range test with reflective chart more applicable for each camera. It is also crucial to note that due to the differences in the test setups between reflective and transparent charts, the reflective test charts are easy to use and cheap to implement. Lastly, our test setup is fixed for four tests and thus, using a reflective chart is more suitable.

As mentioned before, for the dynamic range measuring tests, 21-step gray ramp reflective chart has been proposed. For comparing the tone values and finding the right values for exposure, the products of KODAK Company are widely used for years [7].

Q13 or Q14 Gray Scales are the most common ones which are shown in Figure 2.4. They have 20 steps from white the black with 0.1 increments and gray background. On the other hand, Stouffer Industries produces 12, 14, 21 and 37 step gray charts as testing equipment.

Figure 2.4: Kodak (left) and Stouffer (right) Gray Ramp Test Charts [5] and [6]

Also, there are many other proposed test charts in the literature as shown in Figure 2.5 with nonlinear step sizes or different region arrangements because there is no standardized method to measure the dynamic range of the camera. Furthermore, the testing companies try to unify their test charts in order to use the same chart for testing other features of the cameras. This trend increases the number of charts that can be used for dynamic range measurement in the literature.
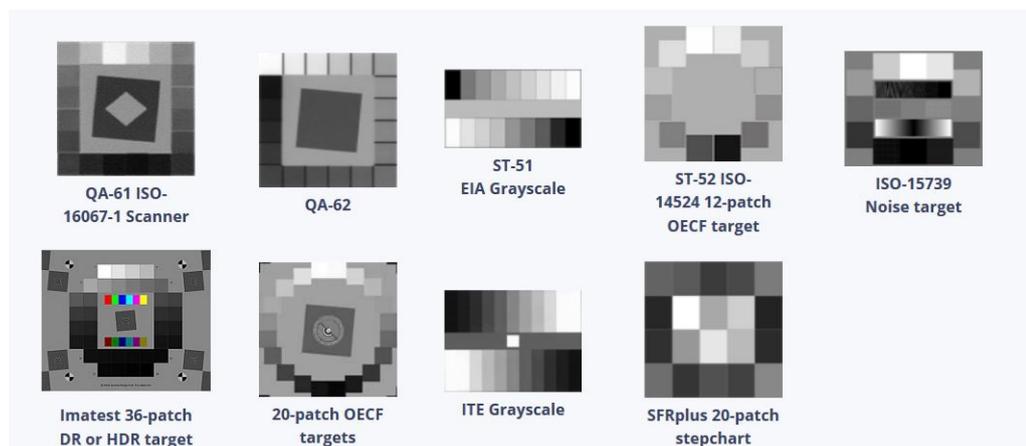
Figure 2.5: Other Test Charts [7]

Having more regions in the test chart provides to measure the higher dynamic range. The most of the test charts in the literature has 20 or fewer regions with the unsuitable placement of regions for distinguishability checking. Just 2 test charts in the market

have 36 and 37 regions from the companies Imatest and Stouffer, respectively. However, while the test chart from Stouffer is too small for our test setup, the test chart from Imatest has unsuitable distribution for distinguishability checking and too many details on it which makes the chart unsuitable for the automatic test procedure. The proposed test chart has 21 steps in the form of a gray ramp with fixed step size as mentioned in the studies of Wong and Lu. In order to determine the orientation of the test chart, four circles have been placed on the corners of the chart. The proposed chart can be seen in Figure 2.6.

Figure 2.6: The Test Chart for the Dynamic Range Tests

2.3 Test Setup

Since the first days of fundamentals of science were started to be determined, all of the hypothesis have relied on the controlled experiments. Throughout this research, all the experiments have been conducted by following this idea. In all of the cases, the environmental effects were tried to be minimized and cameras were the only dependent variable. In order to provide a controlled environment, a dark room has been prepared in Multimedia Research Laboratory of Middle East Technical University. In the dark room, only two LG LED 5000K – 4000 lux flat light sources (Manufacturer Part No: F4050TC1N5A. ALWFT00) have been used. The camera under test has

been placed 45 cm away from the test chart and light sources have been located 2 meters away from the test chart by making the 45 degrees with the line of camera and test chart. Figure 2.7 would clarify the test setup with numbers. Also, Figure 2.8 shows the real experiment setup. The illumination on the test chart is measured by using lux meter and checked whether it is uniform or not. In the test setup, 1000 ± 100 lux value is provided.
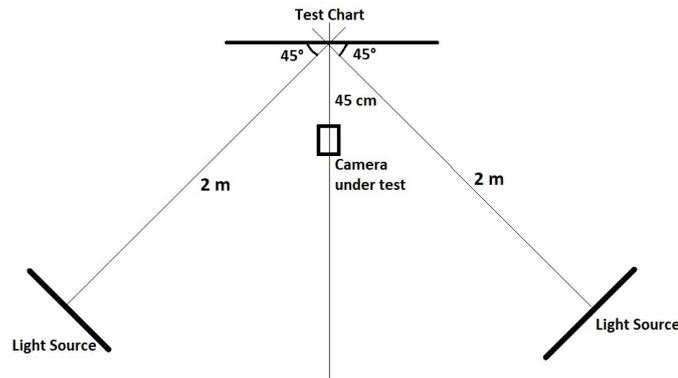
Figure 2.7: Test Setup (not to scale)

Figure 2.8: Real Test Setup

2.4 Test Details

The first part of the test is the image acquisition from the camera. The main purpose of the proposed tests is to be used for all the cameras in the market in order to compare their performance for the different features. On the other hand, the resolution of the cameras shows great variety. In order to overcome this problem, resizing the input image from its original size into the new size which has 720 pixels in width without ratio distortion is used. As explained before, dynamic range is related with white and

black levels that the camera can provide. Thus, the acquired image is converted into grayscale version from R*G*B color space by the following formula [56].

$$Gray = 0.299 * Red + 0.587 * Green + 0.114 * Blue \qquad (2.2)$$

As mentioned before, in order to find the orientation of the test chart, the circles on the corners of the test chart are used. In order to find the circle locations, Circular Hough Transform (CHT) algorithm is used which will be shortly explained in the Appendix part. After finding the corner points, the projective spatial transformation has been performed on the entire image by taking the corner points as the reference. The details of the projective spatial transformation can be found in Appendix part. One of the original images and the result of spatial transformation can be seen in left and middle of Figure 2.9. Since the chart is put into a rectangular form, the rest of the image can be discarded. The Circular Hough Transform is again used in order to find the new locations of the corners in the processed image. This time, the locations of the corners are used in order to crop the image such that only 21-step gray ramp has been obtained as a region of interest. It is important to note that after these image processing operations, the borders of the different gray regions are known by the software. It makes possible to determine whether the gray regions are distinguishable or not. The cropped region of interest for the original image in Figure 2.9 can be observed in the right of Figure 2.9. After obtaining the region of interest, the exact border points are known by the software by simply dividing the length of the region of interest by 21 which is the number of gray steps. This value is used in the next step as the theoretical distance, "D" , between the borders.

Figure 2.9: The Preprocessing the Original Image

Due to the existence of the circles, the region of interest image would have unwanted regions for the processing. Ten percent of the width is excluded from both top and bottom in the rest of the operations. Then, for each one of three row, the gradient

is calculated with the formula below, where "G(i,n)" is the gradient value in the row "i" and column "n" and "I" is the image of region of interest. Using one of three row increases the noise immunity that can partly be formed in a specific region in the image.

$$G(i, n) = I(i, n + 1) - I(i, n) \tag{2.3}$$

In the perfect case, the gradient vector "G" for each row would have zero value in the middle pixels of all the regions and nonzero values only at the borders between the steps. However, in the real case, due to the noise, some extra peaks are observed. On the other hand, if the peak is not repetitive in most of the "G"s, for different rows, then it shows that this peak is not a boundary point but just the effect of the noise. With this idea, the peaks are determined for each row. While determining the peak locations, the information on the distance between the border points is also used. The 70 percent of the theoretical distance between the peaks is considered as a threshold. The algorithm is performed with the steps below:

• The global peak is selected as a start point by assuming that it is a border because simple noise cannot form the global peak

• All the peaks with 70 percent of the theoretical distance from that peak are discarded.

• The procedure is performed iteratively until there is no more peak.

Algorithm 1 Algorithm for Candidate Border Point Detection

for each row, i **do**

$\quad [location, value] = max(G(i, n))$

$\quad$ **while** $value \mathrel{!}= 0$ **do**

$\quad\quad P_{candidate} = location$

$\quad\quad$ **for** $j = (P_{candidate} - (0.7 * D))$ to $(P_{candidate} + (0.7 * D))$ **do**

$\quad\quad\quad G(i, j) = 0$

$\quad\quad$ **end for**

$\quad\quad$ save $P_{candidate,i}$

$\quad\quad [location, value] = max(G(i, n))$

$\quad$ **end while**

end for

After locating the possible borders for each row, the repetitiveness of the peaks is controlled. By using a simple thresholding, the border points are chosen. The 20 percent of the number of rows is used as the threshold value for the beginning of an iterative processing. If a point and its neighborhood of 3 pixels are marked as a border point for the 20 percent of the rows, then this point is considered as a border point.

Algorithm 2 Algorithm for Repetitiveness of Candidate Points

---

    for each column, n **do**

        $search\_key = 0$

        for each row, i **do**

            if $(n - 3, n - 2, n - 1, n, n + 1, n + 2, n + 3) \in P_{candidate,i}$ **then**

                $search\_key = search\_key + 1$

            end if

        end for

    end for

As the next step, the distances between the determined border points are again checked. So, even the information on the theoretical distance between the borders is used while finding the possible border points, there might be some extra points which are chosen as borders. As like mentioned above, 70 percent of the distance between two border points is checked and if an extra point is determined by the previous step then it is considered as the wrong guess. For the next iteration, the threshold value of 20 percent of number of the rows is incremented 1 and the same procedure is performed till the guess does not have 2 border points which are closer to each other than 70 percent of the theoretical distance as shown in Figure 2.10.
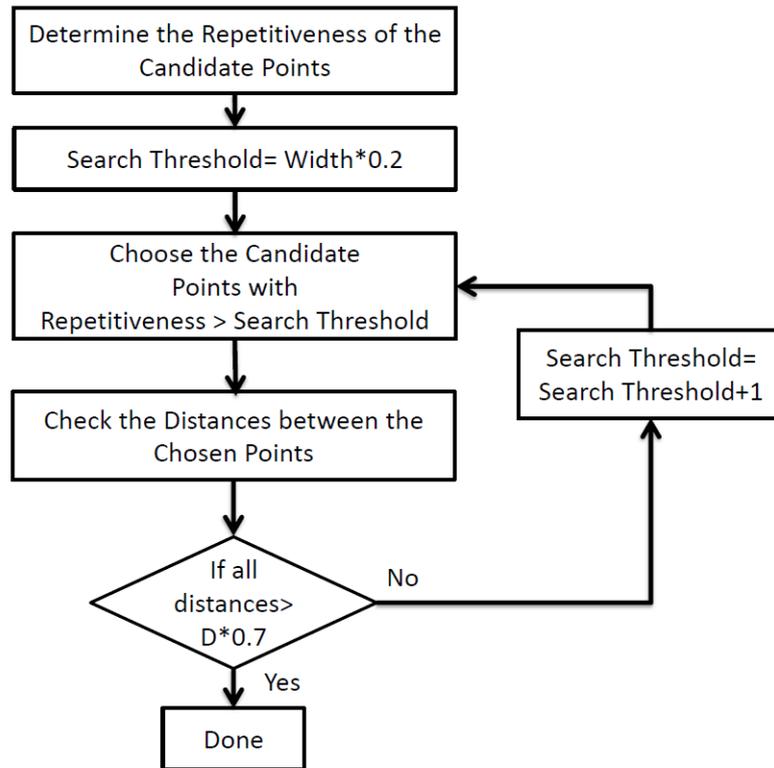
Figure 2.10: Flowchart of the Final Border Point Selection

As the last step, according to the determined border points, the region of interest is cropped in order to obtain the distinguishable steps. The rest of the test steps are only performed on these regions. For each region, the histogram is calculated and the maximum point as taken as the mean value of the region. The difference between mean value and the theoretical pixel value for that region gives a good numerical idea about the dynamic range of the camera. The less difference between the theoretical value shows that the camera has better dynamic range. Also, the variance between the pixels in the same region should be expected less for a good camera performance. Even though it is not directly related to the dynamic range, it is taken into account for the total camera performance criteria.

2.5 Test Results

The explained test procedure in the previous section has been used in order to test eight different cameras in the same test setup. The results of the tests will be given in

this part. The red lines in the histograms show the theoretical values according to the original test chart. The absolute values of the difference between the maximum value of the histogram and the related theoretical value are recorded and their mean value is used for performance comparison. The mean value, μ, is calculated by the following formula where "N" is the number of the distinguishable regions, "I_i" is the pixel level chosen from the histogram in the each region, "i", and "T_i" is the corresponding theoretical pixel level in the region "i".

$$\mu = \frac{1}{N} \sum_{i=1}^{N} |I_i - T_i| \tag{2.4}$$

Moreover, for each distinguishable region, the standard deviation is calculated in order to have an idea about the noise level. For this calculation firstly, the mean pixel value, μ_{region} is calculated with the following formula where "M" is the number of the pixels in the distinguishable region and "I_m" is the pixel values in this region.

$$\mu_{region} = \frac{1}{M} \sum_{m=1}^{M} I_m \tag{2.5}$$

Then, the standard deviation value, "S", is calculated as shown.

$$S_m = \sqrt{\frac{1}{M-1} \sum_{m=1}^{M} (|I_m - \mu_{region}|)^2} \tag{2.6}$$

The mean value of these standard deviation values, σ, is also used as a performance indicator. For both μ and σ, the lower values are desirable.

$$\sigma = \frac{1}{N} \sum_{i=1}^{N} S_m \tag{2.7}$$

2.5.1 The results of Camera-1

The original image taken by the Camera-1 can be seen on the left side in Figure 2.11. By using the algorithm above, the distinguishable regions are determined as shown in Figure 2.11 on the right side. Camera-1 has 18 distinguishable steps out of 21.
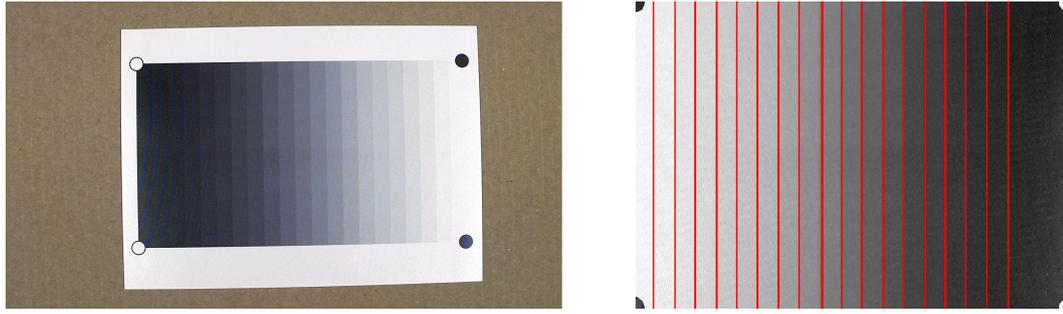
Figure 2.11: Original Image from Camera-1 (left) and Distinguishable Regions (right)

2.5.2 The results of Camera-2

The original image taken by the camera-2 can be seen on the left side in Figure 2.12. By using the same code and test setup, the distinguishable regions are found as shown in Figure 2.12 on the right side. Camera-2 has 19 distinguishable steps out of 21.
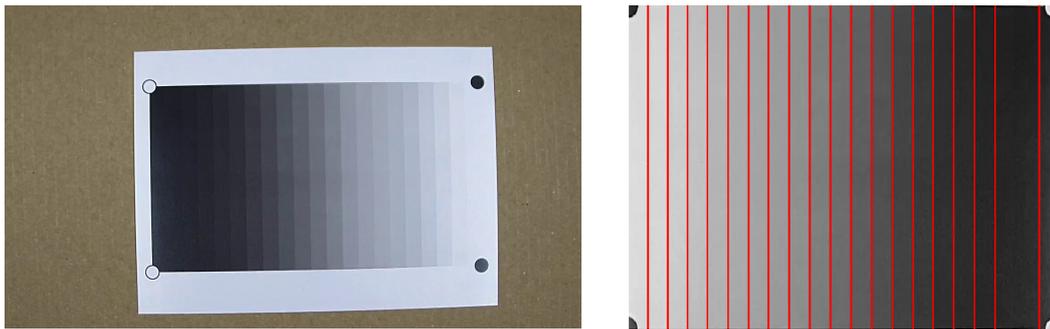
Figure 2.12: Original Image from Camera-2 (left) and Distinguishable Regions (right)

It is important to note that as expected, the performance for the black regions like the 21st step of Camera-2 is dominated by noise. The corresponding variance has a larger value with respect to the previous regions.

2.5.3 The results of Camera-3

The original image taken by the camera-3 and the distinguishable regions can be seen in Figure 2.13. Camera-3 has the best dynamic range performance according to our

tests with 21 distinguishable steps out of 21.
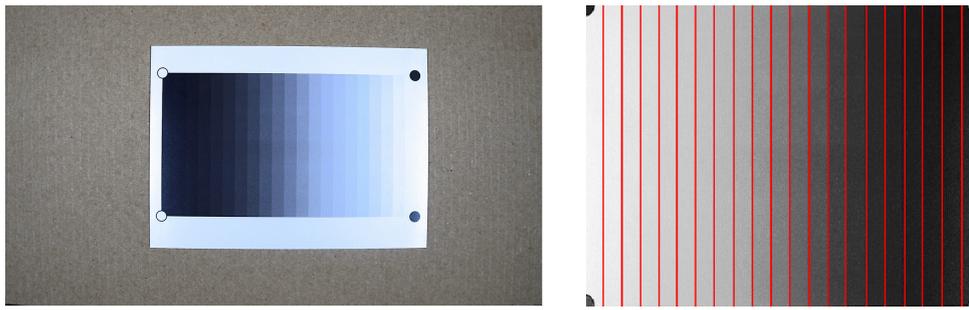
Figure 2.13: Original Image from Camera-3 (left) and Distinguishable Regions (right)

Due to the fact that there are more distinguishable black regions in the result of the Camera-3, averaging operation in σ results with a higher value than Camera-2 because of the noise in these black regions.

2.5.4 The results of Camera-4

The original image taken by the Camera-4 can be seen on the left side in Figure 2.14. The distinguishable regions are shown in Figure 2.14 on the right side. Camera-4 has 18 distinguishable steps out of 21.
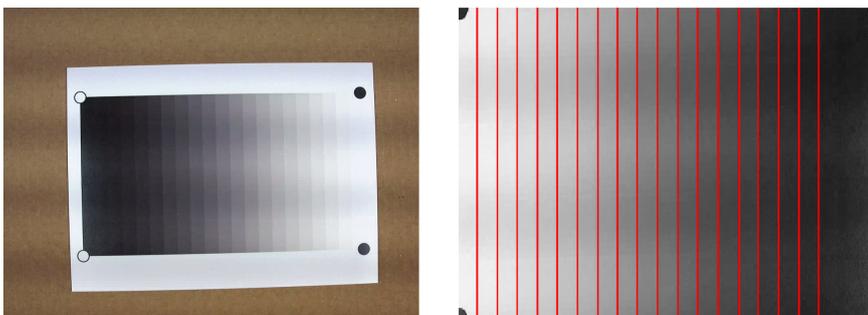
Figure 2.14: Original Image from Camera-4 (left) and Distinguishable Regions (right)

2.5.5 The results of Camera-5

The original image taken by the Camera-5 and the distinguishable regions are shown in Figure 2.15 on the right side. Camera-5 has 14 distinguishable steps out of 21.
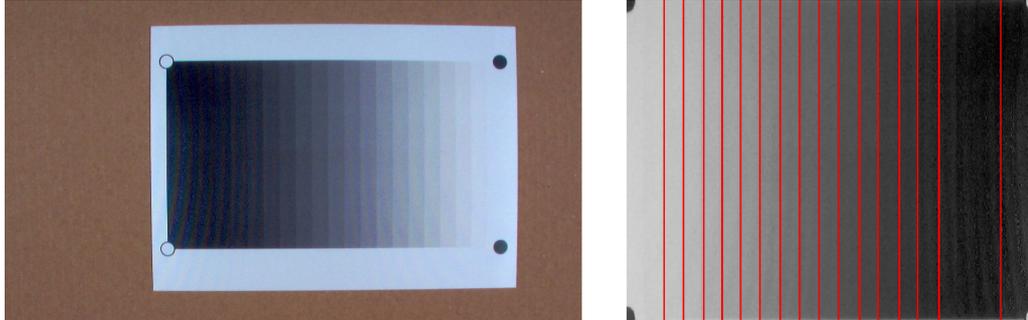
Figure 2.15: Original Image from Camera-5 (left) and Distinguishable Regions (right)

2.5.6 The results of Camera-6

The original image taken by the Camera-6 is shown in Figure 2.16. The distinguishable regions are shown in Figure 2.16 on the right side. It has 17 distinguishable steps out of 21.
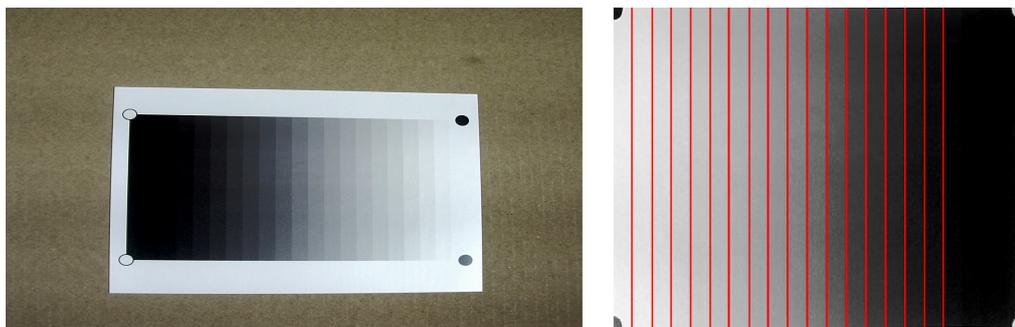
Figure 2.16: Original Image from Camera-6 (left) and Distinguishable Regions (right)

2.5.7 The results of Camera-7

According to our tests, Camera-7 has 16 distinguishable regions out of 21. In Figure 2.17, the original and the processed images can be seen.
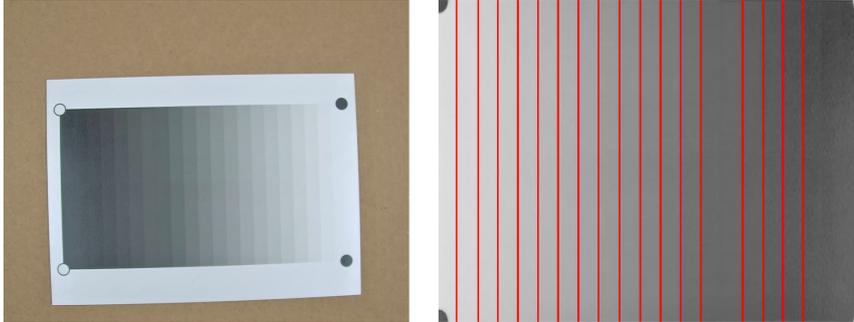
Figure 2.17: Original Image from Camera-7 (left) and Distinguishable Regions (right)

It is crucial to note that Camera-7 has the smallest average standard deviation value among the test result which shows that Camera-7 has very good noise immunity.

2.5.8 The results of Camera-8

The original image acquired by Camera-8 and the determined regions are shown in Figure 2.18. It has 16 distinguishable regions out of 21.
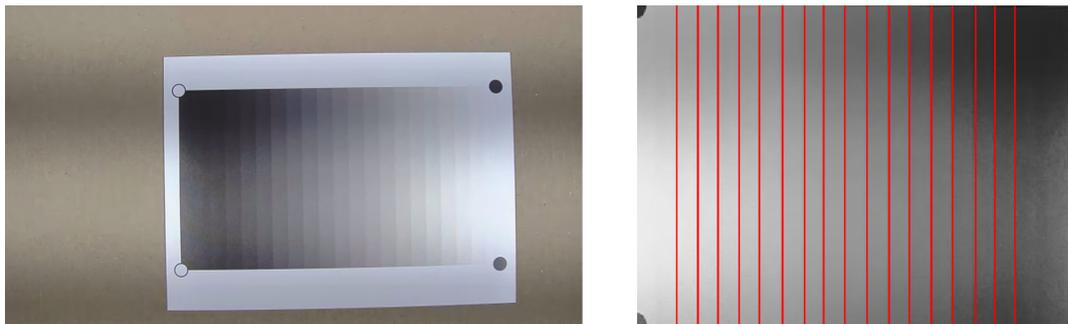
Figure 2.18: Original Image from Camera-8 (left) and Distinguishable Regions (right)

The σ value is the highest value when compared with the others. The tests are performed for the default adjustments of the camera. On the other hand, the only chance is performed for turning on the flicker correction for 50 Hz due to the fact that the light sources which are used for the test are operating with 50 Hz voltage. Because the flicker correction algorithm which is used in Camera-8 is not successful like others, it results with the high variance in the results of the Camera-8.

Table 2.1 gathers the results of the dynamic range tests for eight test cameras.

Camera No	μ	σ
1	11.38	4.08
2	23.42	1.84
3	15.48	3.83
4	18.81	6.27
5	36.86	2.57
6	23.98	3.42
7	16.2	1.42
8	16.83	18.36

Table 2.1: The Results of the Dynamic Range Tests

2.6   Environment Change and Results

The test setup proposed above provides perfect simulation in order to compare the cameras which will be used inside environmental conditions like shopping malls, hospitals, and government buildings. On the other hand, the main purpose of our tests is providing more general results such that one can easily find his/her best choice without considering the environmental conditions. With this idea, by changing the location and the properties of the light sources, the same procedure is used in order to test the cameras.

The highways are generally controlled by using security cameras in Turkey [57]. Another test setup is proposed for simulating the tunnels because we preferred to handle

a real case that cameras are commonly used. Another reason for this choice is that the conditions of the tunnels are totally different than the case simulated before. The lighting and camera system of Taksim Square Tunnel can be seen on the left of Figure 2.19.

Figure 2.19: The Lighting and Camera System in Taksim Square Tunnel, Istanbul [8]

If a car is considered on the right lane, the light source on the right side provides more light on the car while the camera is recording from the up. In other words, this time the luminance value on the test chart should not be distributed homogeneously like the case before. Thus, using only one light source from left or right would sufficient to simulate the lighting conditions of a tunnel. The test chart and the camera are placed in the same places as the previous test in order to obtain the correlated results. The proposed test setup can be seen in Figure 2.20.

As shown in Figure 2.20, the location of the light source is changed in order to simulate the tunnel conditions. On the other hand, as shown in Figure 2.19, high-pressure sodium vapor lamps are generally used in the tunnels in Turkey instead of the LED ones used in the previous test. Thus, changing the type of the light source is also necessary. In this purpose, OSRAM Vialox NAV-T Super 4Y, 70 watts lamp is used (Manufacturer Part No: 4050300015736). The lamp is placed inside Heper Moonlight 70 Watts enclosure (Manufacturer Part No: LS0002.32). It is important to note that sodium vapor lamp has 2000K color temperature and 6600 lumen luminous flux value [58].
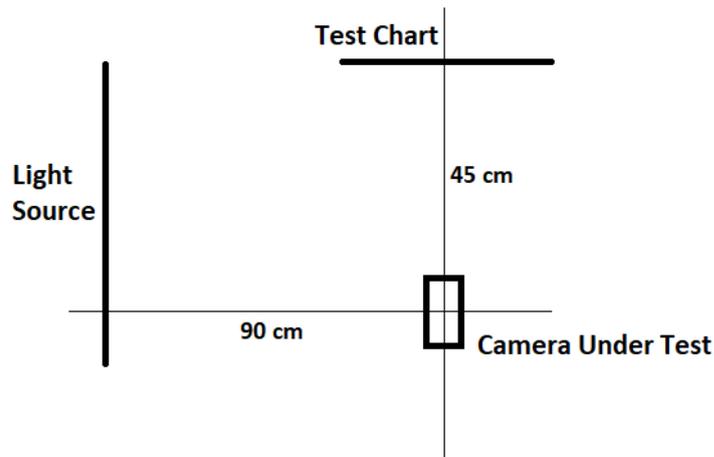
Figure 2.20: Test Setup (not to scale)

There is a main difference between using LED light source and the sodium vapor lamp. While the light leaves the LED source from the whole front panel homogenously by the means of the distribution of LEDs, in sodium vapor lamp case, the light comes directly from the lamp and thus it leaves the enclosure front panel not homogeneously. In other words, the illuminance value in the middle of the front panel is higher than the edges. As stated before, the luminance value is not homogeneously distributed from left to right of the test chart, however, the homogeneity from up to down should be satisfied. This situation can be provided by using a reflective studio umbrella because the explained property of sodium vapor lamp makes unsuitable to use it directly. If the lamp is used directly to illuminate the test chart then some region would have more light according to the others such that the necessary homogeneity cannot be obtained. The homogeneity of the light is checked by using lux meter and 500 ± 100 lux value is provided after 5 minutes since powering up the light source. The real test setup can be seen in Figure 2.21.

27

Figure 2.21: Real Test Setup

The test is performed only with 3 cameras in order to compare the results with the previous test. Camera 3, 4 and 5 are chosen for this test. Due to the color temperature of the high pressure sodium vapor lamp, the colors in the image acquired by the cameras differ than real ones which result in meaningless mean error values calculated with the previous method. Thus, in the next results, the mean error value is not considered as a parameter that can be used in order to compare camera performances. On the other hand, the mean value of the variance in the distinguished regions is still a parameter that can indicate the noise performance in that light.

2.6.1 The results of Camera-3

The original image acquired by the Camera-3 in the test setup shown in Figure 2.21 can be seen on left of Figure 2.22. All the regions are detected by the test software as shown in the right side of Figure 2.22. According to the previous test results, Camera-3 performs the best dynamic range among the others. Also for the second test setup, it has the best performance with respect to Camera-4 and Camera-5.
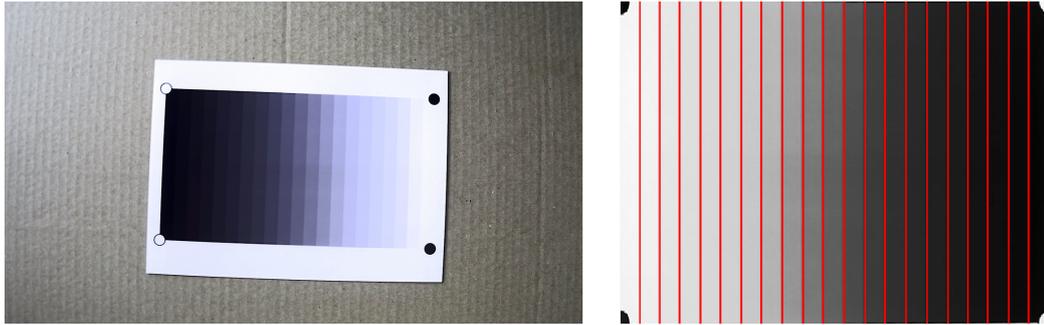
Figure 2.22: Original Image from Camera-3 (left) and Distinguishable Regions (right)

2.6.2 The results of Camera-4

The image taken by Camera-4 and the distinguished regions are shown in Figure 2.23. In this test setup, the light source is placed left side of the chart which results in higher
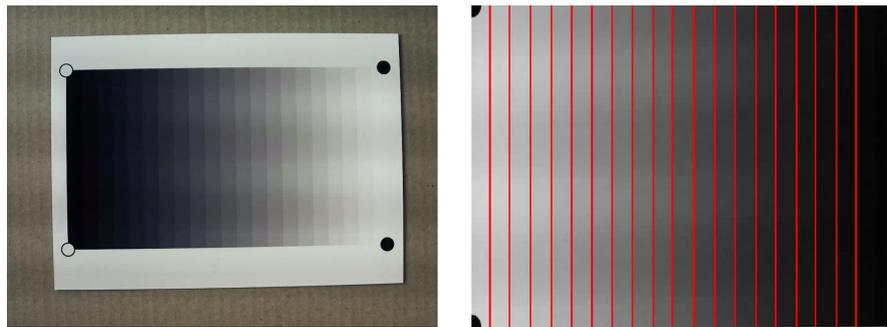
Figure 2.23: Original Image from Camera-4 (left) and Distinguishable Regions (right)

illumination on the darker regions than the lighter ones. As a disadvantage of using a reflective chart, this situation turns the darker regions into a little bit brighter ones because they are illuminated more and lighter regions into a darker version because they are illuminated less. Due to this shrinking, the dynamic range of that can be measured by the test chart decreases. Hence, it is expected that the number of distinguishable regions of the cameras would increase. This effect couldn't have been observed before because Camera-3 had already had 21 regions for both cases. However, the effect of test setup should be considered in the test result of Camera-4. Camera-4 performs to distinguish 19 regions while it has 18 regions in the previous test setup. On the

other hand, this shift in the test results would not change the ranking of the cameras when they are compared with others.

2.6.3 The results of Camera-5

The image acquired by Camera-5 in the second test setup and the corresponding result can be seen in Figure 2.24. It has 18 distinguishable regions out of 21. As explained before, it is expected that the number of distinguishable to increase. Even though it has 14 regions according to the previous test result, it increases to 18 due to the effect of the test setup. On the other hand, the number of distinguishable regions is still less than the Camera-3 and Camera-4 which show that it has the shortest dynamic range among. As can be observed from the results of both first and second proposed test
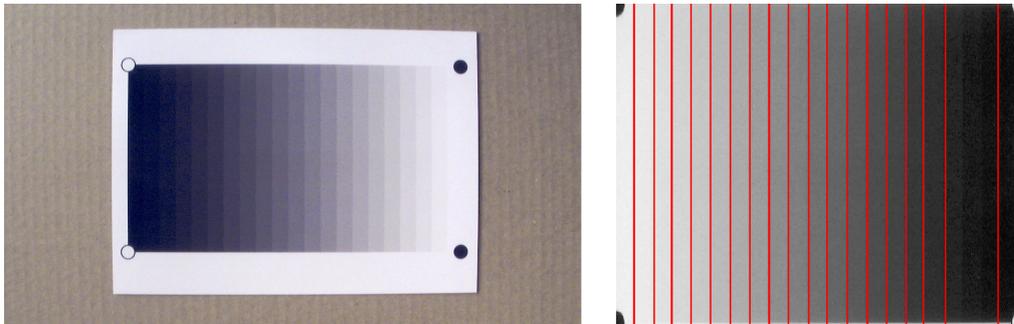
Figure 2.24: Original Image from Camera-5 (left) and Distinguishable Regions (right)

setups, the environment change can partly change the results while the performance comparison between the same cameras would not change as shown in Table 2.2

Camera No	Distinguishable Regions-1	Distinguishable Regions-2
3	21	21
4	18	19
5	14	18

Table 2.2: The Results of the Dynamic Range Tests

CHAPTER 3

CAMERA TESTS ON COLOR AND RESULTS

3.1 Background Information about Color Accuracy and Color Constancy

In this chapter, the importance of color accuracy for the performance of the camera will be examined because the color is one of the most important information about the objects around us. However, before going into the details, the concept of color should be understood clearly in both physically, biologically and psychologically. The shortest definition of the color is the byproduct of the spectrum of the light received by the human eye and processed by the human brain [59]. The definition proposes 3 factors for understanding color from different areas of the science.

The first factor is the spectrum of the light which is the physical component of the color. In general, the light is part of the electromagnetic spectrum ranges from radio waves to gamma rays according to its wavelength [60]. On the other hand, only a small portion of that range can be seen by the human eye as can be seen in Figure 3.1.
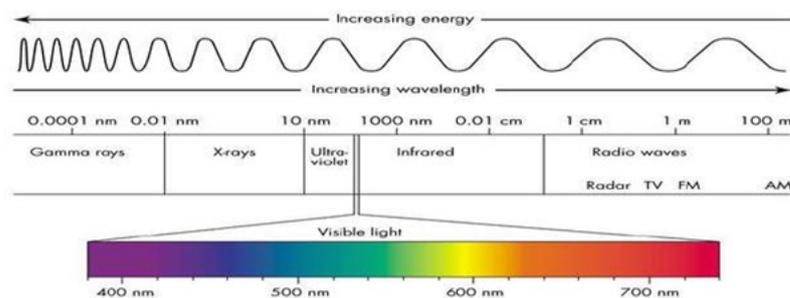
Figure 3.1: The Spectrum of the Light [9]

When the color is considered, the importance of the spectrum of the light differs between two cases. If the object is a light source, then its color depends on the light leaving the light source. In this case, the spectrum of the light which affects the color of the object is the combination of light waves with different wavelengths. However, if the object is not a light source, there should be a light source in order to get the color information of the object. In this case, the color depends on the different light waves reflected from that object. The important point here is that the sensed color depends also on the incoming light because if the incoming light does not have light waves which will be also reflected by the object, then the real color information of the object cannot be obtained. The color signal received by the eye is the wavelength-by-wavelength multiplication of relative energies of illumination on the object and the reflectance of the object as exemplified in Figure 3.2 [10].
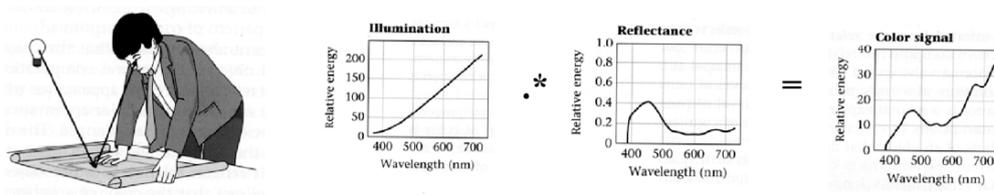
Figure 3.2: Simplified Rendering Model for Reflectance [10]

As explained above, the images of the objects can show differences in color with the different illumination conditions. The images of the same flower under the different light sources can be observed in Figure 3.3 as an example.

The second factor is the human eye which is the biological component of the color definition. The eye is one of the most complex organs of the human body that performs the conversion of light into the neural data. Due to its complex structure, only rod and cone cells where the conversion is actually made will be covered in this chapter. These are two kinds of human photo receptors which are categorized according to their shapes as shown in Figure 3.4.

The rod cells are more sensitive to the light than cone cells however not to the color. They have more photopigments and high amplification such that person can see objects in a dark room with very little light. The vision data acquired by the rod cells give information about motion around [61]. A normal human eye contains around

32

Figure 3.3: The Color Change with the Different Light Sources [11]
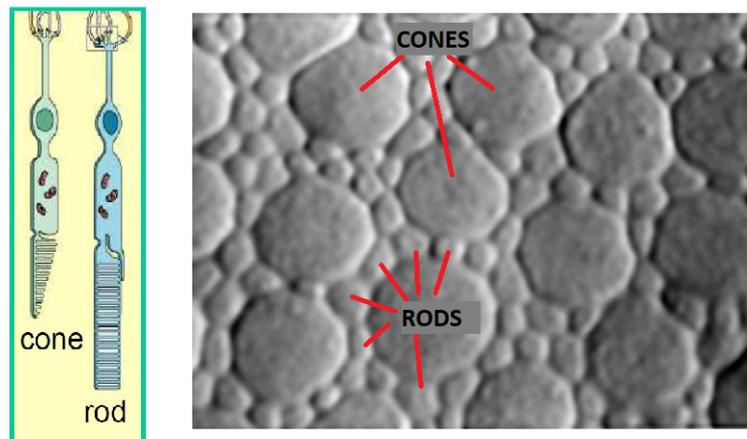
Figure 3.4: The Human Photo Receptors [12]

120 million rod cells [62]. On the other hand, the cone cells are responsible to obtain the color data and the sharp and detailed vision. Although cone cells are less sensitive to light, they provide color data which makes them important for this chapter. There are 3 types of cone cells in a normal human eye according to the color that they are sensitive. 6 million cone cells in the human eye are split into categories as red, green and blue sensitive cones with the percentages of 64 % , 32 % and, 2 % , respectively [62]. They absorb the light and generate related neural data about the color as shown in Figure 3.5 [63]. This behavior gave scientists good insight on how to obtain and store the color information. According to the Tristimulus Theory, under standardized conditions, all the colors can be matched with the linear combinations

of the three primary colors; red, green and blue [64]. On the other hand, even they all use three variables in order to describe a color, there are different color spaces in the literature and they will be discussed later in this chapter.
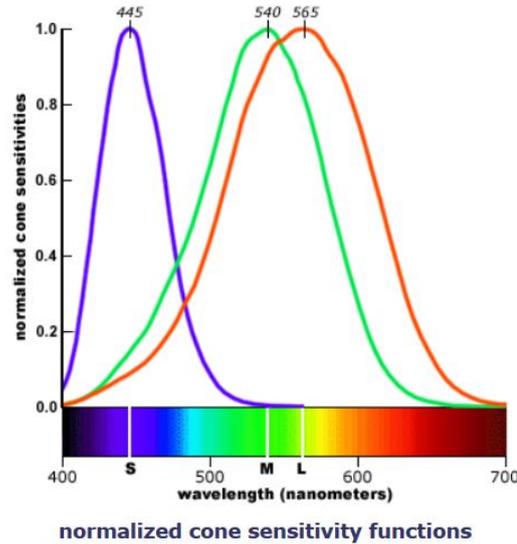
Figure 3.5: The Normalized Sensitivities of Red, Green and Blue Cone Cells [13]

The third factor in the color definition is the human brain. As described above, the process of obtaining the color information starts with the light coming to the eye. Thanks to the photoreceptors, the physical information is converted into neural data and sent to the brain. As the last step, the data is processed in the brain and the color of the object is determined. However, the processing of the neural data is still a mystery for mankind.

The most important effect of the processing in the human brain is called as color constancy in the literature. As explained before, the spectrum of the illumination on the object affects the reflected light from the object thus it affects also the color. On the other hand, the human brain has a complex color perception system which tries to guaranty that the perceived color of the object stays the same under different illumination conditions [65]. This effect was first described by Edwin Land in 1971 and he proposed the "retinex" term in order to explain that the color is the process of both "retina" in the eye and "cortex" of the brain [66]. When the brain takes the data from the eye, it tries to approximate the composition of the illuminating light. Then, the effect of illumination is discounted in order to understand the true color of the

34

object. Today, the same idea is used in the cameras as a post-processing step in order to increase the color accuracy. Some different techniques proposed in [67], [68], [69] such as Fourier transform, derivative and convolutional neural networks are used in order to approximate the illumination on the scene because the process in the brain has not been discovered clearly yet. However, the best results can be obtained by using a test chart and form a color conversion matrix for the different illumination conditions. The details about how the color conversion matrix can be found and used will be given later.

After this background information about the color, it is important to know how the cameras obtain the color data of the image. The cameras try to imitate the working principles of the human visual system. As explained above, the human eye responds differently for 3 colors; red, green and blue; due to the existence of different cone cells. Thus, scientists and engineers invented methods to capture the color information in a similar way. As explained in the previous chapter, the light is converted into electrical data by using the photosites. However, the photosites are color-blind; in other words, they cannot give the color information. Thus, 3 different methods are used [14]. In the first one, a beam splitter and 3 photosites are used in order to get red, green and blue information like the human eye does as shown in the left side of Figure 3.6. In this method, the incoming light passes through the beam splitter and its red, green and blue components are reflected on the corresponding photosites. Even this technique gives the best performance, using a beam splitter and 3 times more photosites result with bigger design and more expensive price according to the others. In the second method, 3 filters corresponding red, green and blue are placed in front of the same photosite array sensor in rapid succession. In using this method, for each color component, the potential differences of the photosites are saved and the image is composed by using this data. However, even the switching between the filters is performed very fast, due to the time difference, both camera and the target scene must be stationary for the best performance. This problem makes the method unsuitable for most of the applications. The last method is placing a Bayer filter in front of the sensor and performing demosaicing algorithm as post-processing. The basic element of Bayer filter can be seen in Figure 3.6 which contains 1 red, 2 green and 1 blue filters because the human eye is more sensitive to green light according

to the other two. Also, this placement results with smoothness and less noise in the image as a result of the demosaicing algorithm.
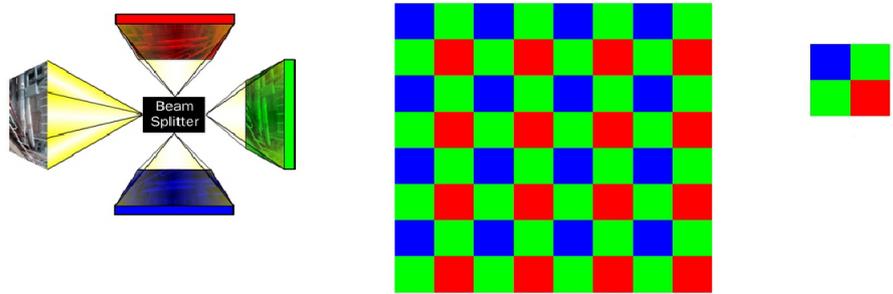
Figure 3.6: The Beam Splitter Method (on the left), Bayer Filter (middle) and Its Basic Element (on the right) [14]

Although there are many other alternatives to Bayer filter in the literature, the idea is the same. Each photosite captures the related color data according to the filter in front of it and then the raw image is processed by a demosaicing algorithm in order to turn the data into a useful image. Generally, 3-by-3 or 6-by-6 neighborhood of the pixel is taken into account in order to obtain the red, green and blue color component values in that pixel. The more detailed explanation will not be given in this thesis about the demosaicing algorithms because the performance testing is not conducted through the raw images. Working on directly R*G*B values provides us to compare also demosaicing algorithm performance plus the sensor of the camera as a whole.

Although R*G*B values are used instead of raw values, the R*G*B color space is not the best color space which can be used for measuring color accuracy. R*G*B color space is the most well-known color space in the literature; however, as mentioned above, there are other color spaces. As mentioned before, according to the Tristimulus Theory, all the colors can be interpreted by the linear combinations of red, green and blue colors. In order to prove the idea, in the 1920s, Guild [70] and Wright [71] , [72] performed some controlled tests and at the end, in 1931, International Commission on Illumination standardized and gathered the results by forming the color space of CIE XYZ 1931. The most important property of CIE XYZ color space is that it is the largest color space in the literature and it depends on the human vision

system directly. Guild and Wright performed color matching experiments in order to find out how to form all the colors in the visible light range from 380 to 700 nm by using only red, green and blue primary lights. Although there are some differences between the versions of 1931, 1964 and 1976, the test setup is shown on the left side of Figure 3.7. In one side of the panel, the test color in the specified wavelength is obtained by filtering the N.L.P white light source; and on the other side of the panel, the red, green and blue light sources are placed [73]. In 1931, the wavelengths of the red, green and blue light sources were determined by CIE as 700 nm, 546.1 nm and 435 nm, respectively; however, in 1964, they rearranged as 645.2, 526.3 and 444.4 nm. It is important to note that these light sources are dimmable such that the observer can adjust the brightness of each light source. According to the proposed experiment, a standard observer with normal vision adjusts the three light sources in order to obtain the same color as the test side of the panel. The vision of the standard observer can be seen on the right side of Figure 3.7. This procedure is repeated for the visible range with 5 nm steps from 380 to 700 nm and also for a large group of the standard observers in order to guarantee the objectiveness. The results by using the light sources specified in 1964 are shown in Figure 3.8.
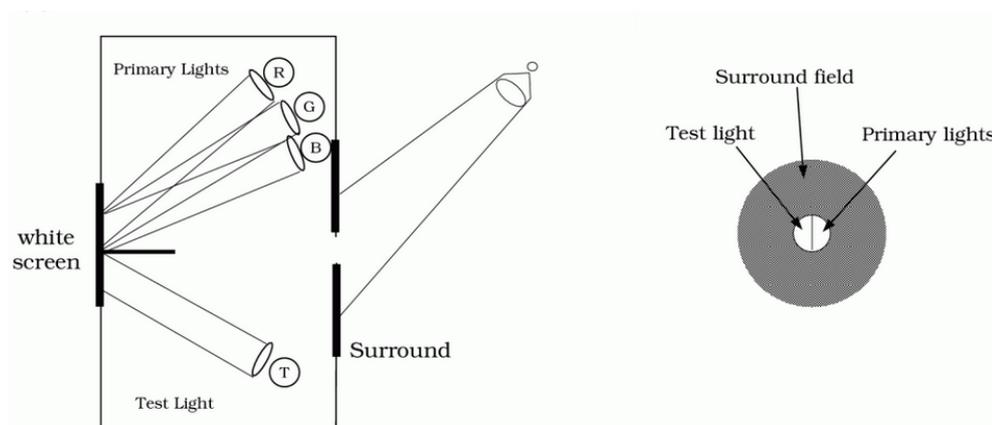
Figure 3.7: The Color Matching Experiment Setup (on the left) and The View of Standard Observer (on the right) [10]

The most important and also interesting regions are one of the primary intensities are below zero. Of course, by just dimming the light source, it is impossible to get negative intensity on the test panel; thus, negative intensity means that the corresponding light source is added to the test light side of the panel as shown in the Figure 3.9. It
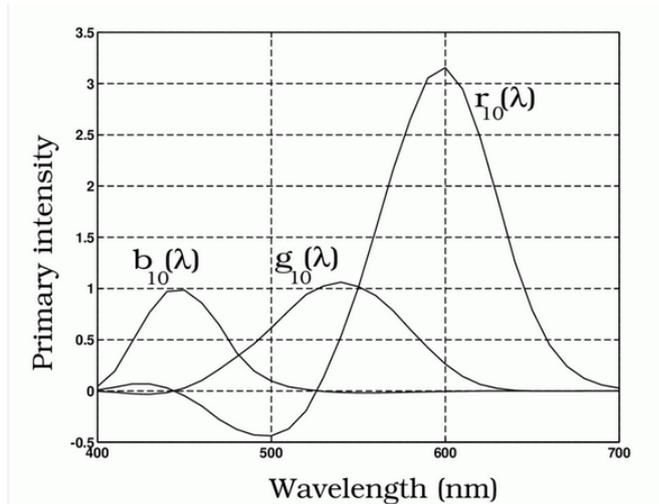
Figure 3.8: The Results of Color Matching Experiments according to Stiles and Burch (1959) [10]

is important to note that p1, p2 and p3 are used in order to show the intensities of the light sources. The results prove that there are colors which can be seen by the human eye but cannot be obtained by just adding the red, green and blue colors. Later, the more detail will be given why this is important for camera performance tests.
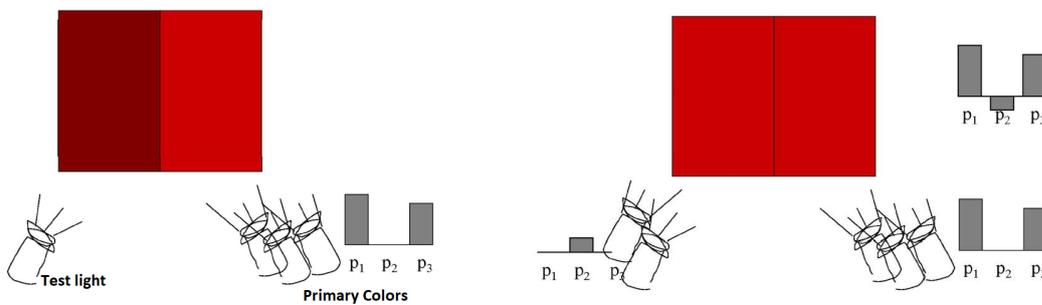
Figure 3.9: The Special Case of Color Matching Experiments [15]

Due to the fact that both representation and working with negative numbers are difficult, the scientists in CIE proposed imaginary primary colors $\bar{x}$, $\bar{y}$ and $\bar{z}$ instead of real light sources with specified wavelengths such that each color can be represented by 3 positive real number. The same graph with Figure 3.8 is plotted by using $\bar{x}$, $\bar{y}$ and $\bar{z}$ primary colors and as expected, each color is expressed by 3 positive components as shown in the left side of the Figure 3.10. Also, by applying some transformations,

the color information is proposed to be formed by 2 variables x and y and one luminance value Y. If the luminance value is ignored, the human vision color space can be represented by x and y variables as shown in the right side of Figure 3.10. It is important to note that due to the transformations mentioned above, the x-y graph has a horseshoe shape. More detailed information about the transformations will not be given in the scope of this thesis.
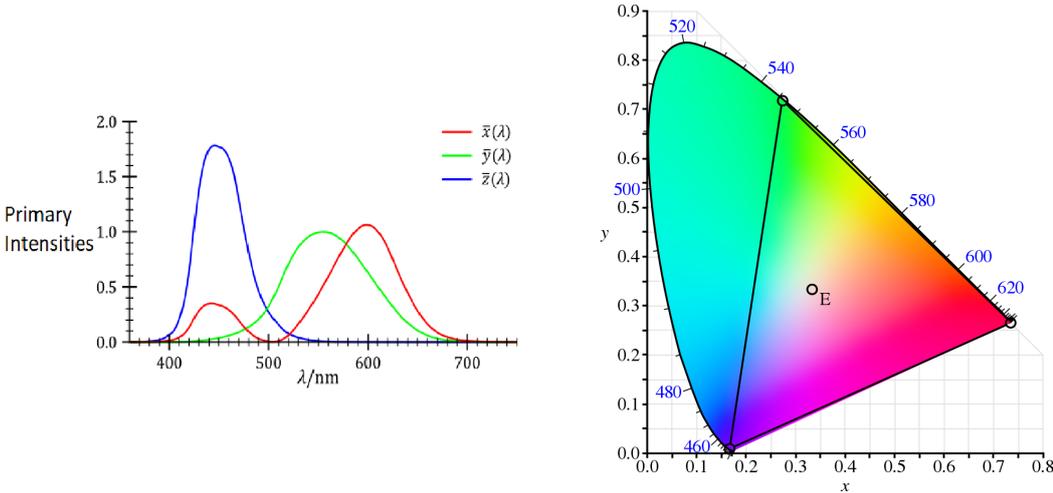
Figure 3.10: The Color Matching Experiment with Primary Colors of CIE XYZ Color Space (on the left) and The CIE XYZ 1931 Color Space(on the right) [16]

As stated before, the color matching experiments were performed by using red, green and blue light sources which have wavelengths, 700 nm, 546.1 nm and 435 nm, respectively. All the colors which can be obtained by adding different but positive intensities of these three primary components are being represented by a triangular in the x-y space as shown in Figure 3.10. The corner points are placed on the corresponding points to the 700, 546.1 and 435 nm. This color space is named as CIE RGB color space or "Wide-Gamut RGB" and obviously, it is a subspace of the CIE XYZ color space. In other words, all the colors are out of "Wide-Gamut RGB" color space cannot be obtained by adding the 3 real light sources. Actually, it is an important point to remember how cameras work and think about its relation with the color space. Like adding colors in the color matching experiments, as discussed before, the cameras have bandpass filters such that they can find the intensities around the specified wavelength, red, green and blue by name. So, it is just the reverse process of the adding colors in the color matching experiment. Thus, if a color cannot be obtained

by adding the primary colors in the color matching experiment, it is also impossible to get that color by using a camera. In other words, there is no camera in the market such that it can perfectly get the color which is out of "Wide-Gamut RGB" color space. Actually, this is the reason why we cannot get the same colors when we use our camera and when we see the view or object with our own eyes.

Until this point, a color space which covers all the colors that can be seen by a normal human is proposed and also a subspace of that color space is determined which can be detected by using a camera. These are the color spaces of CIE XYZ and RGB, respectively. However, these color spaces are a little bit hard to work with for photographers and graphic artists [74]. Also, the spreading of the Internet and public broadcasting forced the scientists to think about bandwidth. As a result, Hewlett-Packard and Microsoft proposed a color space named "sRGB" as a subspace of "Wide-Gamut RGB". This color space has been approved by International Electrotechnical Commission and it is still the most widely used color space [75]. The sRGB color space can be seen in Figure 3.11, and it is important to emphasize that it is a subspace of "Wide-Gamut RGB" which is shown in Figure 3.10, before. So, even if the camera captures the correct color which is outside of the sRGB color space but inside the "Wide-Gamut RGB", the color information will not be able to be transmitted correctly when it is converted into sRGB space.

Although, our proposed performance test can be applied to every kind of camera, our main purpose to use them to be able to compare the performance of security cameras. Nowadays, the most of the security cameras use MJPEG or H.264 video compression standards due to the bandwidth considerations mentioned above. "YCbCr" color space is used instead of directly using sRGB for both of these compression standards. "YCbCr" color space is composed of 3 variables which are consisted with Tristimulus Theory; Y, Cb and Cr. Y is the luminance component of the color which is more important for the human visual system. Cb is the blue component relative to the green component and Cr is the red component relative to the green component [76]. The more detailed information about "YCbCr" color space will not be given in this thesis because it requires a strong background in video processing. On the other hand, "YCbCr" color space is very similar to sRGB space and the conversion can be per-
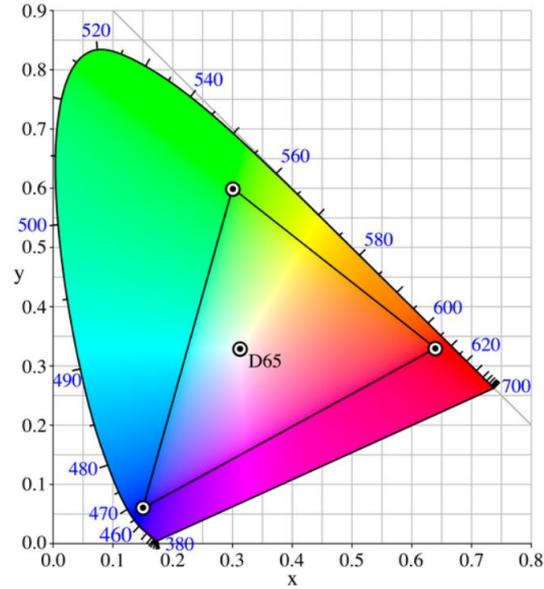
Figure 3.11: The sRGB Color Space [17]

formed by using the following formulas [77].

$$Y = (77/256)R + (150/256)G + (29/256)B \tag{3.1}$$

$$Cb = -(44/256)R - (87/256)G + (131/256)B + 128$$

$$Cr = (131/256)R - (110/256)G - (21/256)B + 128$$

As mentioned above, using directly sRGB values which are acquired and transmitted by the camera provides us to compare not only the lens or sensor performance but also post-processing and video encoding capabilities of the camera. Apart from the perspective of a scientist, this method is more meaningful for the end users of the camera because they mainly care about only the last video decoded and optionally saved in their system for later use instead of raw data of the camera. Although it is stated that R*G*B values are used for performance comparison instead of raw values from the camera, "sRGB" is not the color space that the color accuracy is measured. The main reason is that the equal distance from a point in the RGB color cube does not imply the same color difference according to the human visual system because the human eye is more sensitive to green color than red and blue. It is crucial to note that the same problem of sRGB space is also valid for CIE XYZ color space. In other

words, changing the color with a constant distance would not result with the same amount of color change in CIE XYZ color space. Thus, in 1976, CIE proposed CIE Lab color space by using the observations in the CIE XYZ color space. CIE Lab color space has also all the colors that a normal human eye can see but the components of the color space are more useful in order to understand the color difference between two different colors than CIE XYZ color space. The "L" component represents the luminance value of the color while "a" and "b" components are responsible to indicate the color. CIE Lab color space can be visualized as a sphere as shown in Figure 3.12.
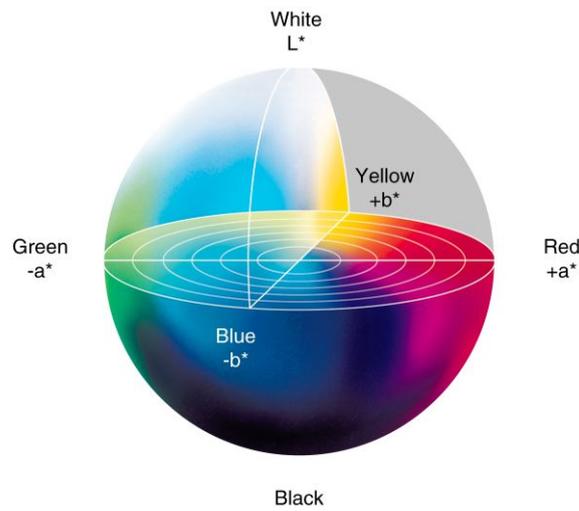
Figure 3.12: The CIE Lab Color Space [18]

According to the CIE 1976 standards, Delta E criteria started to be used for measuring the color differences between two points in the sphere. According to this criterion, Delta E can be calculated for two points say (L1, a1, b1) and (L2, a2, b2) with the following formula [78]:

$$DeltaE = \sqrt{((L_1 - L_2)^2 + (a_1 - a_2)^2 + (b_1 - b_2)^2)} \tag{3.2}$$

CIE made some changes for the formula in 1994 and proposed another metric in 2000. However, due to their concepts and complexities, Delta E formula has been still the most widely used metric in order to measure the color accuracy. Moreover, the Colour Measurement Committee of the Society of Dyers and Colourists defined their color difference metric in CMC 1:c standard in 1984. However, this standard has not been accepted in the photographic industry [78]. Thus, Delta E criterion is

used in our color accuracy calculations. However, due to the fact that Delta E can be calculated by using L, a and b components, the conversion from sRGB to Lab color space should be performed. The conversion is performed in 2 steps; at first sRGB to XYZ and then, from XYZ to Lab.

For image "I", the conversion from sRGB to XYZ is performed by first normalizing the values between [0,255] into [0, 1] for R, G and B components. Then, gamma correction is being applied and linearized components, I'_R, I'_G and I'_B are obtained for each channel R, G and B by the following formula where i represent the channel of image "I".

$$I'_i = \begin{cases} -I_i/12.92 & \text{,if } I_i \leq 0.04045 \\ ((I_i + 0.055)/1.055)^{2.4} & \text{,otherwise} \end{cases} \tag{3.3}$$

After the linearized components are obtained, the conversion can be performed as follows [79].

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.4124 & 0.3575 & 0.1804 \\ 0.2126 & 0.7151 & 0.0721 \\ 0.0193 & 0.1191 & 0.9503 \end{pmatrix} * \begin{pmatrix} I'_R \\ I'_G \\ I'_B \end{pmatrix} \tag{3.4}$$

In the second step, conversion from XYZ to Lab, 3 dummy variables, D_X, D_Y and D_Z have to be calculated with the same formula below with the normalized values N_X, N_Y and N_Z which are equal to (X/96.6797), (Y/100) and (Z/82.5188), respectively. Please note that i represents the channels of X, Y and Z.

$$D_i = \begin{cases} N_i^{(1/3)} & \text{,if } N_i > 8856 * 10^{-6} \\ (7.787 * N_i) + (16/116) & \text{,otherwise} \end{cases} \tag{3.5}$$

As the last step, L, a, b values are calculated from these dummy variables as follows:

$$L = 116 * D_Y - 16 \tag{3.6}$$

$$a = 500 * (D_X - D_Y)$$

$$b = 200 * (D_Y - D_Z)$$

Please note that conversion from CIE XYZ color space to Lab color space formulas are given for D50 white reference point; the reason of it will be more clear in the next part. When the conversion from sRGB to Lab color space is done; by using the Delta E formula, the difference between the color acquired by the camera and the color value in real can be calculated.

3.2 Test Chart

As mentioned above, if we have color patches in our test chart with known Lab values, then capturing an image by the camera and calculating the difference between these two colors by using Delta E formula gives a good idea about the color accuracy performance of the camera. This method is used as the proposed test-2 in our camera performance measurement.

Actually, with the same idea, "Macbeth" chart was proposed in 1976 by McCammy, Marcus and Davidson in the Journal of Applied Photographic Engineering in order to test and /or to calibrate every kind of device-related about color like cameras, monitors, scanners and printers [80]. The "Macbeth" color rendition chart is composed of 4x6 patches each about 50 mm square as shown in Figure 3.13. The colors were selected such that they can simulate real-world subjects like human skin, foliage, sky and flowers.

Today, all rights of the proposed "Macbeth" test chart belong to X-Rite Incorporated. They also proposed different color test charts like Colorcheck Digital SG and Colorcheck Passport with the same purpose of Colorchecker Classic (Macbeth Chart). There are also other companies which produce similar test charts like DataColor and Image-Engineering GmbH. SpyderCheckr series and TE226 can be seen as examples from these companies in Figure 3.14, respectively.

Due to the fact that Colorchecker Classic is the most affordable and widely-used test chart in photography, we preferred to use it in our color accuracy performance tests. It is important to note that the chart was revised in November 2014 and the final reference values were declared in the website of X-Rite Incorporated [81]. These final reference values are composed of L, a and b components of each patch in the test chart

Figure 3.13: The Color Rendition Chart [19]

Figure 3.14: The Alternative Charts in the Market (from left to right, Colorchecker Passport, Colorchecker Digital SG, SpyderCheckr and TE226)

with respect to D50 white reference. This is the reason why the conversion formulas from XYZ to Lab are given above with respect to D50 white reference points.

3.3 Test Setup

The test setup which proposed in the previous chapter is also used for the test of color accuracy tests. The color temperature of the light sources in the proposed test setup is 5000 K which is suitable according to the paper of McCammy, Marcus and Davidson. Other critical points were stated in the website of X-Rite Incorporated with the heading of "Tips for Capturing the ColorChecker Classic Target" [82]. As described, the test chart covers more than 25 % of the captured image in our test setup.

Moreover, the test chart is placed almost parallel to the camera lens in order to prevent any glare in the image and the focusing operation is performed before acquiring the images. After all the conditions are satisfied, the image is obtained and the test which will be covered in the next part is performed.

3.4 Test Details

As explained before, the main purpose of this thesis is to prepare an automated test for measuring the performance of cameras, especially security cameras. The results of the tests can be used in order to compare the different cameras by the end user. In order to automate the system, as proposed in the first test, locating the color patches is the first and the most important step of the test. Due to the fact that almost every camera even from the same brand performs different performances in the color tests. Hence, using only one method based on the R*G*B or grayscale values is never enough to locate the color patches for every camera in the market. Thus, we implemented different methods that can find the location of color patches and merge the data from each method in order to form the final mask which will be used in the rest of the test. In each method, the input image is resized into the size of 720x1280 which is the most common size. Please note that this resizing operation does not affect the image under test but provides some advantages while creating the color patch masks.

In the first method shown in Figure 3.15, the input image is converted from R*G*B color space into grayscale by using the given formula in the previous chapter. Then, Otsu's method is used on the grayscale image in order to obtain the binarized image which will have logic (1) in the color patches and logic (0) in the other regions. It is important to note that instead of a single fixed threshold value, Bradley's adaptive thresholding method is used in order to update the threshold value. After performing the binarization, the opening morphological operation is done in order to discard the noise in the image. As the last step, the regions which are smaller or bigger than the possible color patch size are filtered. Due to the fixed size of 720x1280, the color patches have total pixel numbers between 8000 and 30000; so, these are the limit values which are used in our tests so far. The more details about Otsu's and Bradley's thresholding methods and opening operation will be given in the Appendix part.
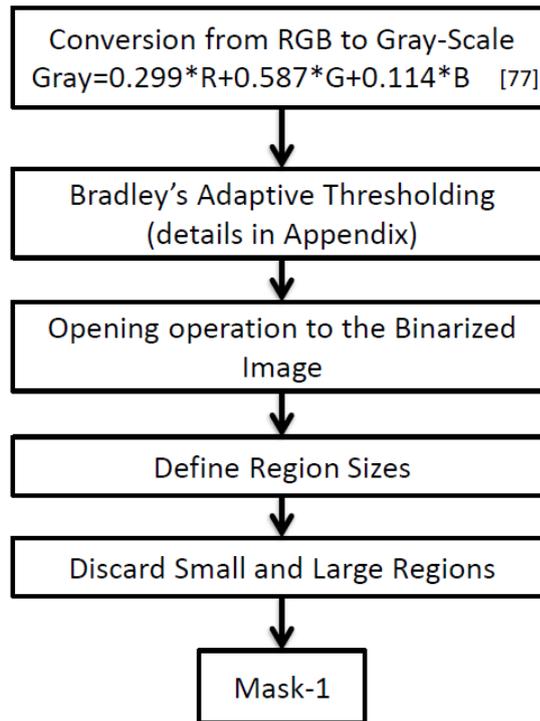
Figure 3.15: The Flowchart of the First Method

During our tests, we recognized that performing the thresholding operation on the grayscale version of the input image results with not successfully finding the location of some color patches. On the other hand, using the thresholding method on each R, G and B color channels separately helped us to locate these color patches. Also, in our tests, we performed histogram equalization on each channel before using the thresholding and this mid-step increased the accuracy. After thresholding, the resulting binary images from each channel are "OR"ed in order to find the output mask as shown in Figure 3.16. Lastly, the same opening and area filtering operations are performed. The details of histogram equalization will be given in the Appendix part.
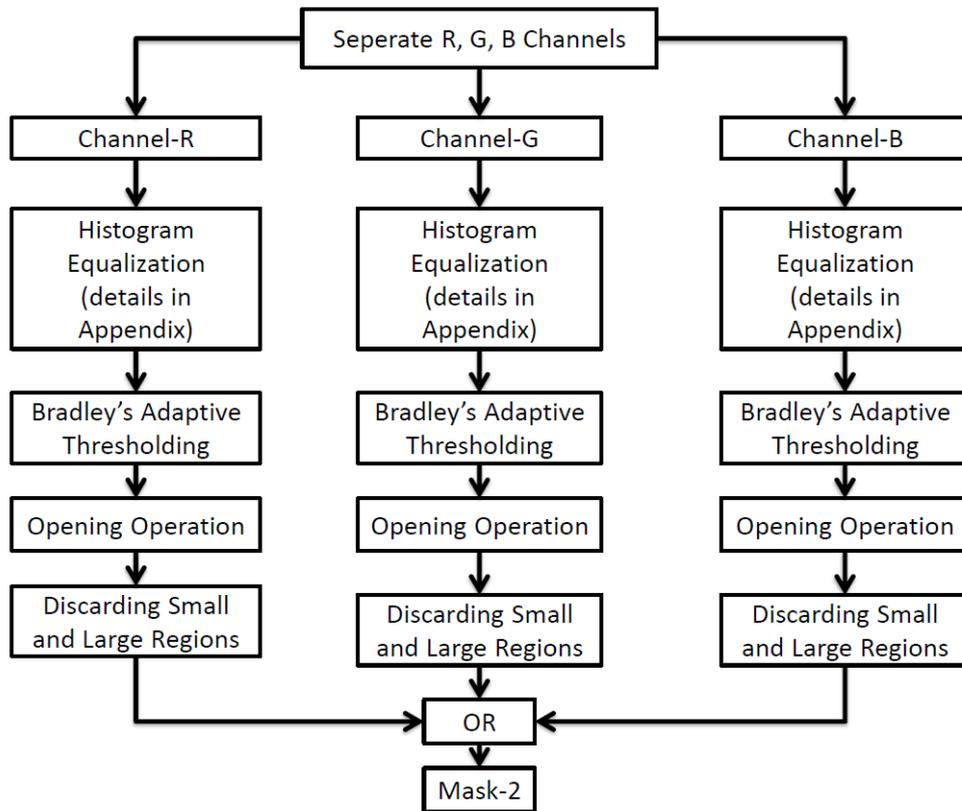
Figure 3.16: The Flowchart of the Second Method

In the last method, k-means algorithm is used in order to classify the regions before using thresholding as given in Figure 3.17. Firstly, histograms of R, G and B channels are obtained. Then, the number of peaks of these histograms is calculated and the maximum of them is used as the class number in the k-means algorithm. Please note that the maximum of peak number does not provide directly the number of color patches in the test chart plus the other regions behind the test chart. It gives just a rough idea about the image and by this idea; a preprocessing step is performed for the thresholding. In other words, not only k-means algorithm is used in order to find the locations of the color patches. There are two main reasons for that. Firstly, we do not know the information about the background of the test chart. There can be a background with different colors or due to some reason different shades of the same color which can increase the number of class numbers. Secondly, even in the best case with the uniform background; there will be 26 different classes in the image (Uniform background + black region which separates color patches + 24 color patches). This number is too large for the k-means algorithm to converge in practice. Thus, a small

number is chosen by using the histogram as class number. After determining the class number, a color space conversion is performed on the input image. As mentioned before, Lab color space is more suitable for working about the color instead of sRGB color space. Thus, using the formulas above, the conversion is done and k-means algorithm is used on the a-b plane with the determined class number. The working principle of k-means algorithm is given in the Appendix part. For each class, the area-size filtering is performed like the other methods while the other classes in the image are turned into black value. In other words, the classification is performed in Lab color space but thresholding is done in R*G*B color space by the same way as the previous method. Using the adaptive thresholding for the different classes separately increases the accuracy of the method.
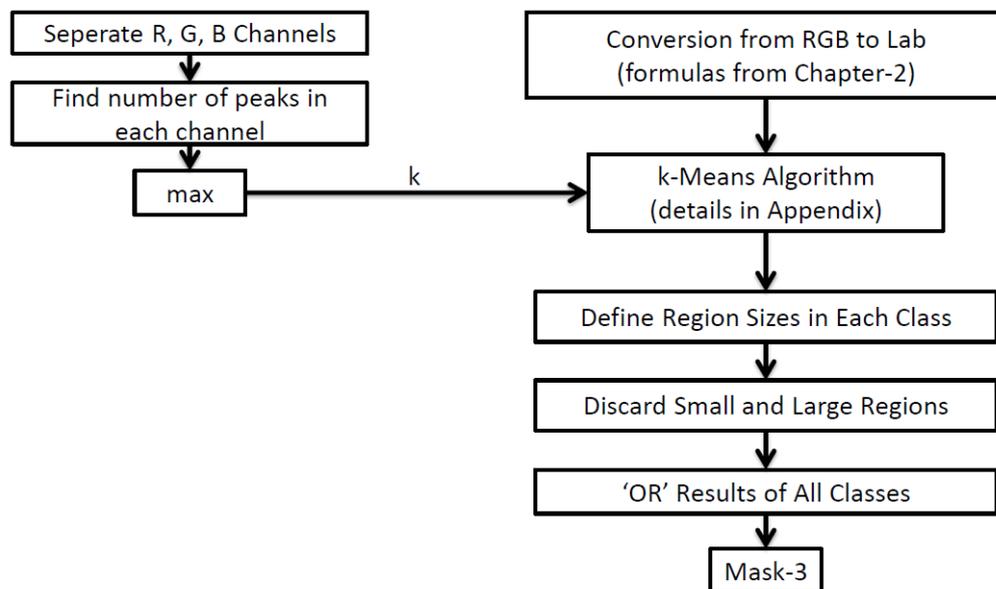
Figure 3.17: The Flowchart of the Third Method

The three methods explained above are used for the input image and their outputs are "OR"ed in order to obtain the resulting mask. Using three different methods provides stability for testing different cameras. In the experiments, two different cases are realized. In the first one, one of the test methods can be enough to localize the color patches. An example of the case is given in Figure 3.18. The results of the methods from 1 to 3 are given on the upper part from left to right and the original image and the resulting mask is given on the lower part of the figure. As can be seen, the result of the third method is almost same as the resulting mask. However, that one method

shows variety between the cameras. For example, the first method can find every color patch in the image taken by another camera.
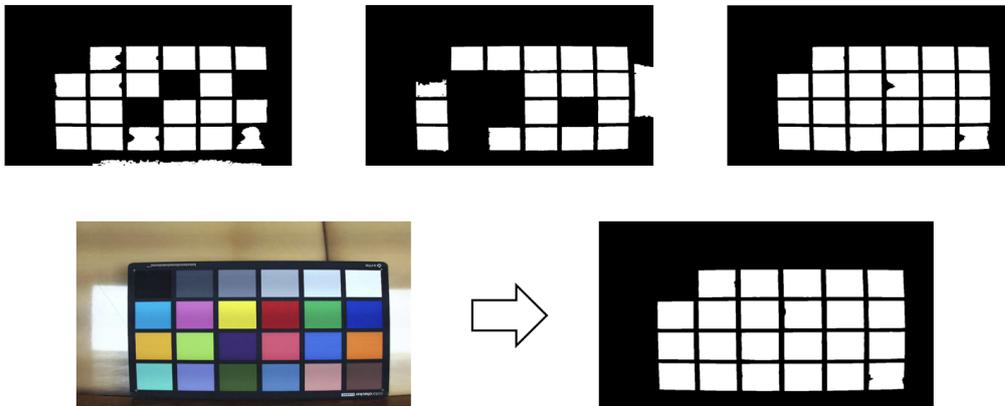
Figure 3.18: The Results of Different Methods-1

In the second case, each method finds the different combinations of the color patches but at the end, each color patch is marked by at least one method. An example of the case is given in Figure 3.19. As shown, the results of three methods complete each other. Thus, "OR"ing the results of the different methods helps to obtain all color patches.
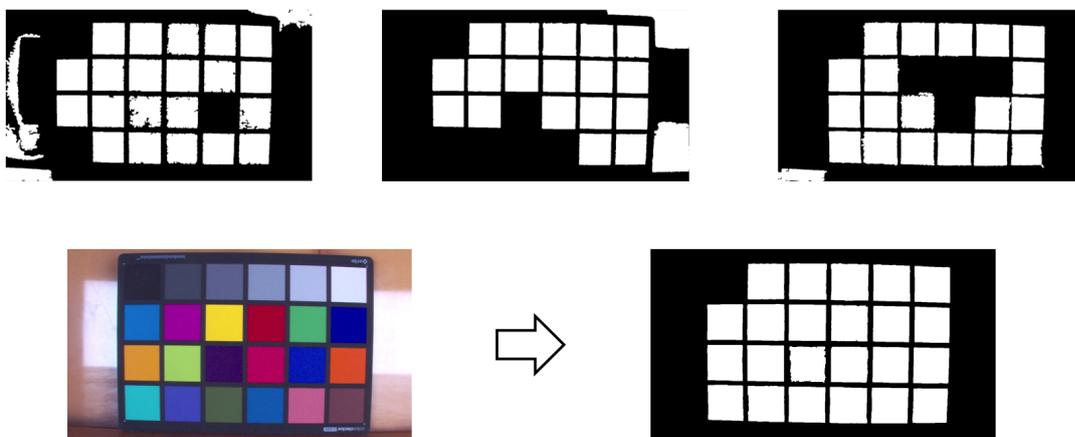
Figure 3.19: The Results of Different Methods-2

However, there can be still some regions as the part of the background and they can be detected as color patches in the resulting mask. In order to prevent them, with a simple assumption of that background would be a lighter color than the black frame of the test chart, the mask of the whole color chart is prepared by using the first method

mentioned above as shown in Figure 3.20. "AND"ing this mask with the resulting image of the 3 methods results with only the color patches in the test chart except for the black patch.
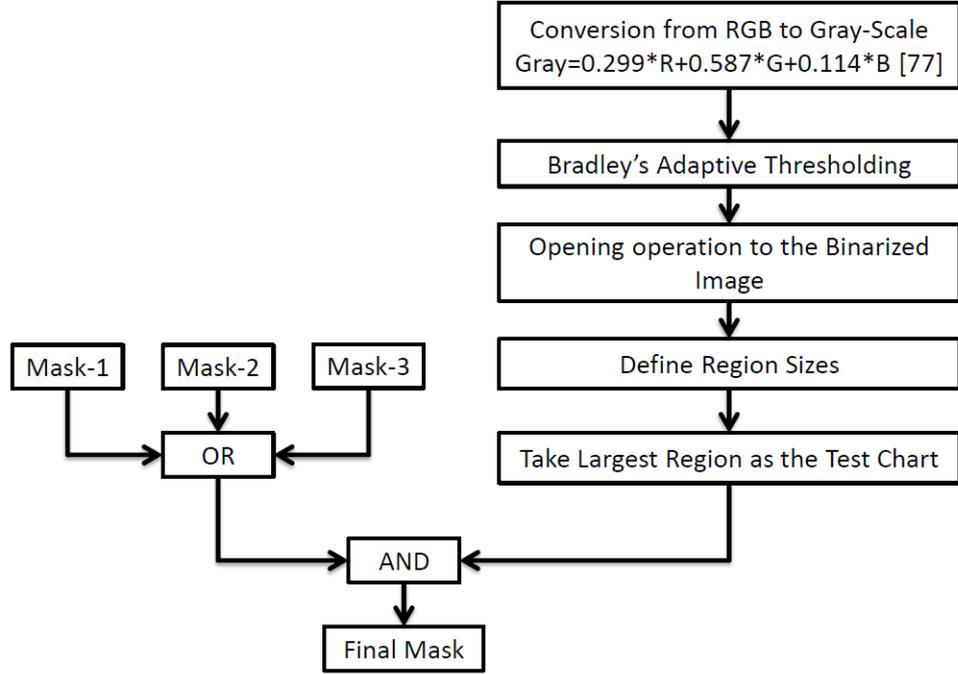
Figure 3.20: The Final Step of Mask Formation

Due to the difficulty of the separating it from the black frame of the test chart, it is excluded from the rest of the measurements. Also, the location of the black patch is used as an indicator for the orientation of the test chart and it is supposed to be located on the upper left corner of the image. When the color patches are determined, the histograms in each patch are obtained. By taking the maximum value in the histogram as the reference, Delta E values in each color patch,$DeltaE_{patch,i}$ values, are calculated by using Formula 3.2. Also, the standard deviation in each color patch is calculated with the formula below because it is also an indicator of the noise level of the camera. Please note that $\mu^c_{patch,i}$ is the mean value of the ith patch for the channel c which can be L, a and b. N represents the number of pixels in that color patch and I^c_u is the pixel values in the related channel c.

$$\mu^c_{patch,i} = \frac{1}{N}\sum_{u=1}^{N} I^c_u \tag{3.7}$$

51

Then, the standard deviation value of the ith patch for the channel c, $S^c_{patch,i}$, is calculated as shown.

$$S^c_{patch,i} = \sqrt{\frac{1}{N-1} \sum_{u=1}^{N} (|I^c_u - \mu^c_{patch,i}|)^2} \qquad (3.8)$$

The average standard deviation of the ith patch, $S_{patch,i}$, is calculated as follows.

$$S_{patch,i} = \frac{1}{3} \sum_{c=L,a,b} S^c_{patch,i} \qquad (3.9)$$

Lastly, the average standard deviation, σ is calculated with the formula below by using the standard deviation values from each color patch.

$$\sigma = \frac{1}{23} \sum_{i=1}^{23} S_{patch,i} \qquad (3.10)$$

Similarly, the average Delta E value, μ_{DeltaE} is calculated as follows by using the Delta E value in each patch, $DeltaE_{patch,i}$.

$$\mu_{DeltaE} = \frac{1}{23} \sum_{i=1}^{23} DeltaE_{patch,i} \qquad (3.11)$$

The smaller values for both μ_{DeltaE} and σ values indicate better performance of the camera for the color accuracy test.

3.5 Test Results

The explained test procedure has been used in order to test and compare the color accuracies of the 8 cameras. Only the input images taken by the cameras and the related test results will be given in this part.

3.5.1 The results of Camera-1

The input image acquired by the Camera-1 can be seen in Figure 3.21.

Figure 3.21: The Test Chart Image taken by Camera-1

3.5.2 The results of Camera-2

The original image taken by the Camera-2 is given in Figure 3.22. Let's take the results of Camera-1 and Camera-2 in order to compare their color accuracy performances. As mentioned before, the smaller Delta E value shows better color accuracy performance. So, Camera-2 has a higher color accuracy performance according to Camera-1 and also their standard deviations are close to each other which also show that their noise immunities are close to each other. As a result, just according to this test, using Camera-2 is a better choice for the end user.

Figure 3.22: The Test Chart Image taken by Camera-2

3.5.3 The results of Camera-3

The image of the test chart obtained by using the Camera-3 can be seen in Figure 3.23.

Figure 3.23: The Test Chart Image taken by Camera-3

3.5.4 The results of Camera-4

The input image taken by the Camera-5 is given in the Figure 3.24.

Figure 3.24: The Test Chart Image taken by Camera-4

3.5.5 The results of Camera-5

The image acquired by the Camera-5 is shown in Figure 3.25.

Figure 3.25: The Test Chart Image taken by Camera-5

3.5.6 The results of Camera-6

The ColorChecker test chart is recorded by Camera-6 as shown in Figure 3.26.

Figure 3.26: The Test Chart Image taken by Camera-6

3.5.7 The results of Camera-7

The taken image of the test chart by using the Camera-7 is given in Figure 3.27. According to μ_{DeltaE} values, the Camera-7 performs the best color accuracy among the others.

Figure 3.27: The Test Chart Image taken by Camera-7

3.5.8 The results of Camera-8

The image of the test chart acquired by the Camera-8 is shown in Figure 3.28.

Figure 3.28: The Test Chart Image taken by Camera-8

Table 3.1 gathers the results of the dynamic range tests for eight test cameras.

3.6 Test Results by Imatest Software

There already exist some companies which present image and video quality tests for their customers like Imatest, Image-Engineering GmbH, and ImageLabs Inc. Imatest is one of the best of these companies in this field and both their test charts and performance measuring software are the most commonly used ones in the market. We consider that the comparison of our test results with a verified and industrialized prod-

Camera No	μ_{DeltaE}	σ
1	26.66	1.05
2	19.65	1.1
3	26.19	0.69
4	16.79	1.68
5	26.42	1.42
6	19.02	1.08
7	12.65	0.85
8	23.43	1.24

Table 3.1: The Results of the Color Accuracy Tests

uct is important in order to prove the parallelism in the results because of the same ideas behind the test procedure. The color accuracy test of Imatest measures μ_{DeltaE} values of the color patches as the common criterion with our proposed test. The main difference is the selection of the region of interest. While the color patches are found automatically by using our software, the user has to indicate the color patches by hand from the user interface of the Imatest software. The results from our software and Imatest software for eight test cameras are shown in Figure 3.29. It can be seen that the results are close to each other and the performance ratings according to μ_{DeltaE} values are the same. The reason for the difference can be the effect of the black patch which is excluded from the results in our test procedure.

3.7 Environment Change and Results

As explained before, the color of a reflective object depends on the illumination on its surface. So, in the case of using another light source other than the ones with 5000K color temperature as we used in our test setup, the reference values provided by the X-Rite Incorporated would not be applicable. Due to the fact that we do not have reference values which belong to each illumination conditions, using the explained color accuracy testing procedure above does not provide useful information for dif-
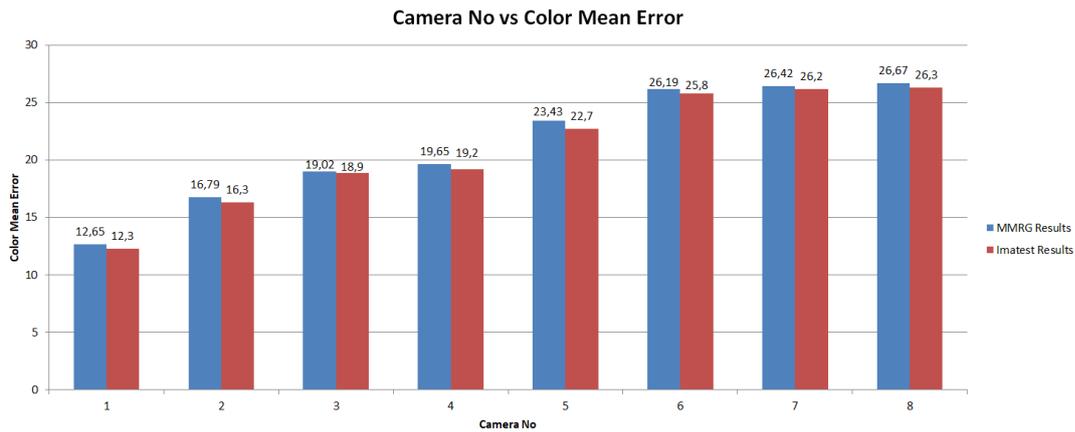
Figure 3.29: The Comparison of Our Test Results (left) and Imatest Software results (right)

ferent environments. In this point, it is important to think about the color constancy phenomenon which described in the background information part. According to this phenomenon, the change of the illumination affects the view of the camera and the human visual system differently. While the colors acquired by the camera changes clearly by the change of the illumination; thanks to the perception process of the human brain, the change in the illumination does not affect the colors seen by the human eye.

The fascinating function of the camera is being the digital replacement of the human visual system. As explained in the background information part, the cameras imitate almost all of the functions of the human eye. However, the processing of the human brain has not been discovered that clearly. On the other hand, also cameras perform some post-processing operations in order to imitate the human brain like noise canceling and flicker prevention. Moreover, as mentioned above, there are some techniques which use Fourier transform, derivatives or neural networks in order to guess the illumination in the scene and perform color correction for ensuring color constancy. Thus, we proposed a modification on the test procedure for measuring the color constancy of the cameras.

The test setup which tries to imitate the environment of the tunnels was proposed in the previous chapter. For the same test setup, the ColorChecker test chart is photographed by three different test cameras. In this test setup, the high pressure-vapor

sodium lamb is used and its color temperature is given as 2000K in its datasheet. By using the images taken by the different cameras, the locations of the color patches are determined by using the same software in the previous part. Then, the maximum values of L, a and b components and variance values of the color patches are calculated. Unlike the previous test, not the reference values of X-Rite Incorporated but the L, a and b values determined in the previous test for the corresponding color patch by using the same camera are used in the calculation of the Delta-E criterion. The comparison of Delta-E values gives a good insight about how these cameras are working suitable to the color constancy concept by applying some post-processing. These results are important because the performance of a camera cannot be considered by just its lens and sensor. The post-processing done in the camera has also an important role in the last image which is the key for the end user. Due to the fact that the cameras are the replicates of the human visual system, they should also work suitably to the whole visual system together with the processing of the human brain. Teh color constancy is one of these effects of the processing of human brain and thus having a small difference for the different illumination conditions for the same object is also desired property for a camera. In this part, the images taken by the three cameras will be given and the results will be gathered by a graph.

3.7.1 The results of Camera-3

The input image acquired by the Camera-3 can be seen in Figure 3.30.

Figure 3.30: The Test Chart Image taken by Camera-3

3.7.2 The results of Camera-4

The ColorChecker test chart is recorded by Camera-4 as shown in Figure 3.31.

Figure 3.31: The Test Chart Image taken by Camera-4

3.7.3 The results of Camera-5

The image of the test chart obtained by using the Camera-5 can be seen in Figure 3.32.
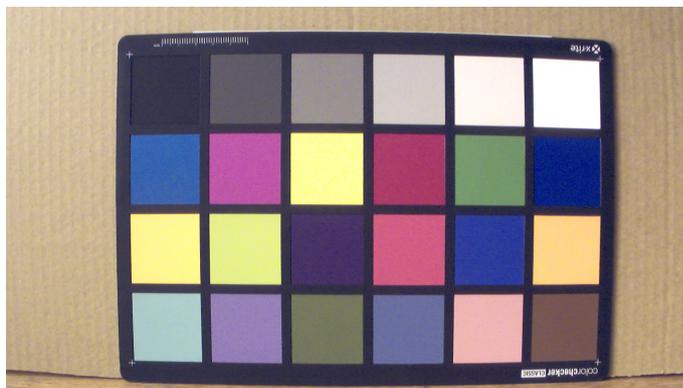
Figure 3.32: The Test Chart Image taken by Camera-5

The color accuracy and color constancy results of these 3 cameras are gathered in 3.33. As observed, when the color accuracy performances are compared, Camera-4 will give the best results, and then Camera-3 will give better results than Camera-5. On the other hand, when the color constancy criterion is considered, Camera-3

will achieve the best results among the others, and then Camera-4 will perform better performance than Camera-5. So, by just checking the color accuracy, it is not possible to judge the total color performance of the cameras and color constancy criterion should be one of these additive concepts to the color accuracy.
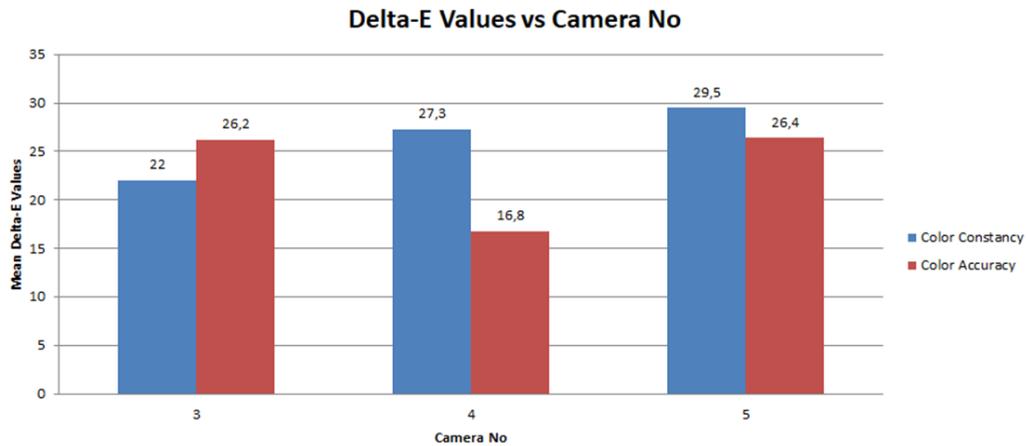
Figure 3.33: The Color Accuracy and Color Constancy Values of Test Cameras

Lastly, the variance values of the cameras change slightly according to the first results due to the fact that the illumination on the test charts is less than the first conditions. The decrease in the illumination increases the noise level in the test results as expected.

3.8 The Color Correction Matrix

As mentioned before, the purpose of the proposing the "Macbeth" color rendition chart is not only testing the cameras but also calibrating them. The calibration can be performed by directly reaching and adjusting the red, green and blue gains of the cameras from their user interface. On the other hand, it is not always possible and easy to do. In such cases, a color correction matrix is the best method to achieve the better results and obtain more accurate colors. Stephen Wolf proposed the Robust Least-Squares Algorithm for color correction matrix in 2003 [83] and it has a common used for digital still and video imaging systems since then. In this part, firstly the details of how the color correction matrix is formed and used will be given and the results for three cameras will be presented later on.

According to this method, firstly a calibration chart with the known values should be obtained like we did in the first test. Then, the algorithm tries to find an optimum solution in order to get each color channel of the reference values by the linear summation of a DC component and all the color components in the obtained image. Due to the fact that there are more than one color patches, the problem is actually an optimization problem. It is important to note that as explained in the background information part, using the CIE Lab color space is more suitable in order to compare the colors. On the other hand, in this method, sRGB color space will be used in order to form the color correction matrix. The reason for this will be given after the details of the algorithm.

The reference values of the color patches provided by X-Rite Incorporated creates the target matrix, O, as shown below. The channels of each color patch fill a row of the "O" matrix and N shows the number of color patches which is 23 in our case.

$$O = \begin{pmatrix} O\_R_1 & O\_G_1 & O\_B_1 \\ O\_R_2 & O\_G_2 & O\_B_2 \\ & . \quad . \quad . & \\ O\_R_N & O\_G_N & O\_B_N \end{pmatrix}$$

Then, in the same way, the L, a and b components obtained from the taken image form the "P" matrix as shown below.

$$P = \begin{pmatrix} P\_R_1 & P\_G_1 & P\_B_1 \\ P\_R_2 & P\_G_2 & P\_B_2 \\ & . \quad . \quad . & \\ P\_R_N & P\_G_N & P\_B_N \end{pmatrix}$$

As mentioned, finding the correction matrix, "A", is an optimization problem which satisfies

$$O \approx O^{'} = \begin{bmatrix} 1 & P \end{bmatrix} A \tag{3.12}$$

So, the fundamental equation for estimating "A" is given below.

$$A = (\begin{bmatrix} 1 & P \end{bmatrix}^T \begin{bmatrix} 1 & P \end{bmatrix})^{-1} \begin{bmatrix} 1 & P \end{bmatrix}^T O \tag{3.13}$$

This function is used in order to find the initial estimation of "A" and in the rest of the algorithm the Delta-E criterion is used as an error function with the channels R, G,

and B instead of L, a and b. Then, a cost vector, "C" is generated by using the Delta-E vector,"E" and a small epsilon which is recommended as 0.1 such that it prevents division by zero.

$$C = \frac{1}{(E + \epsilon)} \quad (3.14)$$

The cost vector is normalized by dividing each element of C with the square root of the sum of squares of all the elements. Then, the element-by-element square of the cost vector C is calculated. The NxN diagonal cost matrix (C^2) which is formed by the cost vector's elements on the diagonal and zeros everywhere is used in the following formula in order to perform cost-weighted least-squares fitting on the next estimate of correction matrix, "A".

$$A = ([1 \quad P]^T C^2 [1 \quad P])^{-1} [1 \quad P]^T C^2 O \quad (3.15)$$

Then, the Delta-E is calculated again with the new values of $O^{'}$ and the procedure is repeated until the elements of color correction matrix converge to four decimal places.

For three of the test cameras, the correction matrix is prepared and the results will be given in the next parts. However, before passing that part, it is necessary to explain the color space choice. As mentioned above, sRGB color space is used instead of CIE Lab color space which is normally the one that the color accuracy is calculated. Stephen Wolf noted that his method for color correction matrix can be used for any color spaces. On the other hand, using the color patches which has saturated channel components, such that any of R, G and B value is greater than 253, in the color correction matrix calculation gives wrong correction matrix. If this correction matrix is applied to the original image, then the saturated patches look very strange [84]. Working on sRGB color space enables us to discard the corresponding rows of these saturated regions from "O" and "P" matrices. In such cases, the color correction matrix may not perform well but at least it does not create any strange regions on the image. On the other hand, in the standard of CIE Lab color space there is no limit for a and b components. Even though the limit between -128 to 127 is taken into the consideration as indicated in [85], the results did not give better results than the result obtained by sRGB color space.

3.8.1 The results of Camera-3

The original image taken by Camera-3 and its color corrected version can be seen in Figure 3.34.
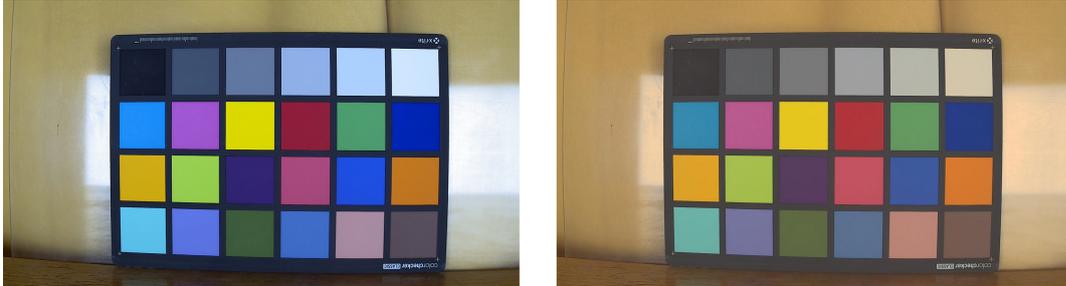
Figure 3.34: The Original Image and the Color Corrected Image for Camera-3

3.8.2 The results of Camera-4

The acquired image by Camera-4 and the resulting image when the color correction matrix is applied are shown in Figure 3.35.

Figure 3.35: The Original Image and the Color Corrected Image for Camera-4

3.8.3 The results of Camera-5

The original image and color corrected image for Camera-5 can be seen in Figure 3.36.
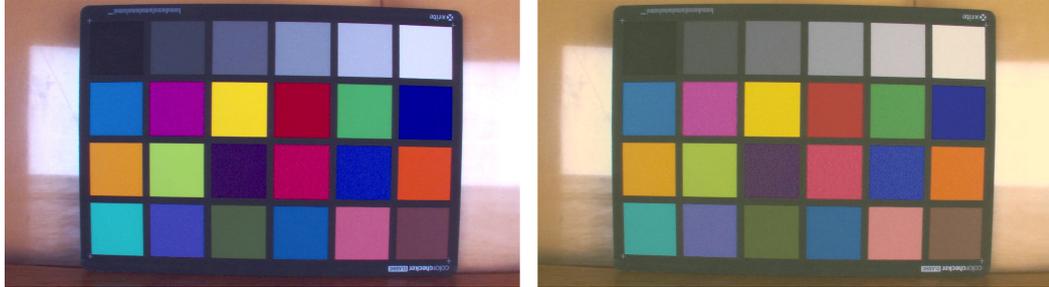
Figure 3.36: The Original Image and the Color Corrected Image for Camera-5

The effect of the color correction on the μ_{DeltaE} values can be observed in Figure 3.37. Actually, the color correction and the white balancing performance of the camera are highly related. A camera with good white balancing performance can correct also the colors in the scene according to the illumination conditions in the environment. If the correction performed in this case is assumed as linear, the best correction which can be provided is the output of the implemented color correction method. In other words, the corrected images represent the best possible white balancing outputs of the cameras. So, by looking at Figure 3.37, it can be said that the camera sensor performances of the test cameras are very close to each other.
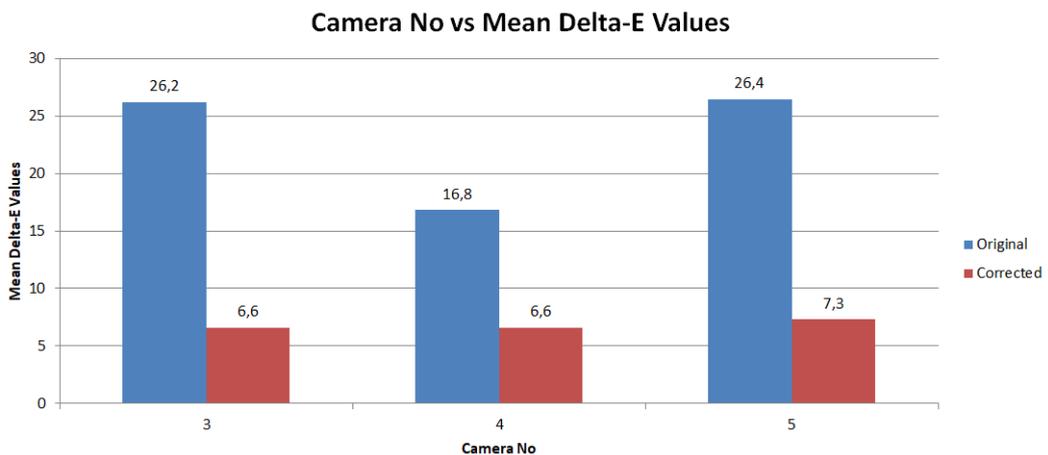
Figure 3.37: The Original and Color Corrected μ_{DeltaE} Values

CHAPTER 4

CAMERA TESTS ON SHARPNESS AND RESULTS

4.1 Background Information about Sharpness

Sharpness is one of the most important features of the image for photographers because it represents the number of details that can be stored in the image [29]. On the other hand, without testing the imaging system, it is hard to evaluate because of the subjective evaluations of the different individuals [20]. In this chapter, the terms of acutance and resolution which affect perceived sharpness will be explained and; the objective measurement of sharpness will be discussed.

Acutance describes how much the transitions are represented clearly at the edges of the image [22]. If the transition has sharp and well-defined borders, then it has high acutance. Figure 4.1 shows the differences in the high and low acutance concepts on a simple edge. So, the acutance decreases from the stepwise transition to ramp-wise transition as shown.
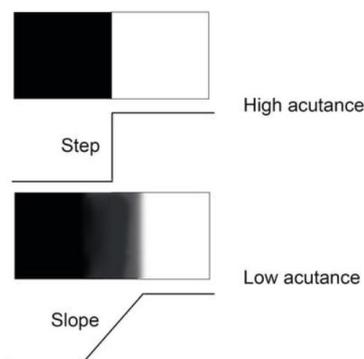
Figure 4.1: The High and Low Acutance Edges [20]

67

Although the acutance is very important for the sharpness, it is impossible to have a perfect step transition by using a camera due to the imperfections of the lens and the sensor. Even the most expensive lens in the market cannot be perfect as it designed and; due to the fabrication processes, the light which is distracted while passing through the lens will create a blurred image on the image plane. The lens is one of the most critical components for the sharpness concept; thus its performance is also measured alone by the manufacturers. The method of how it is measured will be given later. As explained in the previous section, the color formation is performed by using a Bayer filter in front of the camera sensor. Then, the acquired data by the analog to digital converters are used as input to the demosaicing algorithm. However, the effects of the filter, analog to digital converter and the demosaicing algorithm create some blur on the image. On the other hand, there are some post-processing methods that can be used in order to increase the acutance. Thus, the total performance of the camera cannot be understood by only checking one of these components.

The one way to measure the acutance is the rise distance of the edge from the pixel values which go from 10 % to 90 % of the final value. However, this method is not commonly used in the literature, and sometimes,it does not provide enough information due to the post-processing done by the camera. Due to the fact that the human visual system tends to perceive the sharpness higher at the edges with high gradient, some post-processing methods are used inside the cameras. The unsharp masking is one of the most popular of these methods. The more detailed information will be given in the Appendix part about the unsharp masking but its effect can be observed in Figure 4.2.

When the edge points are examined carefully, it is seen that edge profile in the processed image is more stepwise with respect to the original one. On the other hand, there are overshoots near the edges which actually increase the gradient and convince human visual system to be sharper. The existence of those overshoots makes the usage of the rise distance metric hard to measure. The edge profiles are given in Figure 4.3.

Resolution describes the ability of the camera to distinguish between closely spaced elements. The most critical limiting effect on the resolution of the camera is the

Figure 4.2: The Original Image (left), Slightly Unsharp Mask Applied Version (middle), Strong Unsharp Mask Applied Version (right) [21]
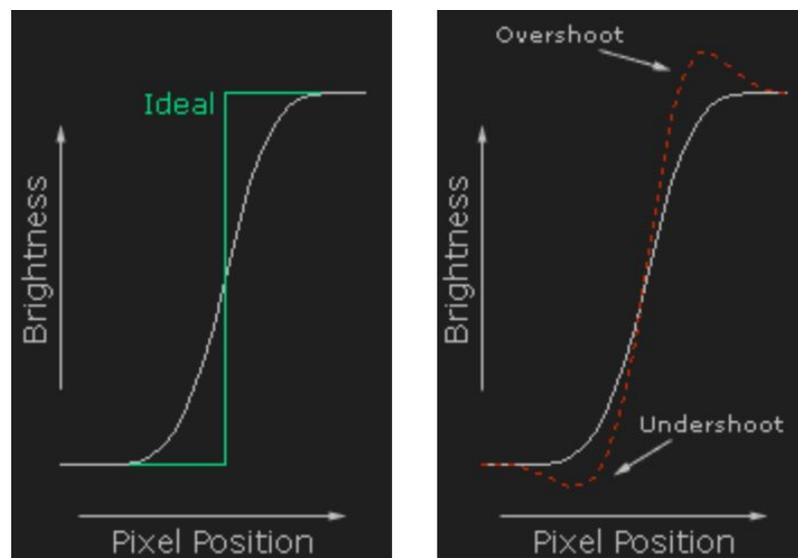
Figure 4.3: The Edge Profiles of Ideal, Original and Processed Images [22]

diffraction of the light when it faces an obstacle [23]. Due to the wave nature of the light, the diffraction occurs when the light passes through a circular opening, for example, aperture of the camera, as shown in Figure 4.4.

This diffraction effect is first discovered by George Biddell Airy in 1835 [86] and thus, the pattern obtained on the image plane is called as Airy disc pattern. The Airy pattern can be obtained with the squared modulus of the Fourier Transform of the circular aperture. The final mathematical expression of the Fourier Transform is given below for $I_0 = \frac{P_0 * A}{\lambda^2 * R^2}$ where P_0 is the total power from the light source, A is the area of the aperture, λ is the wavelength of the light and R is the distance from

69
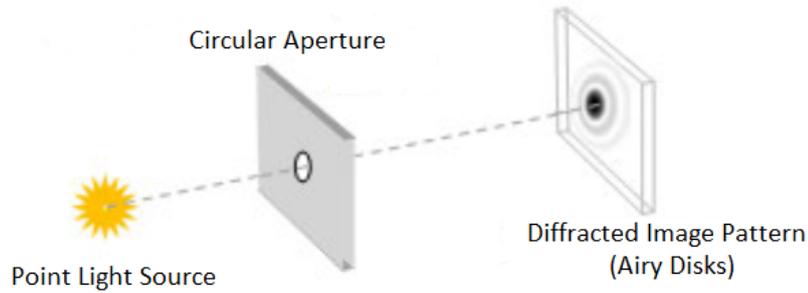
Figure 4.4: The Diffraction of Light Passing Through Circular Aperture [23]

the aperture. Due to the property of Bessel function, the discs around a center are observed as shown in Figure 4.5.

$$I(\theta) = I_0(\frac{2J_1(x)}{x})^2 \quad where \quad x = \frac{\pi * q}{\lambda * N} \tag{4.1}$$

It is important to note that q is the radial distance from the optic axis in the observation plane and N is the is the ratio of the observation distance to the aperture diameter.
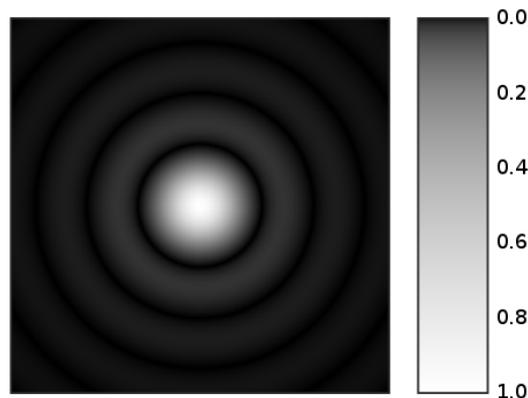
Figure 4.5: The Airy Discs [24]

If two point light sources are considered instead of one, the existence of the Airy Discs obviously limits the resolution of the system because if two light points are very close to each other, their projections in the image plane become unresolvable. Thus, the minimum distance between the separable two point light sources can be used as an indicator of the resolution of the system. This phenomenon is firstly studied by Lord Rayleigh in 1879 and "Rayleigh Criterion" has been proposed in order to calculate the

70

theoretical minimum distance of resolvable two point light sources [87]. According to this criterion, the light sources can be separable with minimum distance when the maximum of the Airy pattern of one of them is on the top of the first minimum of the other Airy pattern. Figure 4.6 shows the details about Rayleigh Criterion in 1D.
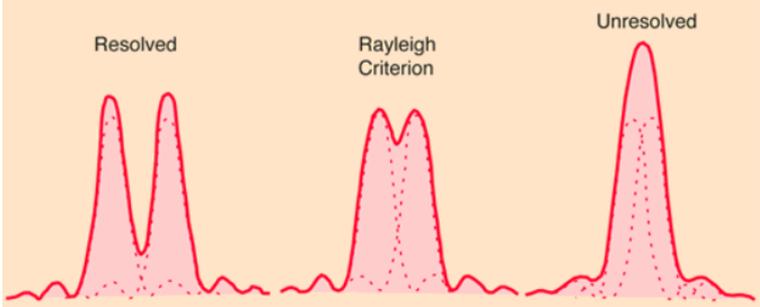
Figure 4.6: The Rayleigh Criterion [25]

Mathematically, the theoretical smallest separation mentioned above is satisfied with the formula below where X is the separation, f is the focal length, lambda is the wavelength of the light and d is the diameter of the aperture [88].

$$X = 1.22 * \frac{(\lambda * f)}{d} \tag{4.2}$$

On the other hand, due to the imperfections of the lenses, all the cameras have lower resolution limit than this theoretical value. Although the sensor and the software affect the resolution of the camera too, it is hard to compensate for the difference between high-quality and low-quality lenses. Hence, the lens manufacturers test their products and modulation transfer function is the most commonly used metric while comparing the performance of the lenses. The modulation transfer function of lenses is measured by using special testbeds as shown in Figure 4.7.
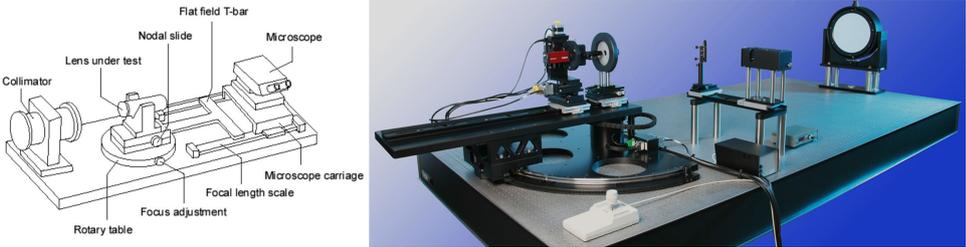
Figure 4.7: The Testbed for Measuring Resolution of Lenses [26], [27]

Although all the details are given until this point based on the point source, these testbeds use a slit in front of the lens in order to measure the resolution of the lens under test. The same idea is also used in the method of slanted-edge which will be given later. The light sent from the collimator passes through a slit and reaches to the lens. It is distracted while passing through the lens and the light intensity on the image plane has more spread than the original split. The intensity on the image plane is used in order to find the line spread function and the real part of its Fourier transform, Modulation Transfer Function (MTF), gives the information about the resolution of the lens. The results of two different lenses can be observed in Figure 4.8. There are some companies which provide testbeds and test software for their customers like Trioptics GmbH [89] and Image Science Ltd [27].
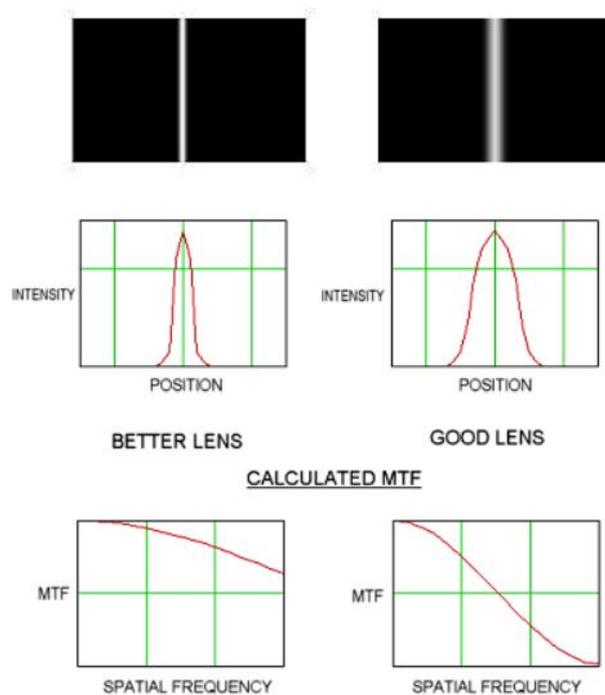
Figure 4.8: The Splits on the Image Plane, Intensity Values and MTFs of Two Different Lenses [28]

If the device under test is a camera instead of the lens, a set of black and white lines are used in order to indicate the resolution. As expected, if the pairs of lines are placed very close to each other, the camera would not be able to differentiate them. The better understanding can be provided by the Figure 4.9.
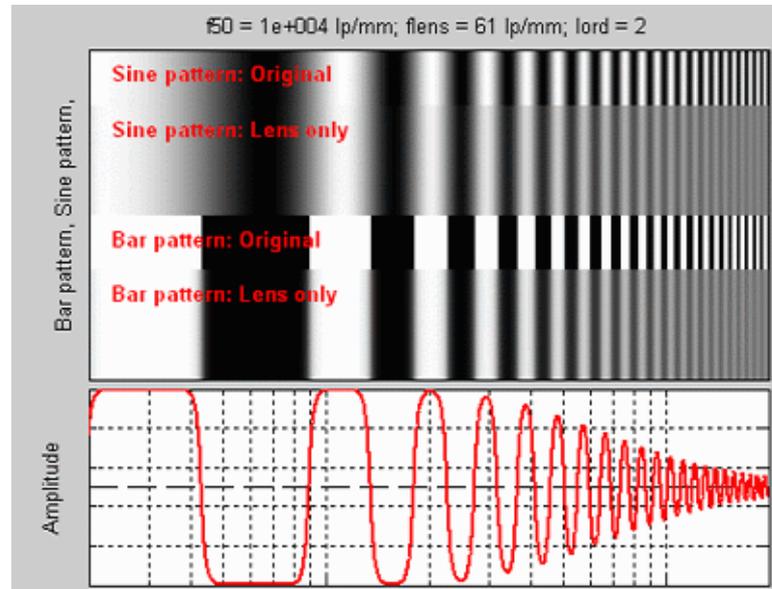
Figure 4.9: The Sine and Bar Patterns with Pixel Amplitudes [29]

As shown in Figure 4.9, the sine or bar patterns are used in order to find out the resolution performance of the camera. If the black lines or regions are far away from each other like in the left part of the test patterns which corresponds to low frequency, then the test chart and the image acquired by the camera becomes similar. The amplitude of the pixel values swings between limits, 0 and 255 for 8-bit representation. On the other hand, if lines are close to each other which corresponds to high frequency as shown on the right side of the test pattern, then differentiating the lines becomes harder and harder. This effect can also be observed from the amplitude plot. By increasing the frequency, the difference between the maximum and minimum pixel values becomes smaller. This observation is the fundamental of the modulation transfer function. The modulation transfer function (MTF) is generally identical to the spatial frequency response [1]. It is used as the key for the resolution and thus, sharpness measurements. The modulation transfer function can be obtained from the luminance modulation, 'C(f)', which is also calculated from the amplitude

73

of pixel values, 'V(f)', for the spatial frequency 'f'.

$$C(f) = \frac{(V(f)_{max} - V(f)_{min})}{(V(f)_{max} + V(f)_{min})} \qquad (4.3)$$

$$MTF(f) = \frac{C(f)}{C(0)} \quad where\, C(0) = 1\, theoretically \qquad (4.4)$$

It is important to note that for the analog cameras, the line pair per mm or line per mm is used as the unit for MTF. On the other hand, for digital cameras, line pair (cycle) per pixel is the most commonly used unit for MTF. By using a bar or sine chart as shown in Figure 4.9 and the provided equations 4.3 and 4.4, the MTF of the camera can be obtained as shown in Figure 4.10.
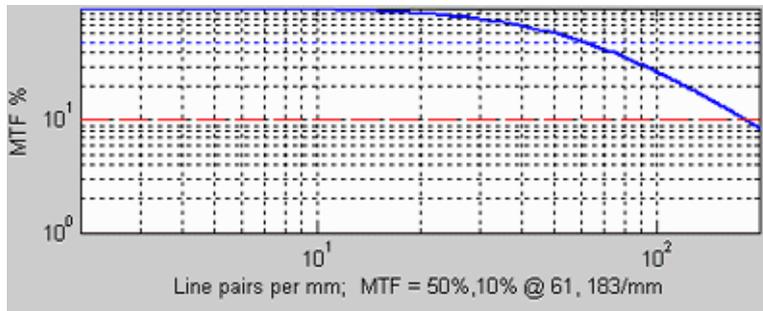
Figure 4.10: An Example of Modulation Transfer Function of a Camera [29]

The other important point to think about is the relationship between the MTF and the acutance. If the acutance of the camera is low, then it is expected that its MTF will approach zero much sooner than a camera with high acutance because it is harder to distinguish the black and white lines even for the low frequencies due to the larger transition areas between fully black and fully white regions. For higher frequencies, the transition regions cover the maximum and minimum areas and dominate the MTF. So, MTF can be used as an objective metric which includes both resolution and acutance concepts.

Although a method how to measure the MTF by using a sine or bar test pattern is explained above, it is not the only way to calculate the MTF. It is clear that using sine or bar pattern is very sensitive to the noise and thus the results can be affected dramatically for high frequencies. Due to this consideration, a more stable method is proposed by "ISO 12233- Photography- Electronic still picture imaging- Resolution

and spatial frequency responses" which is called as slanted-edge method [31]. Unlike the sine/bar pattern method, the slanted-edge method uses a single edge captured by the camera under test. Instead of increasing the frequency of test pattern, all the calculations are performed by the means of Fourier transform in the light of Rayleigh Criterion.

Hence ISO 12233 standard is the commonly accepted way to measure the MTF of the camera and its performance is proven with the well-defined test procedure, it is chosen for our test of sharpness measurement. As mentioned above, the optic lens is the most critical component of the camera which creates the blur on the image. On the other hand, slanted edge method treats the cameras as a system because only lens performance cannot represent the total sharpness performance of the camera. In other words, the results of the tests also cover the sensor and the software effects like unsharp masking and demosaicing algorithm and this ability is also a reason to choose the slanted edge method. The commonly-accepted implementation of ISO 12233 which works on MATLAB is prepared by Peter Burns who was a member of ISO 12233 committee [90], [34]. In our sharpness test, his method is used without change. Thus, the fundamentals of the procedure will be explained in "Test Details" part but the implementation details will not be provided.

Although MTF is used in our camera sharpness performance test, it is important to mention about the Subjective Quality Factor (SQF) and the acutance metric of Camera Phone Image Quality (CPIQ) group. These are two measures of perceived image sharpness and both include the MTF of the imaging system, the Contrast Sensitivity Function (CSF) of the human visual system, image display height and viewing distance. Historically, the idea of the taking the human visual system into account was proposed by Granger et al in the 1970s [91]. Like any camera, also the human visual system has a MTF as shown in Figure 4.11. Due to the effect of the lens in the eye, it is expected to see the decrease in the high-frequency details like every camera. However, unlike the typical behavior of the cameras, the human brain processes the data coming from the human eye and filters the low-frequency changes as well. Thus, after a certain value, the sensitivity for the low frequency also decreases. Mainly, the human visual system has a peak at 6 cycles/degree. This response of the human visual system is called as Contrast Sensitivity Function (CSF) in the literature.
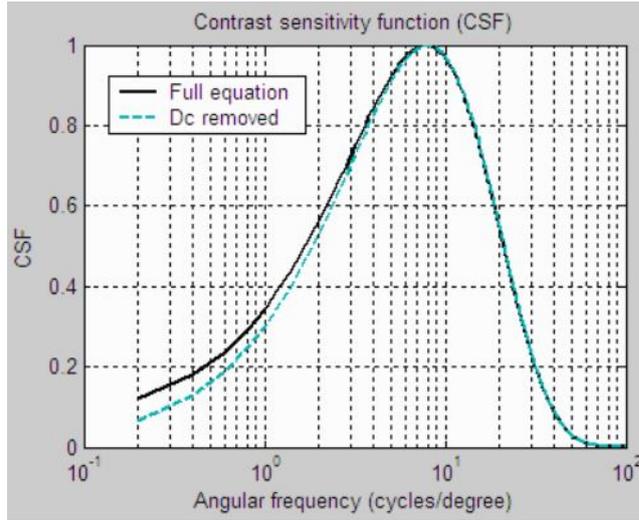
Figure 4.11: The Contrast Sensitivity Function [30]

On the other hand, the full equation of the CSF is a little bit controversial. Actually, Granger [91] used a print and asked about its sharpness to the viewers. Furthermore, due to the calculation complexity at that time, they used a simplified version of the collected data while calculating the SQF. Mannos and Sakrison proposed the formula below for the CSF which is also shown in Figure 4.11 as "Full Equation" [92].

$$CSF(f) = 2.6 * (0.0192 + 0.114f) * e^{-(0.114*f)^{1.1}} \qquad (4.5)$$

However, Imatest GmbH proposed to eliminate the DC term in that equation by resulting in a small change as shown in Figure 4.11. This simplification provides stability to the SQF calculation by preventing the blowing up of the integral because of the DC term. The resulting equation which is used by Imatest GmbH is given below.

$$CSF(f) = 0.114f * e^{-0.1254f} \qquad (4.6)$$

Lastly, the SQF equation is given below. Although the SFQ was used by Kodak and Polaroid Companies in product development tests, it is not a commonly used metric today. Thanks to the standardization by CPIQ, acutance metric is preferred instead of the SQF.

$$SQF = K \int_0^\infty \frac{CSF(f) * MTF(f)}{f} \mathrm{d}f \qquad (4.7)$$

$$where \quad K = \frac{100\%}{\int \frac{CSF(f)}{f} \mathrm{d}f} \qquad (4.8)$$

76

The formulization of the acutance metric is similar to the SQF but it is well defined and given below. The Contrast Sensitivity Function formula gives slightly different results according to the Figure 4.11.

$$Acutance = \frac{\int_0^\infty MTF(v) * CSF(v)\mathrm{d}v}{\int_0^\infty CSF(v)\mathrm{d}v} \tag{4.9}$$

$$where \quad CSF(v) = \frac{75 * v^{0.8} * e^{(-0.2v)}}{34.05} \quad cycles/degree \tag{4.10}$$

It is important to note that the procedure for calculating the MTF explained above provides the MTF value in the units of cycles/pixel. However, obviously, for acutance and SQF, the conversion from cycles/pixel to cycles/degree should be performed with the following formulas where η_{PH} is number of vertical pixels along picture height in landscape orientation, d is viewing distance and PH is the picture height.

$$f(cycles/degree) = \frac{f(cycles/pixel) * (\pi * \eta_{PH} * d)}{180 * PH} \tag{4.11}$$

As mentioned in the previous chapter, determining the performance of the camera according to the human visual system is important if the cameras are thought of as the replacement of the human visual system. However, we decided to prefer to use MTF as the performance criterion instead of acutance metric of CPIQ and SQF. The reason behind is that the results of the acutance and SQF are strongly dependent on the test chart size and the viewing distance. As explained in [28], due to the characteristic of the CSF, for different sizes of the test chart or for different viewing distances, the same cameras can give different results. For example, while Camera A performs better results than Camera B for the same test chart from 30 cm viewing distance, Camera B might perform better results than Camera A for the same test chart with 45 cm viewing distance. Thus, using the acutance metric of CPIQ and SQF is not possible to use as an objective metric without knowing the application details and environment conditions. On the other hand, the MTF gives objective and stable results which are invariant to the test chart size and viewing distance.

4.2 Test Chart

The oldest test chart in the literature which is used for the measurements of the modulation transfer function of the cameras is 1951- USAF test chart. It is shown on the left side of Figure 4.12 and at that time, it was used as a transmissive test chart. This test chart is used in order to compare the lenses by a user/operator and it is hard to measure the MTF directly because the film which the test chart is printed on also contributes the MTF.

The most commonly used test chart in the literature for the measurement of the sharpness of the still images is the ISO 12233:2000 test chart which is shown on the right side of the Figure 4.12. It was proposed in 2000 by the acceptance of the ISO 12233 standard and since then, it has been widely used by the scientists and photographers. Although it was marked as legacy after the revision of ISO 12233:2017 standard last year, it is the test chart used in our performance test. There are two main reasons for this. Firstly, the calculation procedure of the MTF stays almost the same between the two versions of the standard. Secondly, there are lots of software and test results that can be used in order to compare our results whether there is a problem or not. These factors lead us to use this legacy test chart which would also provide good knowledge of the MTF measurement procedure.
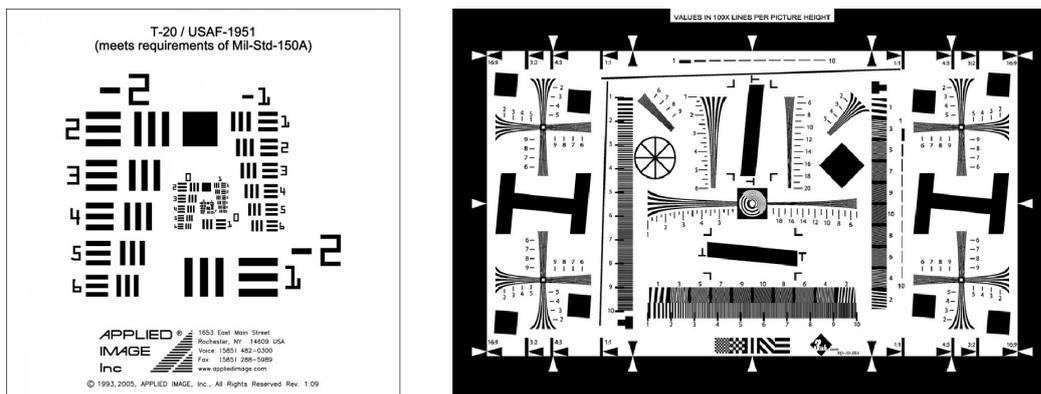
Figure 4.12: The 1951- USAF Test Chart (left) [28] and ISO 12233:2000 Legacy Test Chart [31]

The revision of ISO 12233:2017 proposes three different charts in order to measure the MTF of the camera under tests. The test chart shown on the left side of Figure

4.13 is for the usage of the slanted edge method of the MTF calculation which is the same method used in our performance test. The experiments show that the MTF changes according to the place of the slanted edge on the screen due to the physical properties of the lenses. Thus, 9 identical slanted edges are placed separate from each other and this enables to measure the MTFs for slanted edges distributed to the scene which is photographed. In our tests, the slanted edge is placed into the middle for all the cameras and thus, the change in the MTF is tried to be eliminated. The second test chart proposed by ISO 12233:2017 is shown in the middle of Figure 4.13. It has 1 vertical, 1 horizontal and 2 diagonal bar patterns. How the MTF can be calculated by using a bar pattern was explained before and the same procedure is followed also with this test chart. As different from the legacy chart, diagonal bar pattern provides better information about the MTF of the camera, because MTF changes according to the organization of the bar pattern. In other words, by using the same camera with horizontal, vertical and diagonal bar patterns gives different results. Thus, this test chart includes 4 different bar patterns that can be used in order to measure the MTF. The third test chart is shown on the right side of the Figure 4.13 which is also called as Siemens star. The idea of the Siemens star is the same as the sine and bar pattern. The sine pattern is applied from the outer radius of the circle to the center of the test chart. So, in order to go from low frequency to high frequency, the experimenter should go from outside to the center. The orientation problem of the bar pattern is not observed when all the points on the concentric circles are examined.
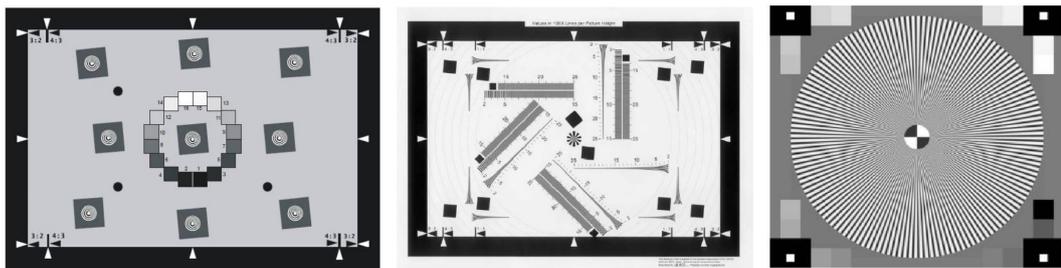
Figure 4.13: The Test Charts of ISO 12233:2017 [32]

Although the Siemens star test chart which is another test chart for sharpness measurement is not used in the scope of this thesis, giving some implementation details might be useful. The procedure is explained as simple and clear in [33]. Firstly, the test chart is photographed as shown in Figure 4.14. Choosing the center of the

Siemens star which is the red point in Figure 4.14 is very important because the rest of the work is done by taking this point as the center of concentric circles which are the green circles shown in Figure 4.14. From the outside of the chart towards the center, for each concentric circle, the number of pixels and the number of complete cycles are known. So, checking the amplitude on the same cycle gives the MTF value at that frequency. The only problematic point is the very low frequency (almost DC) response of the camera. This problem is solved with the little circle in the middle of the test chart, around the center. By using the black and white samples, the MTF values are normalized and the final result is obtained. The most critical and hard to provide point is that the camera lens should be parallel with the test chart. Otherwise, the concentric circles would not fit theoretical data, even if the center of the test chart is found correctly. This situation gives wrong measurements and thus, the MTF would not be valid.
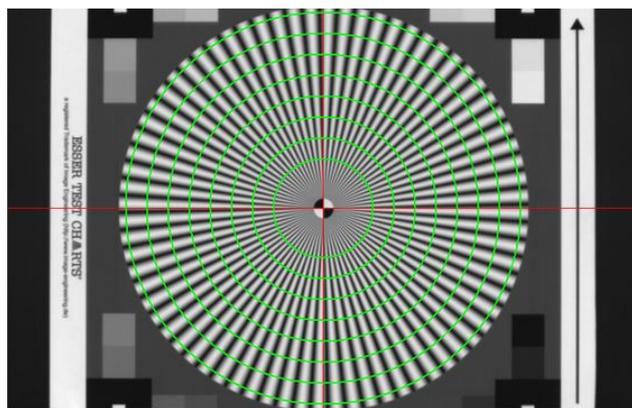
Figure 4.14: A Siemens Star Test Chart [33]

4.3 Test Setup

The test setup which is proposed for the first proposed test is also used for the sharpness measurements. The details of the test setup were given in the previous chapter and thus, it will not be repeated here. As stated before, the provided conditions are very good approximations of the standard viewing. Furthermore, like the previous tests, the test chart covers more than 50 % of the captured image in our test setup such that the slanted edge would place in the middle of the scene. Moreover, the

test chart is placed almost parallel to the camera lens in order to prevent any glare in the image and the focusing operation is performed before acquiring the images. The distance of 45 cm is a reasonable viewing distance for the test chart even if the MTF results are not dependent on this distance. After all the conditions are satisfied, the image is obtained and the test which will be covered in the next part is performed.

4.4 Test Details

As stated before, the ISO 12233:2000 test chart is used in order to measure the MTF of the camera as the sharpness metric. For this purpose, the slanted edge method proposed by ISO 12233 is used. Although the implementation of the method provided by P. D. Burns is used directly, the region of interest is required to be chosen by the user in the original version. The main purpose of this thesis is preparing the automatic camera performance test as explained before. Thus, pre-processing of the image taken from the camera should be performed in order to automate the process. At the end of the pre-processing, the region of interested (the slanted edge) is obtained and the MTF calculation method works only on this patch. After the details about finding the region of interest are explained, the necessary information about the procedure of measuring MTF by slanted edge method will be given.

The image is acquired from the camera by using the same way as the other proposed tests. Due to the fact that the pixel numbers of the cameras are different in the market, the resizing operation without distorting the width/height ratio is necessary in order to perform the pre-processing operations on the same base. The ISO 12233:2000 test chart has lots of high-frequency patterns, and thus it is easy to eliminate most of them by closing and opening operations which are explained in the Appendix part. The pre-processing procedure will be explained by using the image acquired by Camera-3. The original image captured by Camera-3 and its binary version after these operations can be seen in Figure 4.15. Then, the regions which are larger and smaller than certain levels are eliminated. The filtered version is shown on the left side of Figure 4.16. As stated before, the slanted edge should be located in the middle of the screen in order to provide the same conditions between the cameras. Thus, at this point, the object which is the closest one to the center of the image can be marked as the slanted

edge. However, in order to ensure that the closest area to the center is the slanted edge, one more step is added for stability. In this step, the filtered image is taken and the eccentricities of the objects are calculated. The eccentricity of the object is determined by the ratio of the foci of the object with its major axis length. So, the value of the eccentricity changes between zero and one. If the eccentricity is close to zero, the object has a circular shape. Similarly, if it is close to one, this shows that the object looks like a line. The more mathematical details can be found in Appendix. Thus, in this step, the objects which have less eccentricity than a certain level is also excluded from the resulting image. The example image can be seen on the right side of Figure 4.16.
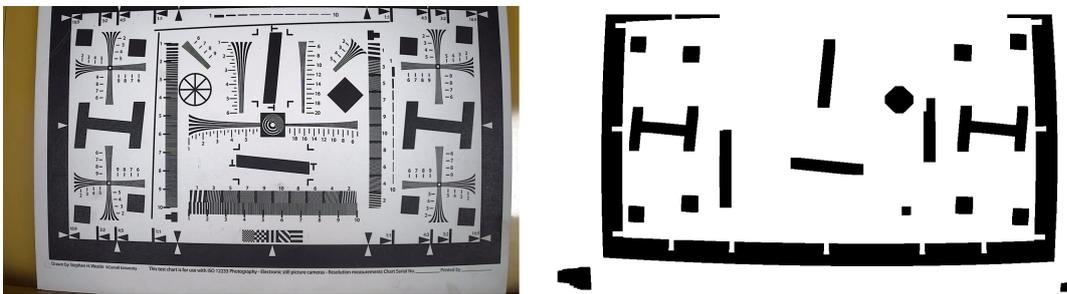
Figure 4.15: Original Image (left) and Its Binary Version with Morphological Operations
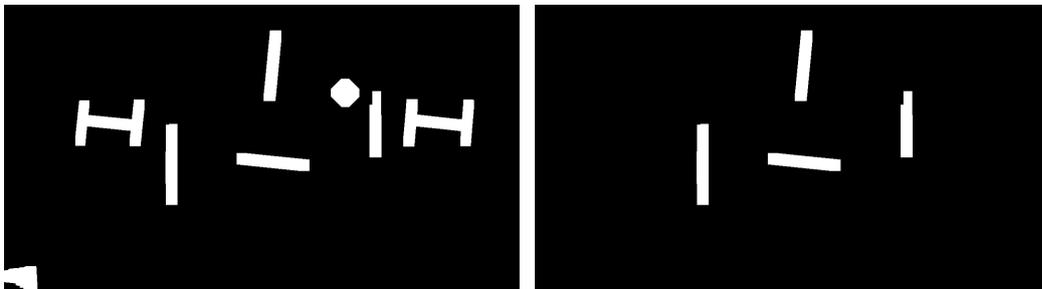
Figure 4.16: The Area-Size Filtered Image (left) and the Eccentricity Filtered Image (right)

The filtered image according to the size and eccentricity values is now ready to use for determining the object closest to the center of the image. Actually, due to our future work plans, not only the closest but also the second closest object is determined. The second closest object is also a slanted edge which can be used for MTF calculation, but in this thesis, only the vertical one is used. The usage of the second slanted edge

82

will be covered in the Conclusion part. The vertical slanted edge is cropped such that it can be used directly by the MTF calculation method of Burns. It is important to note that the determined area is the black part of the slanted edge. However, the MTF calculation requires the transition from white to black for each column. Hence, the cropping is performed as it is shown on the left side of Figure 4.17. The width and height of the area are used for the determination of the cropping points. Moreover, the cropped slanted edge patch is shown on the right side of Figure 4.17. The cropped image is used for the MTF calculation.
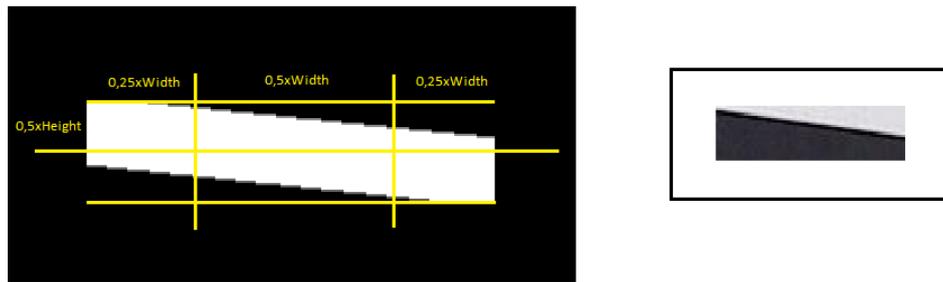
Figure 4.17: The Cropping Details (left) and the Cropped Image (right)

The block diagram of the implemented code for the calculation procedure of MTF is given in Figure 4.18. Before starting to explain the implementation steps, it is obvious that the code requires the selection of the region of interest which needs to cover the slanted edge. It is the most problematic part of the sharpness test if the full automation is aimed. Hence, it will be the main concern of this proposed test and the details of how to find the location of the slanted-edge will be covered later in the Test Details part. At this point, it is assumed that the slanted edge is successfully chosen and given to the implemented code as an input. If the opto-electrical conversion function (OECF) of the camera is known, then it can be the second input of the procedure. Otherwise, tonal response is assumed as linear. If the input has R, G and B channels, then the luminance (Y) is calculated with the formula given in the previous chapter. The Hamming Window is applied on each channel in order to increase the noise immunity. Then, the finite-difference scheme of [-1, 0, 1] is used for finding the gradient [93]. The centroid for each line is determined and a linear equation is found which fits the centroid locations. The linear equation and the location of the centroids are used for the binning algorithm. The centroids are separated into 4 groups according to the fractional errors between the real centroid location and
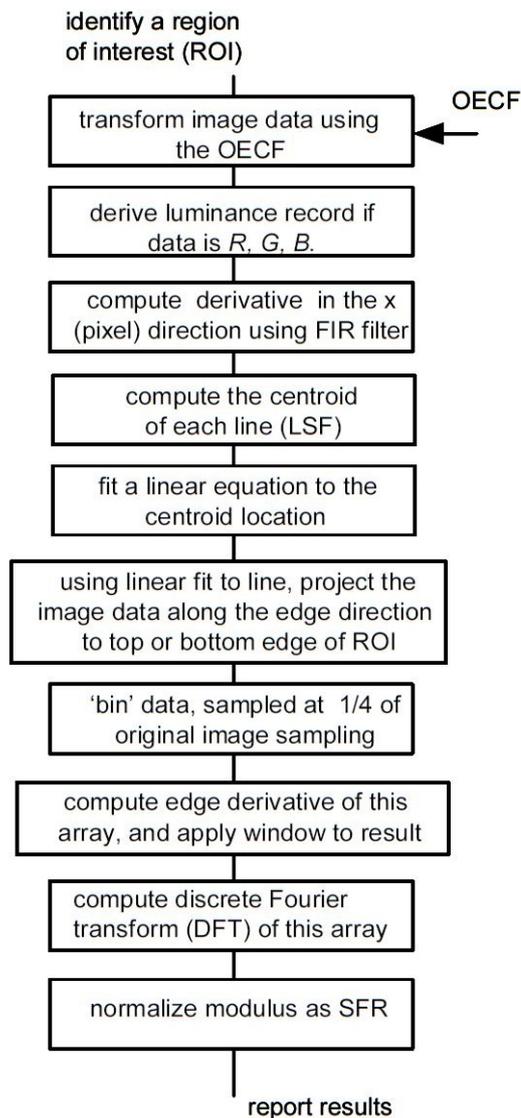
Figure 4.18: The MTF Calculation Procedure Implemented by Peter Burns [34]

its corresponding value determined by the linear equation. If the fractional difference between the centroid and the corresponding point on the linear equation is less than 0.25, the centroid is placed into "Bin 1". If it is larger than 0.25 and less than 0.5, it is placed into "Bin 2". Same wise, if the fractional part is between 0.5 and 0.75, the centroid is placed into "Bin 3" and if it is larger than 0.75, then it is placed into "Bin 4". The four bins are combined to calculate an averaged 4x oversampled edge. This method enables the measurement of the MTF beyond the normal Nyquist frequency. Again, the derivative of the edge is calculated and Hamming Window is used to the result. It is called the Line Spread Function. The absolute value of the Fourier Trans-

form of the Line Spread Function gives the Modulation Transfer Function of the total imaging system.

The MTF curves of the cameras indicate the sharpness performance of the cameras. For the comparison, the most suitable value is obtained at the frequency value which has the 50 percent of the MTF. This point is called as MTF50 point and used as a key to sharpness performance by Imatest GmbH [29]. Thus, it is the numerical value which is used for the comparison also in this thesis.

4.5   Test Results

The explained test procedure has been used in order to test and compare the sharpness performance of the 8 cameras by using their MTF curves. The input images taken by the cameras and the related MTF graphs will be given in this part.

4.5.1   The results of Camera-1

The input image acquired by the Camera-1 and the MTF graph can be seen in Figure 4.19. As shown in Figure 4.19, the MTF curve takes values greater than 1 which is not possible without using a post-processing method like unsharp-masking. Checking Figure 4.3 in this manner explains the situation. So, it is possible to understand whether the camera performs post-processing in order to increase the sharpness by looking at the MTF graph of it.
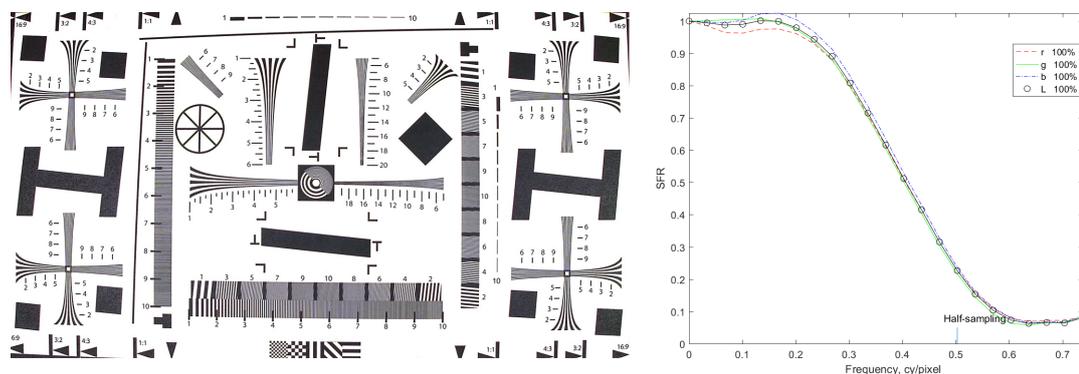
Figure 4.19: The Original Image from Camera-1 (left) and Its MTF Graph (right)

4.5.2 The results of Camera-2

The photographed ISO 12233:2000 chart by Camera-2 and the calculated MTF graph are shown in Figure 4.20. The processor inside the camera performs edge sharpening and this post-processing action can fail for some frequency values and create valley like shown in Figure 4.20.
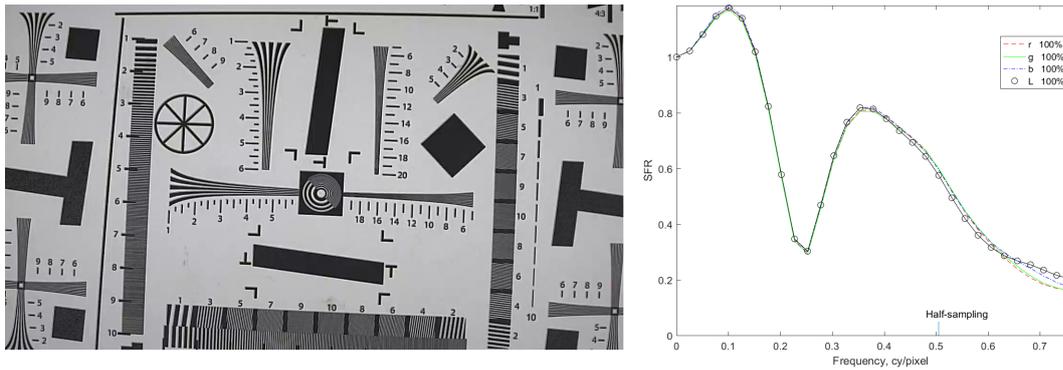
Figure 4.20: The Original Image from Camera-2 (left) and Its MTF Graph (right)

4.5.3 The results of Camera-3

The captured image by Camera-3 and its corresponding MTF graph is given in Figure 4.21. The effect of the post-processing for increasing the sharpness can be recognized.
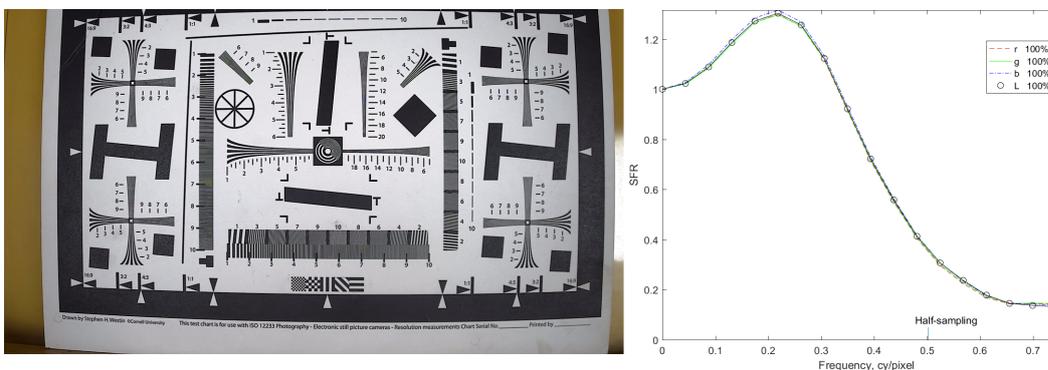
Figure 4.21: The Original Image from Camera-3 (left) and Its MTF Graph (right)

4.5.4 The results of Camera-4

The original image acquired by Camera-4 and the corresponding MTF graph can be seen in Figure 4.22.
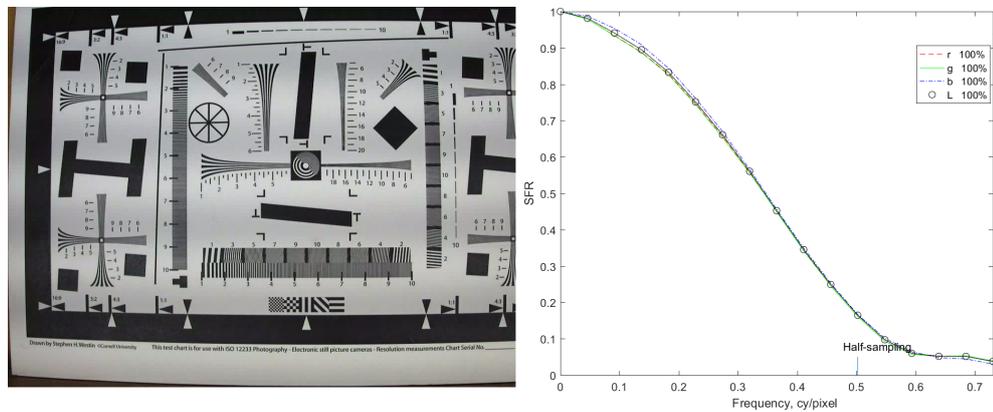
Figure 4.22: The Original Image from Camera-4 (left) and Its MTF Graph (right)

4.5.5 The results of Camera-5

The image of the test chart taken by Camera-5 is shown in Figure 4.23. Also, the MTF curve of Camera-5 is given in the same figure.
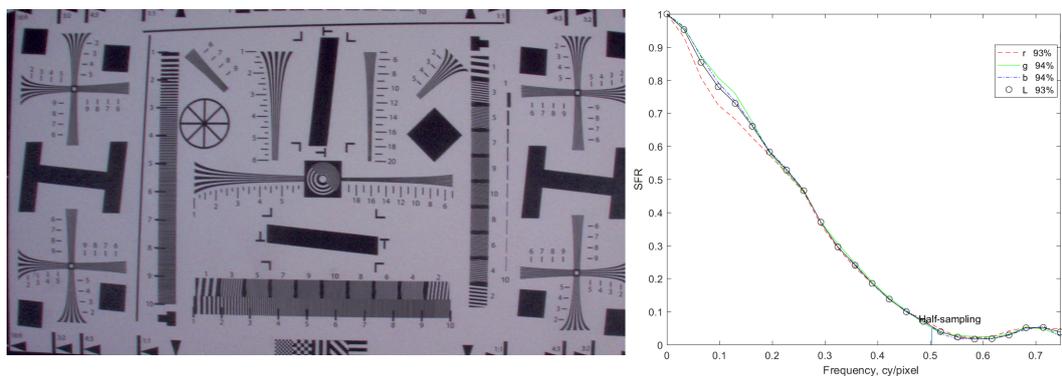
Figure 4.23: The Original Image from Camera-5 (left) and Its MTF Graph (right)

4.5.6 The results of Camera-6

The acquired image of the test chart by Camera-6 and its MTF characteristic are given in Figure 4.24. The camera performs post-processing in order to increase the sharpness.
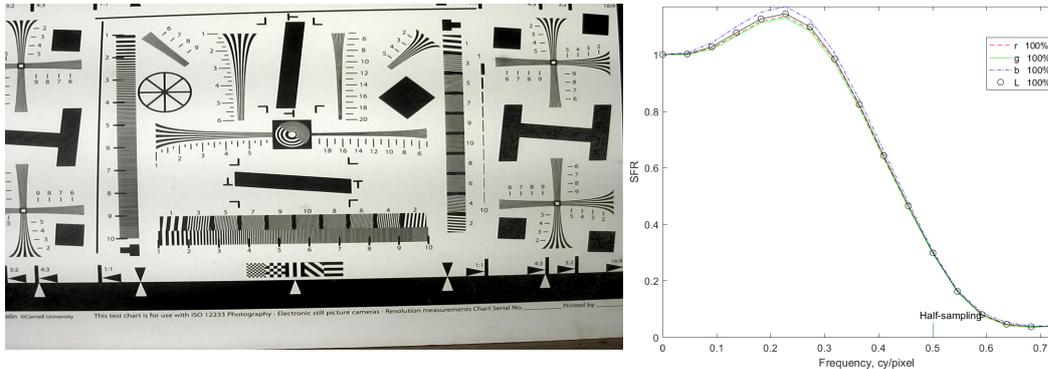
Figure 4.24: The Original Image from Camera-6 (left) and Its MTF Graph (right)

4.5.7 The results of Camera-7

The ISO 12233:2000 test chart is photographed by using Camera-7. The original image and the MTF graph can be found in Figure 4.25. It can be understood from the MTF graph that the camera uses post-processing for increasing sharpness.
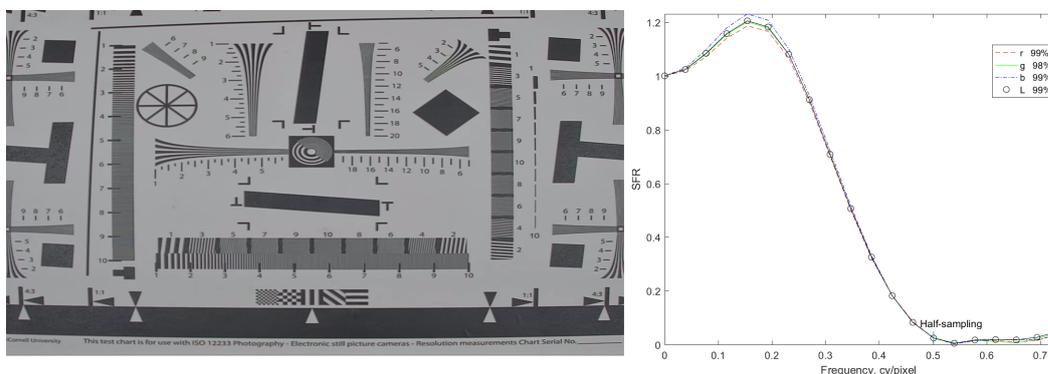
Figure 4.25: The Original Image from Camera-7 (left) and Its MTF Graph (right)

4.5.8 The results of Camera-8

The image of the test chart taken by Camera-8 and its corresponding MTF characteristic are given in Figure 4.26. It can be considered that the cameras perform better sharpness performances when they also apply post-processing. However, as can be observed in this example, although Camera-8 performs post-processing on the taken image, its performance is worse than Camera-4 which does not use any post-processing actions in order to increase image sharpness.
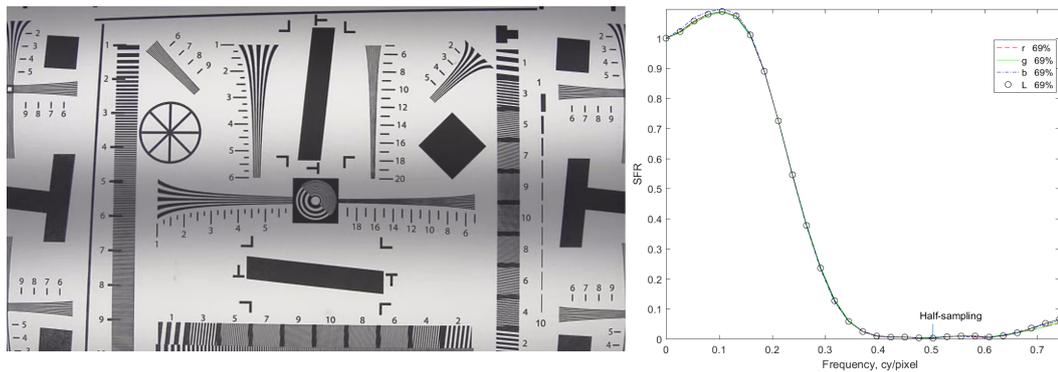
Figure 4.26: The Original Image from Camera-8 (left) and Its MTF Graph (right)

Table 4.1 gathers the results of the dynamic range tests for eight test cameras.

Camera No	MTF50 Value
1	0.405
2	0.209
3	0.452
4	0.344
5	0.238
6	0.445
7	0.347
8	0.246

Table 4.1: The Results of the Sharpness Tests

4.6 Test Results by Imatest Software

As stated before, there are some companies which measures camera performances and Imatest GmbH is one of the biggest of them. Like the previous test procedure, Imatest GmbH also has test software in order to calculate MTF of cameras. The MTF50 value is provided as the test result and it is stated that MTF50 value is the best indicator of the camera sharpness like our test procedure. Thus, it is important to satisfy the consistency between the given results in this thesis and the corresponding results provided by a verified product. Although the slanted edge has to be selected through a user interface manually in Imatest software, it is done by our software automatically. The test results show that there is a consistency between two products. The MTF50 values which are calculated by Imatest software and our test method for eight test cameras are given in Figure 4.27.
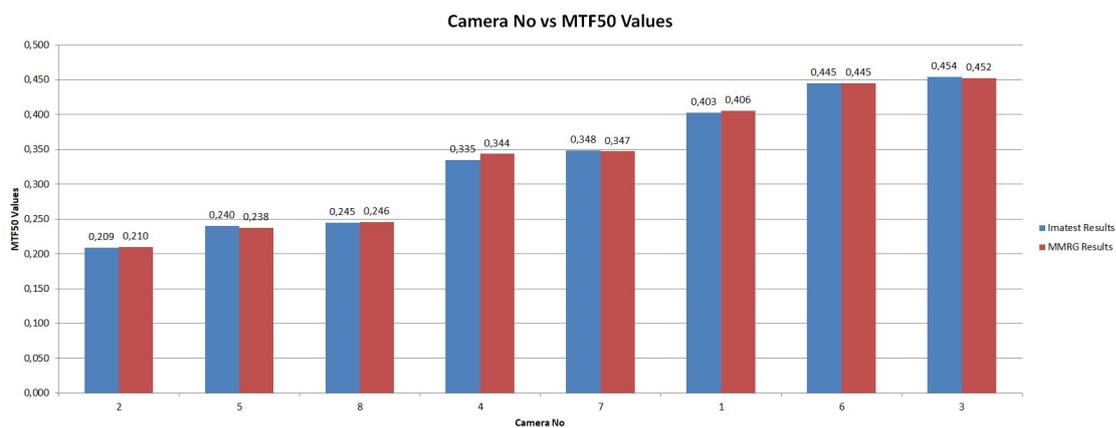
Figure 4.27: The Comparison of Our Test Results (left) and Imatest Software results (right)

CHAPTER 5

CAMERA TESTS ON CHANGE SENSITIVITY AND RESULTS

The proposed tests in the previous chapters are aimed to measure some still image properties of the camera under test. Thus, the images are acquired while both camera and the test chart are not moving. On the other hand, the performance of the camera when there is a change in the scene is also important in order to find the right choice according to the application which it is used in. Hence, in this and the next chapter, two tests will be proposed in order to measure the performance of the camera when there is a change in the test setup. For this purpose, the videos acquired by different cameras will be examined frame-by-frame and their characteristics will be discussed.

5.1 Background Information about Change Detection

The change detection is one of the most critical processing actions performed on the videos acquired by the digital cameras. The performances of the computer vision applications like people counting, action recognition, anomaly detection, room occupancy monitoring, fall detection, parking occupancy detection and tracking depend on their fundamental pre-processing step of change detection [94]. Without a successful change detection algorithm, these computer vision actions cannot be performed reliably because the identification of the region of interest is provided by the change detection algorithm.

Obviously, the motion detection algorithms have more importance for security cameras because the working principles of the security cameras and the other digital cameras show some difference [95]. Although the digital cameras record video after it is enabled until it is disabled manually; the security cameras perform enabling/disabling

the video actions automatically if there is a motion in the scene. The security cameras are always powered up because they should be able to provide video whenever an unusual action happens in the view of the camera. On the other hand, if they provide and save the video all the time while they are powered up, the end user faces with two problems. Firstly, saving that much data in a system requires too much storage. At this point, it is vital to notice that the storage elements are expensive, and thus, it is not practical. Secondly, checking the stored data takes too much time when it is necessary. Using the motion detection algorithms implemented in the camera solves these problems by capturing image or video when there is a motion in the scene [96].

It is important to note that the most of the security cameras require some variables to be adjusted by the user according to the environmental conditions in order to start recording or ignite the alarm outputs. These variables represent some threshold values possibly used in the algorithms implemented in the processor in order to detect the motion. In general, three adjustments are needed to be done for the motion detection [97]. First of all, the region of interest is selected by rectangle because examining the whole scene might be time-consuming and it increases the power consumption. The power consumption is very critical criteria for a device which is always turned on. On the other hand, the region of interest can only be determined by the end user. The second variable which should be adjusted is the percentage of the region of interest in order to trigger an alarm. The smaller motions than this threshold do not trig the alarm. The third variable is the sensitivity of the motion detection algorithm. Although the percentage variable defines the number of pixels for the motion detection, the sensitivity variable determines how much change in the pixel value results with the marking as motion in this pixel. For example, by this adjustment, the alarm trigger caused by the little illumination change due to the headlight of a car can be prevented.

Although the adjustments explained above help to increase the performance of the motion detection performance of the camera, the motion detection algorithms implemented in the high-performance processors of the cameras have some common problems. These problems can be listed as follows [98]:
• Small and distant objects might not be detected.
• Detection accuracy might be affected by weather conditions such as wind, heavy

rain, and snow.

• There are minimum lighting requirements which should be satisfied by the end user.

• For a good motion detection performance, the camera should not be subjected to excessive vibrations.

Because of the listed reasons, the false alarms are triggered. At this point, it should be considered that if the image quality of the camera is not good enough, developing a motion detection algorithm becomes harder for that camera. On the other hand, the alarm trigger software is developed by the software engineers of the camera manufacturer and thus, it seems like it is not the problem of the end user. However, this situation makes impossible to obtain a reliable and robust performance for the different cases. So, the end user faces with problems in reality. The false alarms create the problem of storage and examination as mentioned before; on the other hand, decreasing the sensitivity and percentage variables accordingly results in the missing the real thread. In other words, although fine-tuning of the variables in the algorithm has great importance, a good motion detection algorithm mainly depends on the image quality of the camera. These facts form the idea of using the motion detection algorithms as an image quality metric while comparing the performance of cameras.

As explained above, each security camera has its own motion detection algorithm inside its high-performance processors. However, the output of the algorithm provides a binary output which is called alarm trigger. Due to the fact that the motion masks are not reachable to the end user, it is hard to measure the performance of both motion detection algorithm and the image quality of the camera. Thus, we consider that taking the video of the same action in the same conditions with different cameras and using the same motion detection algorithm on these videos can give a good idea about the performance of the camera for motion detection algorithms. As mentioned above, motion detection is the first step of the other higher level video processing studies. So, the camera achieved better performance for this algorithm can give also better results for the other video processing actions and it also proves the higher image quality. In the rest of this part, some information about motion detection algorithms will be given. It is important to note that we need an algorithm such that we do not need to update the variables for different cameras. So, the algorithm which is proposed by Huang is used in our tests as the examiner and its details will be also mentioned [35].

Before going into the details of the motion detection algorithms, it is crucial to state that the importance of the motion detection algorithm for the security cameras is explained, till this point. On the other hand, the proposed test method does not require that the camera under test is specifically a security camera. So, any camera which can record a video can be used in this test and the result is directly related to its image quality.

A common approach for the motion detection problem is based on the background subtraction, where each frame of the video is compared with a background model. The pixels which have significantly different values from the background model are marked as they belong to the moving object. On the other hand, the background subtraction algorithm has many challenges. Firstly, it should work robust and fast such that it can achieve real-time processing. Secondly, it must have immunity to the change in illumination. Thirdly, it should avoid detecting non-stationary background objects such as swinging leaves, rain, snow and shadow of moving objects. Lastly, the background model should react fast enough to the change in the background [99]. There are many methods in the literature used for background modeling [35].

5.1.1 Simple Background Subtraction

Binary motion detection mask D(x,y) is obtained by subtracting the current frame I(x,y) and fixed background model B(x,y).

$$D(x, y) = \begin{cases} 1 & \text{,if } |I_t(x, y) - B(x, y)| > \tau \\ 0 & \text{,otherwise} \end{cases} \tag{5.1}$$

The simple background subtraction method has the following problems:
- Determining the threshold value, τ,
- Affected by noise
- Non-stationary objects in the background model

5.1.2 Running Average

In this method, the background model B(x,y) is updated by each new frame by using the following formula [100]:

$$B_t(x, y) = (1 - \beta) * B_{t-1}(x, y) + \beta * I_t(x, y) \tag{5.2}$$

Then, the motion detection mask is obtained in the same way as the previous method. On the other hand, this method also has problems of determining the threshold value, τ, and the update coefficient, β.

5.1.3 Sigma-Delta Estimation

The temporal statistics of the pixels of the original video sequence is calculated while the background model is updated by the following formula [8].

$$B_t(x, y) = B_{t-1}(x, y) + sgn(I_t(x, y) - B_{t-1}(x, y)) \tag{5.3}$$

Absolute difference image, $\Delta(x,y)$, and time variance image, $V_t(x,y)$, are also calculated by using the following formulas.

$$\Delta_t(x, y) = |I_t(x, y) - B_t(x, y)| \tag{5.4}$$

$$V_t(x, y) = V_{t-1}(x, y) + sgn(N * \Delta_t(x, y) - V_{t-1}(x, y)) \tag{5.5}$$

Lastly, the motion mask is obtained by using the following formula.

$$D(x, y) = \begin{cases} 1 & \text{,if } \Delta_t(x, y) > V_t(x, y) \\ 0 & \text{,otherwise} \end{cases} \tag{5.6}$$

5.1.4 Simple Statistical Difference Method

In this algorithm, the background image is obtained with the mean of the previous frames, μ_{xy}. The variance value, σ_{xy}, is also calculated and it is used as the threshold

value [101].

$$\mu_{xy} = (1/K) * \sum_{k=0}^{K-1} I_k(x, y) \tag{5.7}$$

$$\sigma_{xy} = \sqrt{(1/K) * \sum_{k=0}^{K-1} (I_k(x, y) - \mu_{xy})^2} \tag{5.8}$$

It is important to note that K represents the frame number in these formulas. As a final step, motion mask is created like the below:

$$D(x, y) = \begin{cases} 1 & \text{,if } |I_t(x, y) - \mu_{xy}| > \lambda\sigma_{xy} \\ 0 & \text{,otherwise} \end{cases} \tag{5.9}$$

As can be seen from the formula, the determination of the predefined parameter, λ, is critical in this method.

5.1.5 Temporal Median Filter

This technique uses 2 timers for short and long-term variances in the background model and updates the long-term and short-term background models according to these timers. If the short-term timer T^S(x, y) is greater than the long-term timer T^L(x, y), then the pixels of B_t^S(x, y) and B_t^L(x, y) are updated by the new incoming frame I_t(x, y) and the short-term timer T^S(x, y) can subsequently be reset to zero. If the long-term timer T^L(x, y) is greater than the tolerance μ, then the long-term timer T^L(x, y) can be reset to μ. The timer threshold, μ, is determined according to the frame rate of the video. The determination of the pixel threshold, τ, is critical like the other methods. The augmentation of the timers is performed according to the following formulas [102].

$$T^S(x, y) = \begin{cases} T^S(x, y) + 1 & \text{,if } |I_t(x, y) - B_t^S(x, y)| \leq \tau \\ 0 & \text{,otherwise} \end{cases} \tag{5.10}$$

$$T^L(x, y) = \begin{cases} T^L(x, y) + 1 & \text{,if } |I_t(x, y) - B_t^L(x, y)| > \tau \\ 0 & \text{,otherwise} \end{cases} \tag{5.11}$$

In his paper, Huang proposed a method for motion detection by trying to unify all these algorithms in one technique. This makes the algorithm robust and reliable. Thus, it is selected as the examiner in our test. The block diagram of the algorithm can be seen in Figure 5.1.
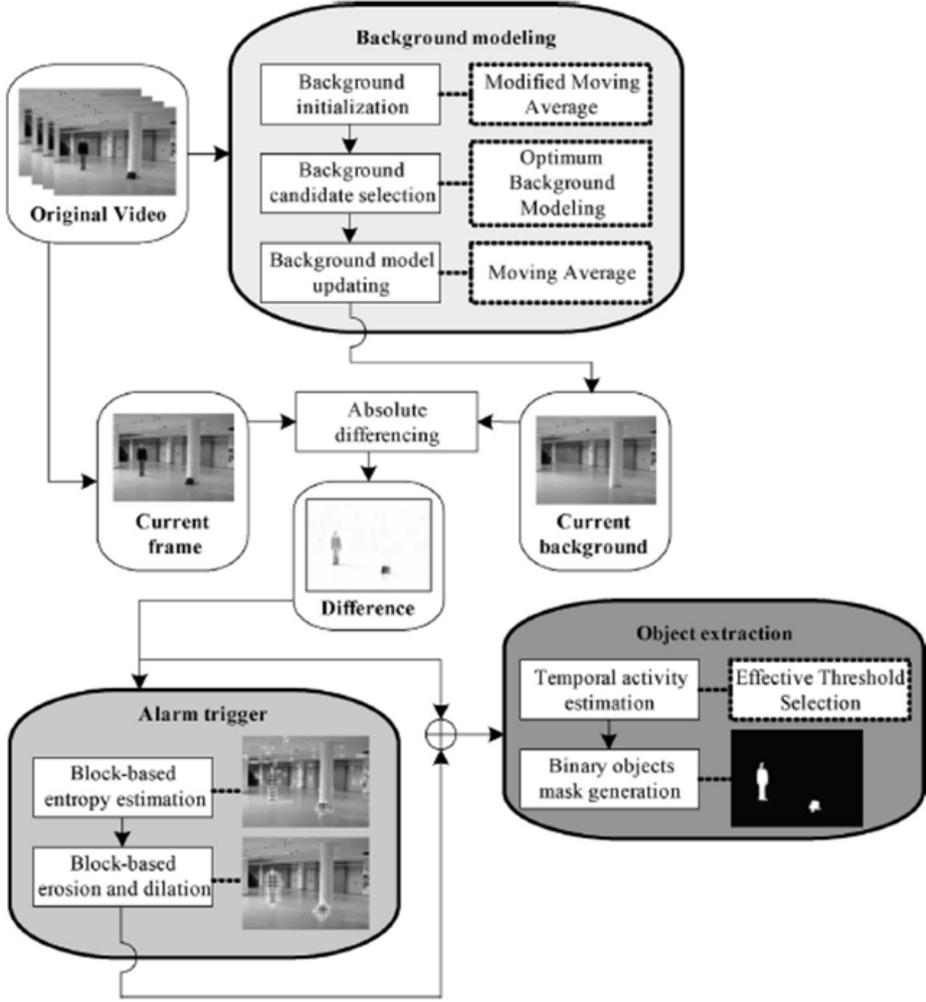
Figure 5.1: The Block Diagram of the Algorithm [35]

As shown in Figure 5.1, the algorithm can be separated into 3 sub-blocks. The first one is background modeling. The background initialization is done by using the formula below for the first 50 frames where t is the frame number in the video.

$$B_t(x,y) = B_{t-1} + \frac{1}{t}(I_t(x,y) - B_{t-1}(x,y)) \qquad (5.12)$$

Background updating is done only for the pixels which are determined as background candidates. Background candidates are determined in 3 steps like follows:

i. The pixel location (x,y) should satisfy the rapid matching condition:

$$I_t(x, y) - I_{t-1}(x, y) = 0 \qquad (5.13)$$

ii. All the pixels from the set of background candidates selected via the rapid matching procedure are then trained through the stable signal trainer (Note that p is predetermined parameter as 1).

$$M_t(x, y) = \begin{cases} M_t(x, y) + p & \text{,if } I_t(x, y) > M_{t-1}(x, y) \\ M_t(x, y) - p & \text{,if } I_t(x, y) < M_{t-1}(x, y) \end{cases} \qquad (5.14)$$

iii. Lastly, the pixels which satisfy I_t(x,y) $-$ M_t(x,y) = 0 are being updated by the formula below where α is predetermined as 8.

$$B_t(x, y) = B_{t-1}(x, y) + 1/\alpha(I_t(x, y) - B_{t-1}(x, y)) \qquad (5.15)$$

The second sub block is alarm trigger module. In this part, Δ_t(x,y) is computed simply by absolute differentiating I_t(x,y) and B_t(x,y). The corresponding Δ_t(x,y) frame is divided into w x w blocks, where w is determined as 8. Then, block-based probability density function, P_h(i,j), is calculated as follows where h (an arbitrary element of L_0, L_1, L_2, L_3 , . . . , L_{v-1}) represents the arbitrary gray-level within each w x w block (i, j) and nh(I,j) denotes the number of pixels corresponding to arbitrary gray-level h. Note that h is reset to 0 when it is smaller than τ, where τ is set to 10.

$$P_h^{(i,j)} = \frac{n_h^{(i,j)}}{w^2} \qquad (5.16)$$

After finding the block-based probability function, block-based entropy, E(i,j), is calculated by the formula:

$$E(i, j) = -\sum_{h=0}^{L_{v-1}} P_h^{(i,j)} log2(P_h^{(i,j)}) \qquad (5.17)$$

According to the threshold T, determined as 1, the first binary mask is obtained like the following.

$$A(i, j) = \begin{cases} 1 & \text{,if } E(i, j) > T \\ 0 & \text{,otherwise} \end{cases} \qquad (5.18)$$

Lastly, the erosion operation by using 1-pixel-disk structure element and then the dilation operation by using 2-pixel-disk structure element is done on image A(i,j). The necessary information about the erosion and the dilation operations can be found in Appendix part.

The last sub-block is the object extraction module. Like in the temporal median filter, short and long time variance images are used where p is indicated in background updating part as 1 and N is determined as 2 in the paper.

$$v_t^s(x,y) = \begin{cases} v_{t-1}^s(x,y) + p & \text{,if } N * \Delta_t(x,y) > v_{t-1}^s(x,y) \\ v_{t-1}^s(x,y) - p & \text{,if } N * \Delta_t(x,y) < v_{t-1}^s(x,y) \end{cases} \tag{5.19}$$

$$v_t^l(x,y) = \begin{cases} v_{t-1}^l(x,y) + p & \text{,if } v_t^s(x,y) > v_{t-1}^l(x,y) \\ v_{t-1}^l(x,y) - p & \text{,if } v_t^s(x,y) < v_{t-1}^l(x,y) \end{cases} \tag{5.20}$$

The best variance value, V_t(x,y), is then calculated like as follows and it is used for creating the resulting binary motion mask, D(x,y).

$$V_t(x,y) = N * min(v_t^s(x,y), v_t^l(x,y)) \tag{5.21}$$

$$D(x,y) = \begin{cases} 1 & \text{,if } \Delta_t(x,y) > V_t(x,y) \\ 0 & \text{,otherwise} \end{cases} \tag{5.22}$$

The explained algorithm is used for each frame acquired by the camera and motion masks are obtained.

5.2   Test Device

As mentioned before, till this part all the test charts are used in order to measure the still image performances of the cameras. On the other hand, in this test, we need a device which can create change in the scene while the camera is recording. For this purpose, a light source is prepared by using a LED lamp inside a cylindrical case. The reflective cover inside the cylindrical case provides almost uniform light intensity coming out from the top of the case. The top is covered by a special transparent plastic

and a piece of A4 white paper in order to prevent the strong illumination change in the scene. The used LED lamp shown in Figure 5.2 is able to change its color and the green color in the maximum intensity is used during the tests. The image of the front view of the test device can be seen in Test Results part.

Figure 5.2: The Used LED Lamp in the Test Device

5.3 Test Setup

The test setup which is proposed before is also used for the image quality test by using motion detection algorithm. The details of the test setup are given in the first chapter and as stated, the provided conditions are very close to the standard viewing conditions. Furthermore, like the previous tests, the test device would place in the middle of the scene. Moreover, the test device is placed almost parallel to the camera lens in order to prevent any glare in the image and the focusing operation is performed before acquiring the video. After all the conditions are satisfied, the video is obtained and the test which will be covered in the next part is performed.

5.4 Test Details

In the proposed test setup, the video is recorded by the camera under test with 25 or 30 frames per second (fps) adjustment. It is important to note that the exposure time selection is left as auto exposure mode for all the test cameras. The more detail about the exposure time and the reason of why it is left in auto mode will be explained in the next chapter. For the fixed exposure time cameras, it can be selected as 20 milliseconds for 25 fps or 16.6 milliseconds for 30 fps. It is also important to note that H.264 is chosen as the video coding standard. As described in the previous chapters, in order to reach a common pixel number, the original frames are resized into a size which has 720 pixels in height without distorting the image ratio. The video is started when the lamp is on and it stays on for approximately 10 seconds. Then, the lamp is turned off and the next 10 seconds is recorded. Then, the lamp is turned on again and the video records for 10 more seconds. By using the motion detection algorithm explained above, the video is examined frame by frame and the motion masks are formed for each frame. Then, from the first frame, a mask which represents the top of the cylindrical case is obtained by using the image processing techniques.

The formation of the mask for the top of the cylindrical case is very important for the result of the test. The results of the motion detection algorithm and the mask are compared pixel by pixel and thus, obtaining a wrong mask would directly affect the result of the test. As mentioned above, in the first frame, the lamp is turned on and by the help of the color of the lamp, the top of the cylindrical case becomes easily separable from the surroundings, the white part of the paper coverage. The color is the main difference with the surroundings, so it is used as the discriminative property. As mentioned in the second proposed test, k-means algorithm is again used in order to separate the regions as shown in Figure 5.3.

Unlike the second test, we have only a few different colors in the scene, using only the k-means algorithm provides enough information in order to form the mask. With the same idea of the second proposed test, the histogram of the R, G and B channels are used. The maximum number of the peak is used as the number of class. Then, the conversion from sRGB color space to Lab color space is performed because it is more convenient to work on the color. For the conversion formula, the second
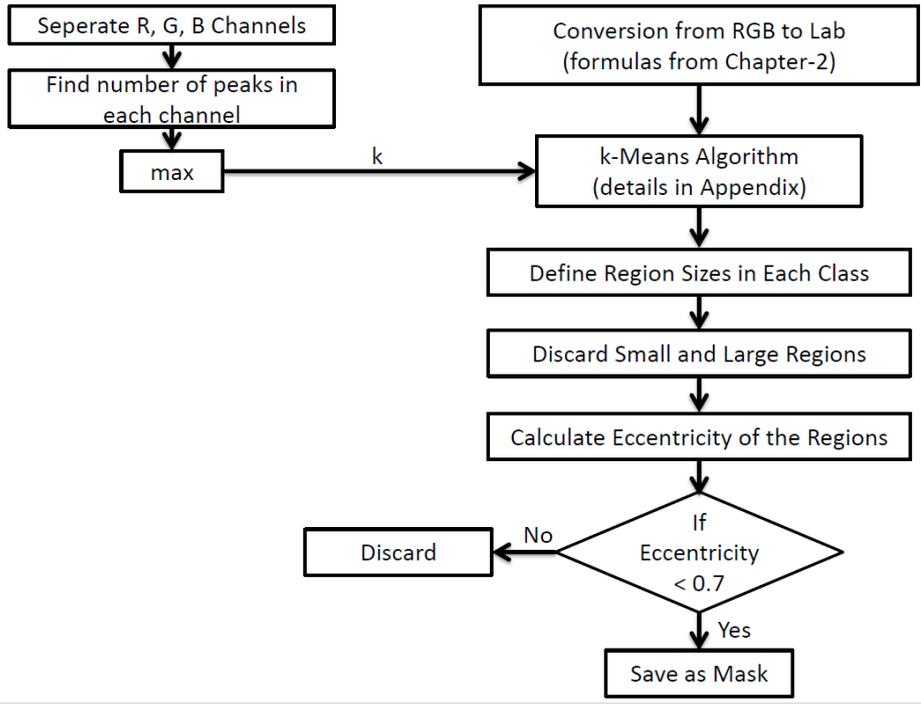
Figure 5.3: The Flowchart for Groundtruth Generation

proposed test can be recovered. The regions are separated into classes by using k-means algorithm explained in the Appendix part. The original first frame from the camera and the corresponding k-means algorithm result can be seen in Figure 5.4. After this step, regions smaller and greater than certain limits are discarded. Then, each region is examined according to its eccentricity as in the third proposed test. Unlike the previous test, this time the regions which have eccentricity less than 0.7 are chosen as the candidates for the mask. At this point, normally, we do not expect any other region to exist in the resulting image. On the other hand, in some cases, some other areas can be obtained like in Figure 5.5. Thus, the region which is the closest to the center of the image is determined as the mask of the cylindrical case, as shown in Figure 5.5.

As the next step, the obtained mask is compared with the motion masks of each frame. It is important to note that only the frames after the switching of the lamp have motion masks because the scene does not contain any change other than the illumination change due to the lamp. Also, the time difference between the switching of the lamp (10 seconds) is long enough to give the algorithm to absorb the change
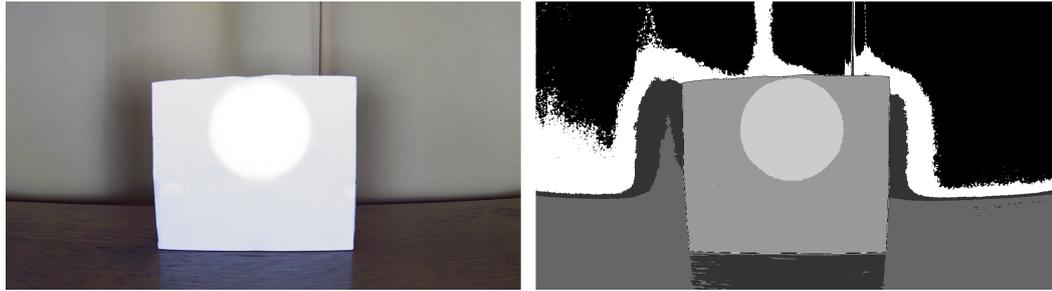
Figure 5.4: The First Frame of the Video (left) and The Result of k-means Algorithm (right)
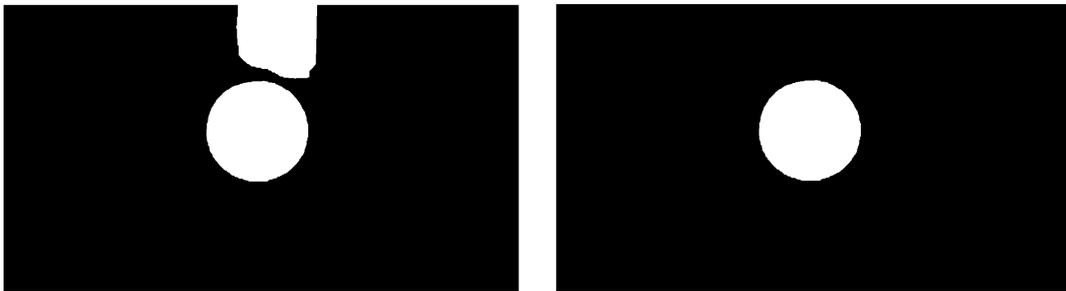
Figure 5.5: The First Version of the Mask (left) and The Last Version of the Mask (right)

as background. So, the illumination change creates a motion mask when the lamp is turned on or off; and then it fades away in time till the next turning on or off. Thus, all the processed frames of the video are checked with the mask of the cylindrical case and the frame with the maximum coverage is chosen for processing.

Pierre-Marc Jodoin, Janusz Konrad, Prakash Ishwar and Fatih Porikli organized a challenge in 2012 in order to measure the performance of the motion detection algorithms [94]. They provided a different set of images as dataset taken by different cameras in different conditions. They actually separated the examples into 6 categories according to the problem types of most of the motion detection algorithms have. These categories are baseline, dynamic background, camera jitter, intermittent object motion, shadows and thermal. In the challenge, the idea was that each algorithm evaluates the dataset and the performance of the algorithm is determined by comparing the results with the ground truths provided by the organizers. For the comparison of the performances, some terms were defined and by using them a metric

was proposed.

Although the applications are different, the ideas of this challenge and our performance test are quite similar. In the challenge, the image set is constant for all the motion detection algorithms and the performance of these algorithms are compared by checking the difference between the ground truths. On the other hand, in our test, the motion detection algorithm and the test setup is constant while the camera is the variable. In the same manner, the ground truth is prepared by using some image processing actions instead of doing it manually for each camera. So, for each case, the ground truth and the motion mask are obtained as output. Thus, using the same metric for both cases is reasonable. Hence, the following terms should be defined as proposed in the challenge.

• True Positive (TP): The total number of pixels which are marked as motion in both the ground truth and the motion mask (output of the motion detection algorithm)

• False Positive (FP): The total number of pixels which are marked as motion in motion mask but not marked as motion in the ground truth

• False Negative (FN): The total number of pixels which are marked as motion in the ground truth but not marked as motion in motion mask

• True Negative (TN): The total number of pixels which are not marked as motion in both the ground truth and the motion mask

• Recall: TP/(TP+FN)

• Precision: TP/(TP+FP)

• F-Measure: (2*Precision*Recall)/(Precision+ Recall)

Each pixel in the chosen frame and the mask of the cylindrical case is compared and the above terms are calculated. F-Measure is the key metric that we use in our performance measurements because it provides a nice numerical representation which has the effect of both true and false results without being affected by the size of the cylindrical case. The F-Measure Metric converges 1 when two conditions are satisfied. At the first one, the created change mask by the change detection algorithm fits the groundtruth of the system which is the opening of the cylindrical case of the test device. The second conditions is that the change detection algorithm does not detect any false alarms which means there is no change in the area other than the groundtruth.

5.5 Test Results

The proposed test method used in order to measure the performance of eight test cameras. The videos taken by these cameras in the proposed test setup with using the test device are examined according to the details explained in the previous part. The results will be given in this part.

5.5.1 The results of Camera-1

The first frame of the video taken by the Camera-1 is given on the left side of Figure 5.6. One frame from the interval of the lamp is turned off is also given on the right side of Figure 5.6. The ground truth image is shown on the left side of Figure 5.7. The image which has the maximum coverage of the ground truth according to the motion detection algorithm is given in Figure 5.7. In Figure 5.8, the change on the F-Measure metric can be observed through the frames of a sample video taken by Camera-1. As expected, F-Measure metric stays closer to zero values while there is no change in the scene. It takes the maximum values when the change occurs and then, it converges to zero because the new scene becomes the part of the background.

Figure 5.6: The First Frame of Camera-1 (left) and a Frame after Illumination Change (right)

Figure 5.7: The Ground Truth for Camera-1 (left) and Detected Change (right)

Figure 5.8: The F-Measure Change

5.5.2 The results of Camera-2

The first frame of the video recorded by the Camera-2 and a frame after the illumination change are shown in Figure 5.9. The corresponding ground truth image is given on the left side of Figure 5.10. The image of the best accuracy of motion detection algorithm is given also in Figure 5.10. The change on F-Measure values is given in Figure 5.11.

106

Figure 5.9: The First Frame of Camera-2 (left) and a Frame after Illumination
Change (right)

Figure 5.10: The Ground Truth for Camera-2 (left) and Detected Change (right)

Figure 5.11: The F-Measure Change

5.5.3   The results of Camera-3

The first frame of the video taken by the Camera-3 is given on the left side of Figure 5.12. One frame after illumination change is also given on the right side of Figure 5.12.The mask of the cylindrical case is shown on the left side of Figure 5.13. The best match of the output of motion detection algorithm is given in Figure 5.13. The F-Measure values change through the frames of the video as shown in Figure 5.14.

Figure 5.12: The First Frame of Camera-3 (left) and a Frame after Illumination Change (right)
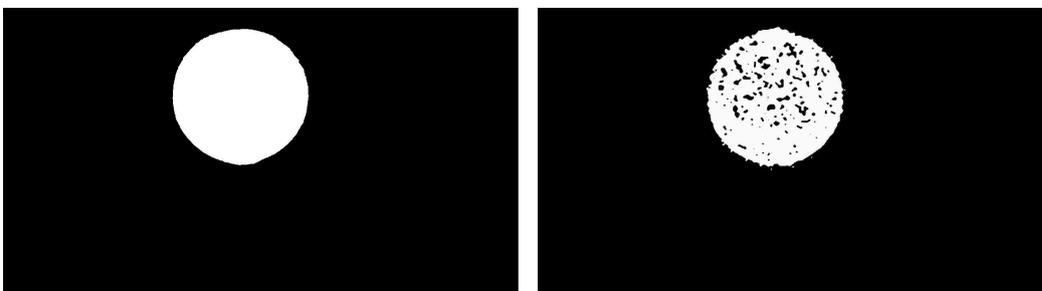
Figure 5.13: The Ground Truth for Camera-3 (left) and Detected Change (right)

Figure 5.14: The F-Measure Change

Such impulsive characteristics obtained in the results of the Camera-1, the Camera-2, and the Camera-3 indicate that the response time of the camera is short. In other words, the frame which is acquired at the moment of the change is very close to the steady-state output of the camera. Thus, the change in the scene becomes a part of the background in a short time period. If the F-Measure metric does not converge to zero in a short time after the change in the scene, it shows that the pixel values changes in the region of interest for a long time until they reach the steady state values. This can be the result of the auto-exposure procedure of the camera or the effect of the flicker correction algorithm performed by the processor inside the camera.

5.5.4 The results of Camera-4

The first frame of the video recorded by the Camera-4 and a frame after the illumination change are shown in Figure 5.15. The corresponding ground truth image is given on the left side of Figure 5.16. The image of the best accuracy of motion detection algorithm is given also in Figure 5.16. In Figure 5.17, all F-Measure values can be observed.
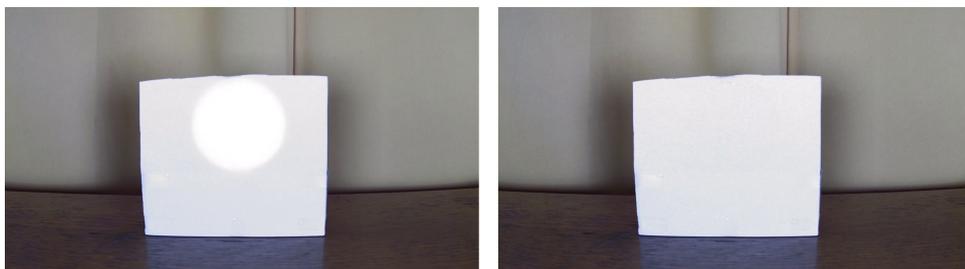
Figure 5.15: The First Frame of Camera-4 (left) and a Frame after Illumination
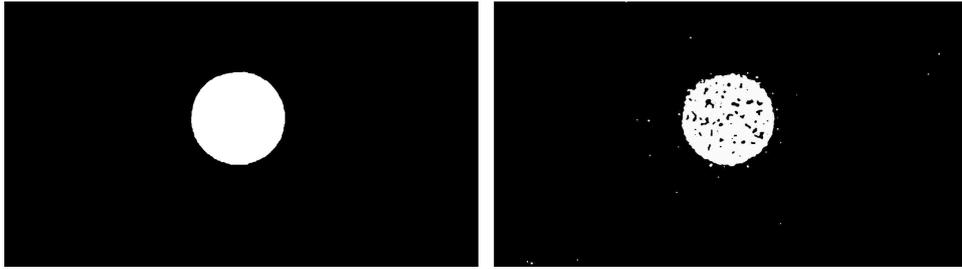Change (right)
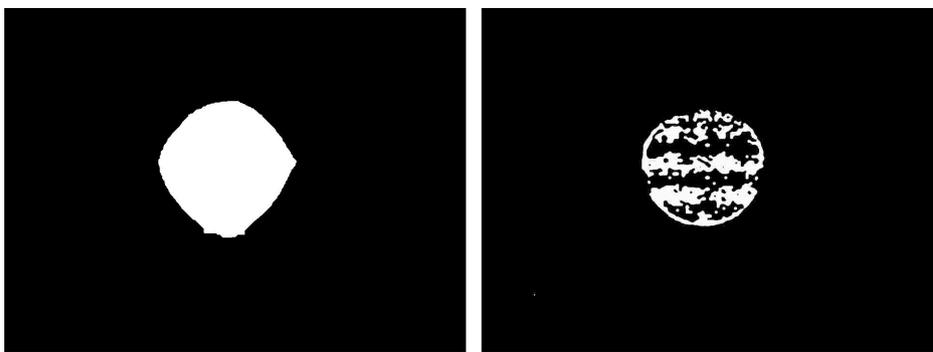
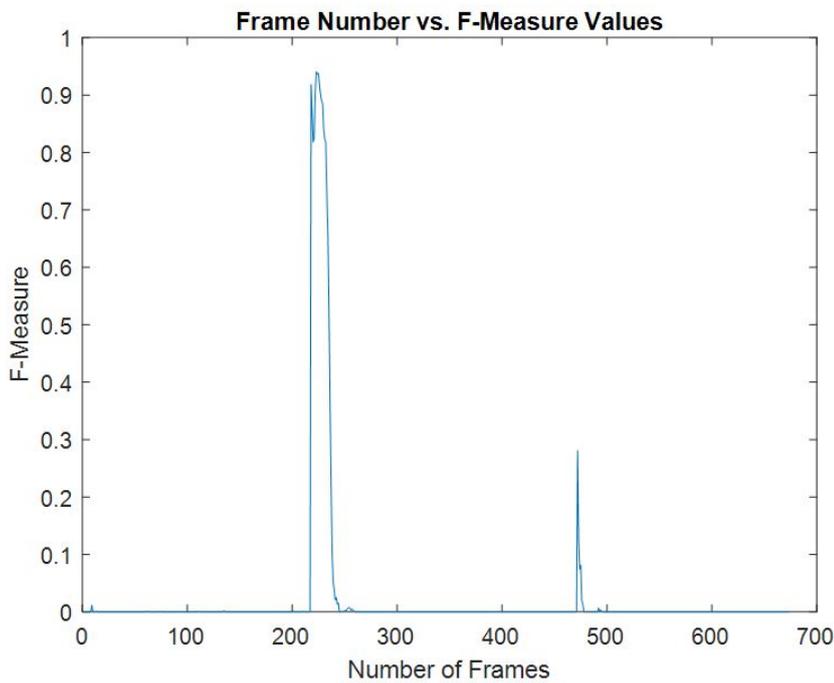Figure 5.16: The Ground Truth for Camera-4 (left) and Detected Change (right)
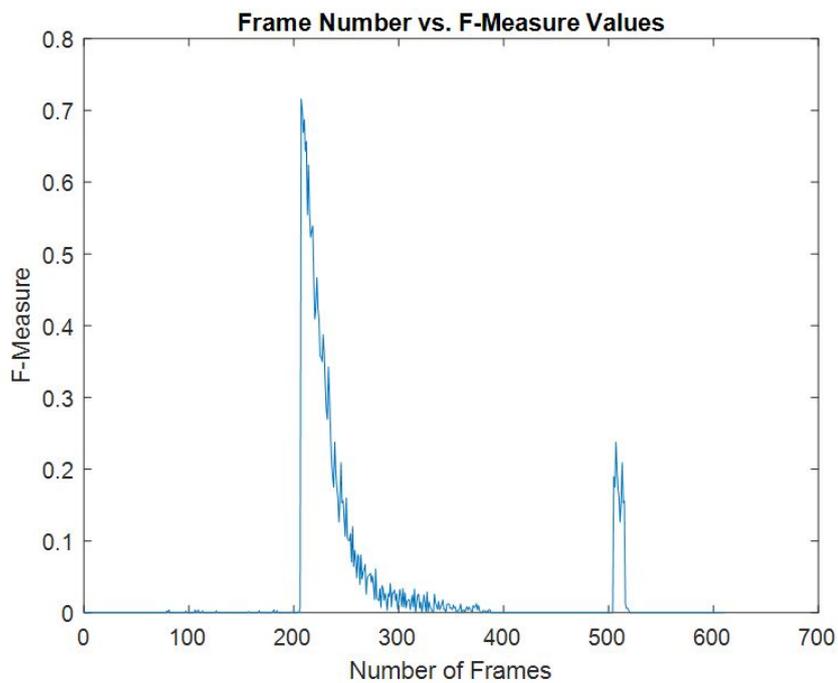
Figure 5.17: The F-Measure Change

5.5.5 The results of Camera-5

The first frame of the video taken by the Camera-5 is shown on the left side of Figure 5.18. A frame from the interval of the lamp is turned off is also given on the right side of the Figure 5.18.The cylindrical case mask image is given on the left side of Figure 5.19. The frame which has the maximum coverage of the ground truth according to the motion detection algorithm is given in Figure 5.19. The change in F-Measure values is given in Figure 5.20.
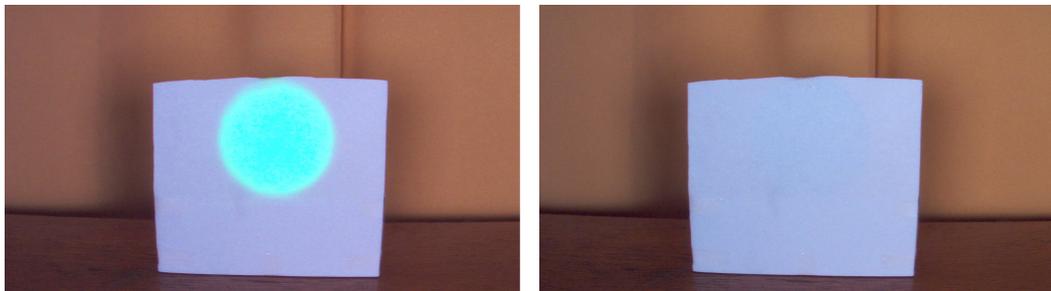
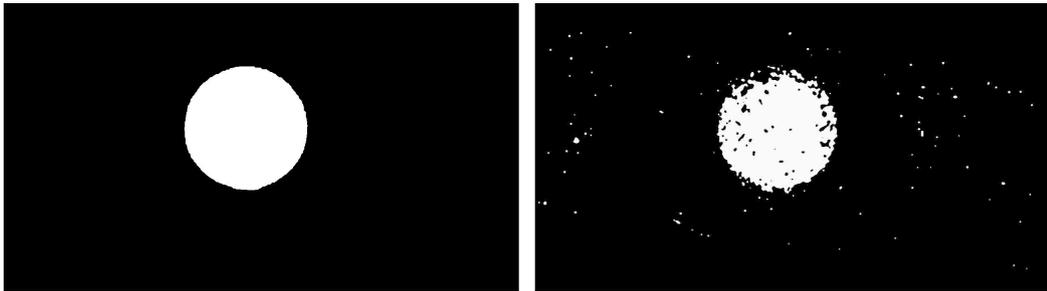Figure 5.18: The First Frame of Camera-5 (left) and a Frame after Illumination Change (right)
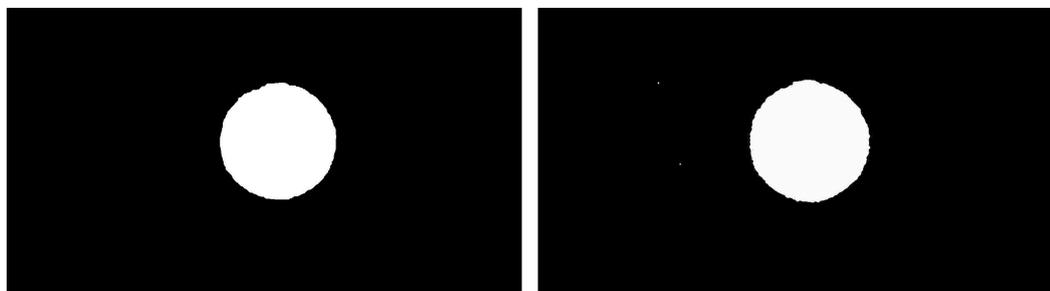
Figure 5.19: The Ground Truth for Camera-5 (left) and Detected Change (right)

111
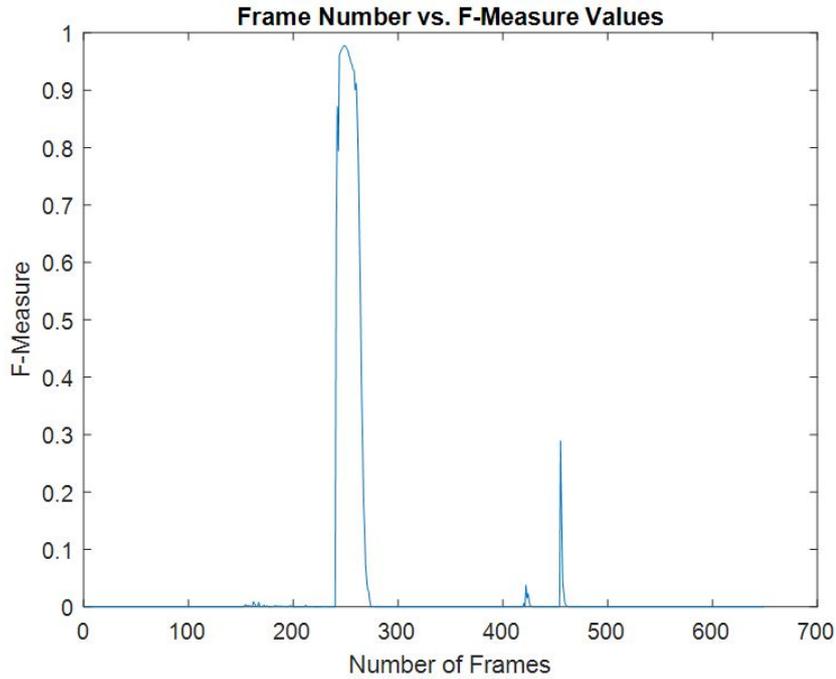
Figure 5.20: The F-Measure Change

5.5.6 The results of Camera-6

The first frame of the video recorded by the Camera-6 and a frame after illumination change are given in Figure 5.21.The mask of the cylindrical case is shown on the left side of Figure 5.22. The best match of the output of motion detection algorithm is also given on the right side of Figure 5.22. The F-Measure values can be observed in Figure 5.23.
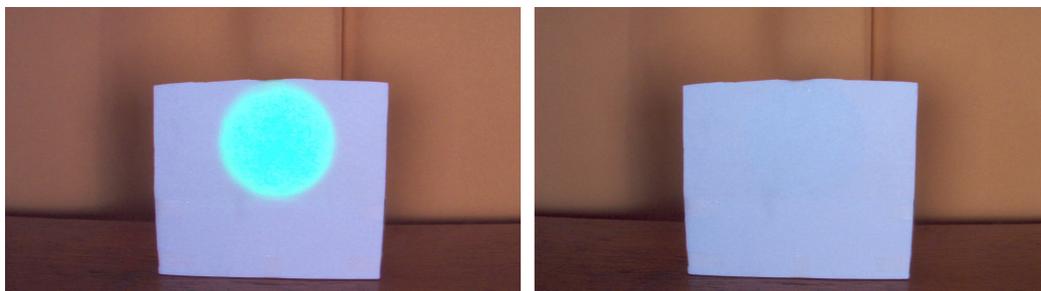
Figure 5.21: The First Frame of Camera-6 (left) and a Frame after Illumination Change (right)
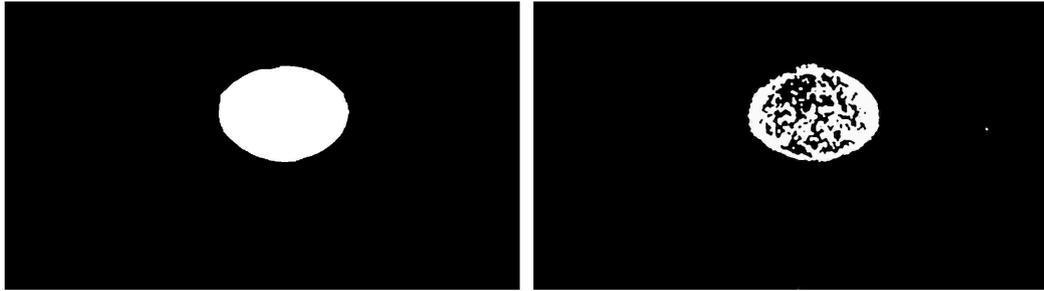
Figure 5.22: The Ground Truth for Camera-6 (left) and Detected Change (right)
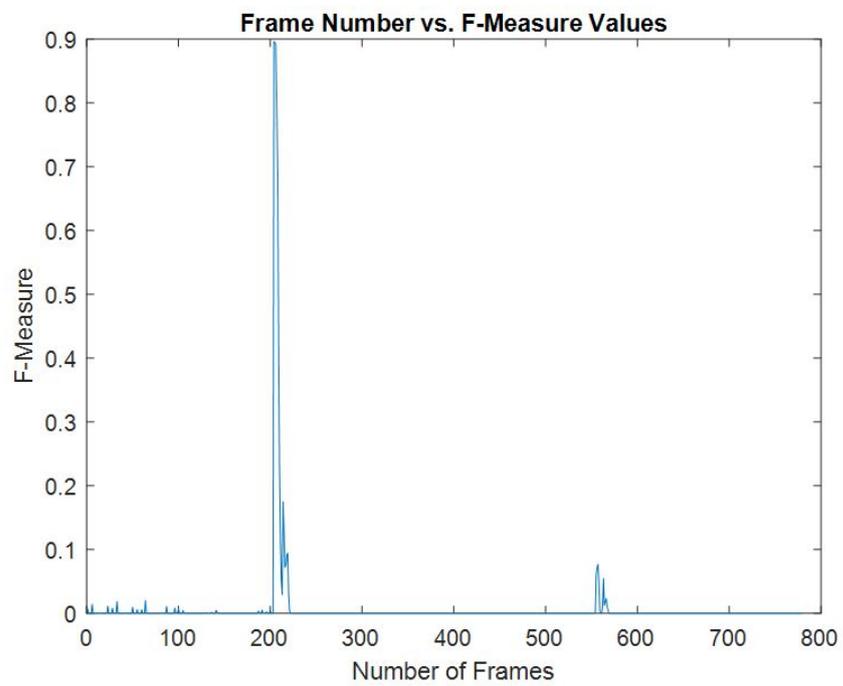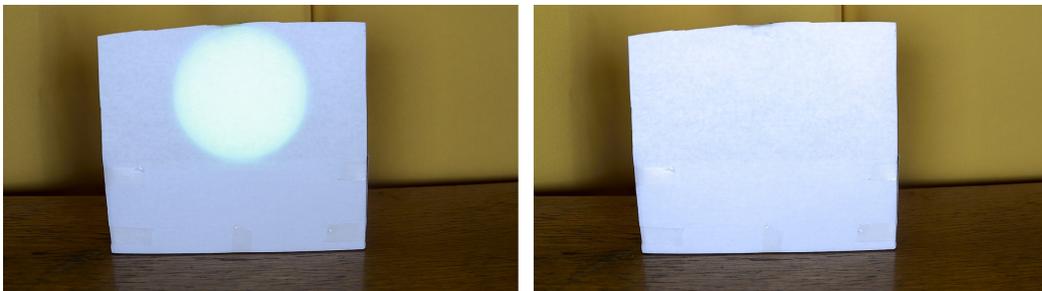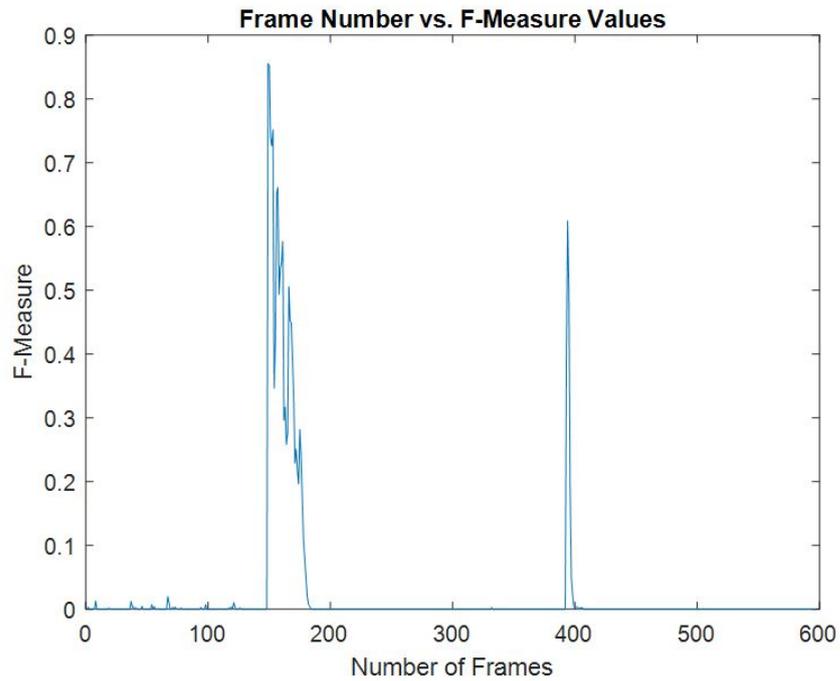

Figure 5.23: The F-Measure Change

5.5.7   The results of Camera-7

The first frame of the video taken by the Camera-7 is shown on the left side of Figure 5.24. A frame from the interval of the lamp is turned off is also given on the right side of Figure 5.24. The cylindrical case mask image is given on the left side of Figure 5.25. The image which has the maximum coverage of the ground truth according to the motion detection algorithm is given in Figure 5.25. For a sample video acquired by Camera-7, the F-Measure values are given in Figure 5.26.
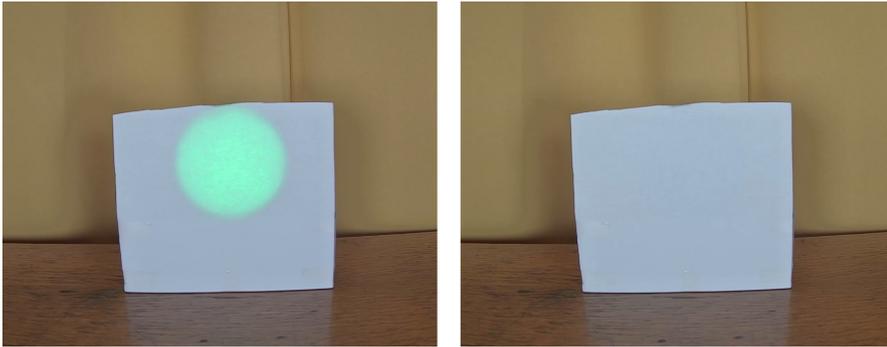
Figure 5.24: The First Frame of Camera-7 (left) and a Frame after Illumination Change (right)

Figure 5.25: The Ground Truth for Camera-7 (left) and Detected Change (right)
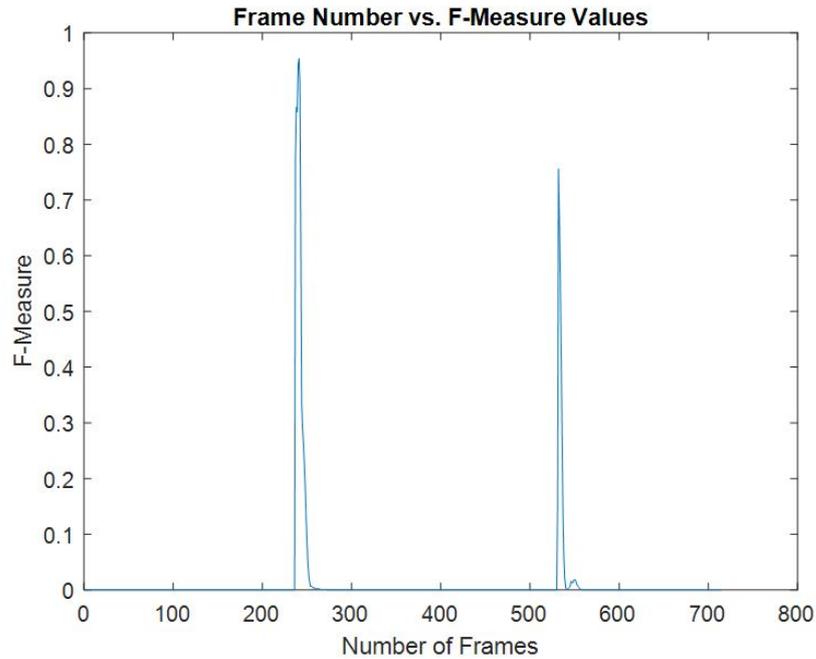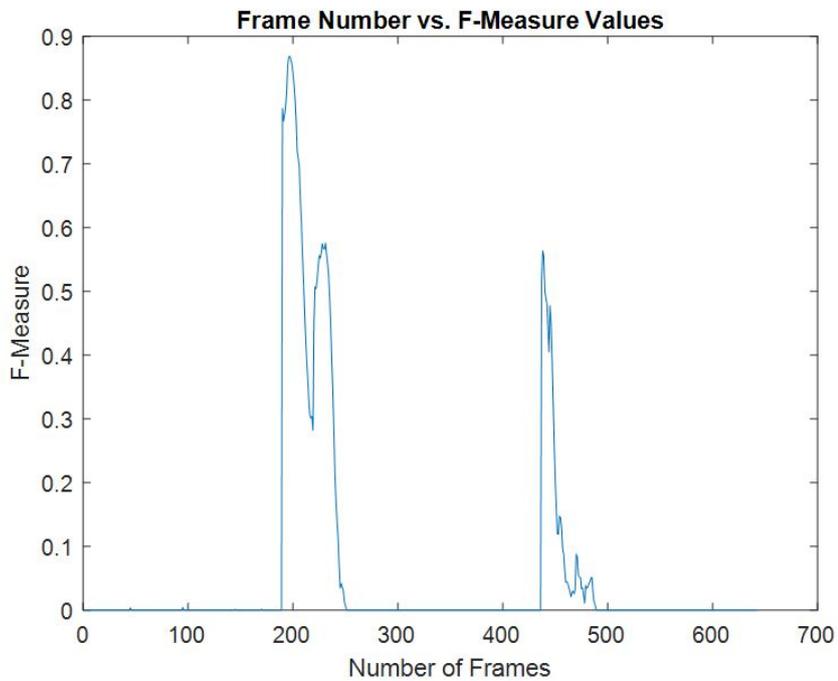

Figure 5.26: The F-Measure Change

114

5.5.8 The results of Camera-8

The first frame of the video recorded by the Camera-8 and a frame after the illumination change are shown in Figure 5.27. The corresponding ground truth image is given on the left side of Figure 5.28. The image of the best accuracy of motion detection algorithm is given also in Figure 5.28. The F-Measure change between the frames of a sample video is shown in Figure 5.29. Due to the effect of the flicker algorithm of the camera, the results of the Camera-8 gives the worst change sensitivity performance.
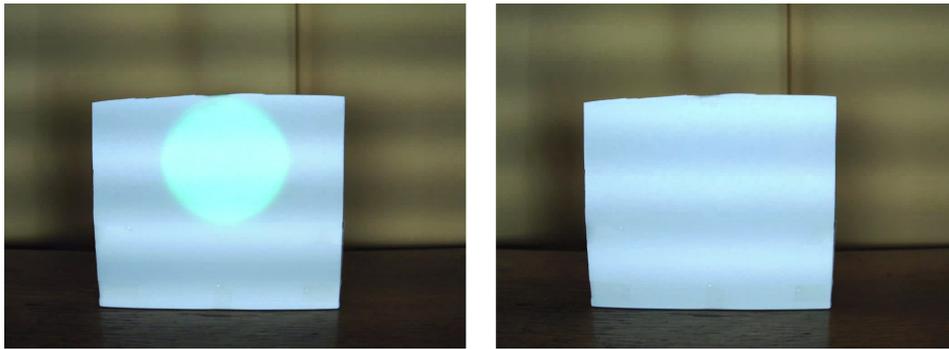

Figure 5.27: The First Frame of Camera-8 (left) and a Frame after Illumination Change (right)
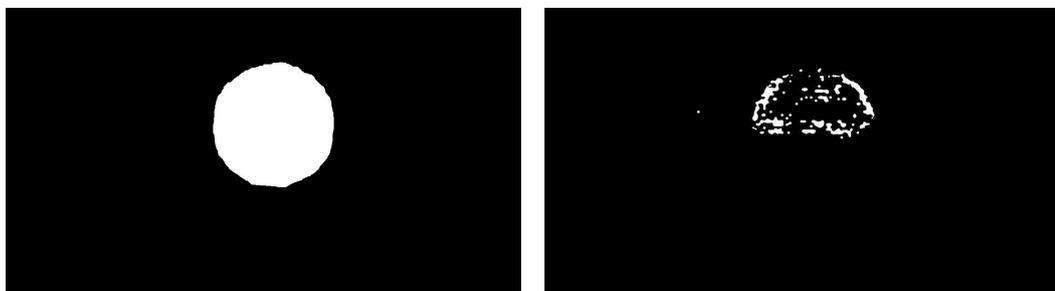
Figure 5.28: The Ground Truth for Camera-8 (left) and Detected Change (right)
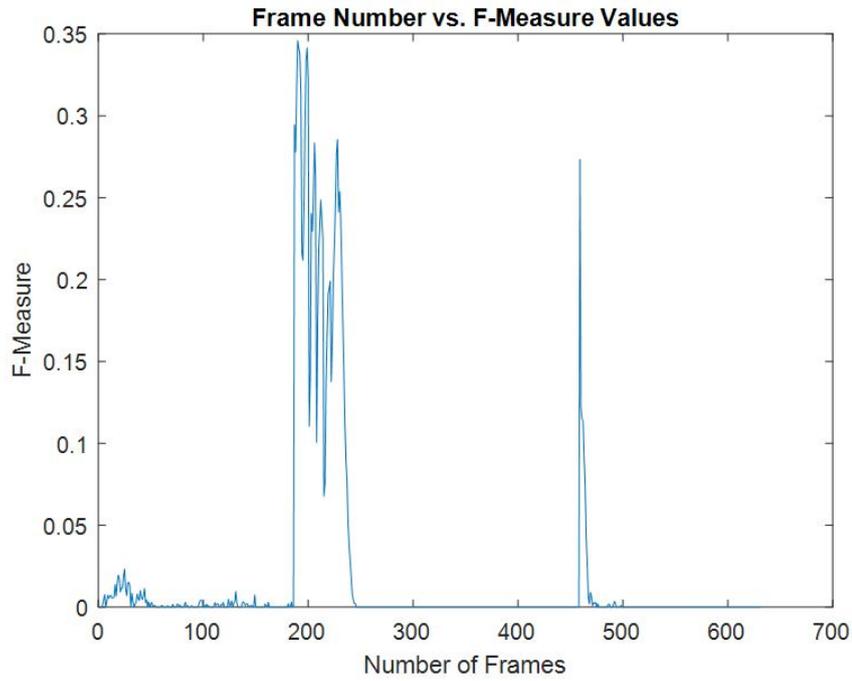
Figure 5.29: The F-Measure Change

Table 5.1 gathers the results of the change sensitivity tests for eight test cameras.

Camera No	F-Measure Value
1	0.941
2	0.895
3	0.954
4	0.716
5	0.961
6	0.856
7	0.855
8	0.346

Table 5.1: The Results of the Change Sensitivity Tests

CHAPTER 6

CAMERA TESTS ON MOTION BLUR AND RESULTS

6.1   Background Information about Motion Blur

The importance of the motion detection algorithms was discussed in the previous chapter. As mentioned before, the motion detection algorithms are used in order to find the region of interest in the frames which is the input of the more complex image processing applications. However, it has to be considered that if there is a motion in the scene, there always exists the corresponding motion blur. Therefore, even though the change detection algorithm works fine for the illumination change, its performance is affected by the motion. Actually, it was the reason to use only illumination change in the previous test in order to test the change detection performance of the camera without the effect of the motion blur.

Firstly, it is important to understand why the frames acquired by the camera have motion blur if there is a motion in the scene. Although the camera is considered as the replica of the human eye, due to the difference between their working principles, their performances are not the same, unfortunately. While the human eye provides analog (continuous) data to the human brain, the camera can form an image by collecting data during a time interval (digital) [36]. Thus, the image obtained by using the camera is actually the average of the scene in this time interval which is called the exposure time. In other words, the light coming from the scene passes through the lens and excites the photosite in the camera sensor. The photosites are being excited by the light during a certain time which is theexposure time. Therefore, the resulting image represents the average scene during the exposure time. So, a moving object affects all the pixel values which cover the trajectory of the moving object during the exposure

time. For example, during the exposure time, the trajectories of the hands of the guy in Figure 6.1 cover the region surrounded by the red line. Thus, only these pixels experience the motion blur.

Figure 6.1: An Example of Motion Blur [36]

It is obvious from the previous example that if the exposure time is kept the same, increasing the speed of the moving object results with increasing the number of pixels which are affected by the motion blur because the change in the scene covers more pixels due to the larger displacement. This effect can be observed from the images taken from the studies of Gultekin [37]. For all the four case given in Figure 6.2, the exposure time is kept as constant at 100 milliseconds and the chart is moved at different speeds. The images from (a) to (d) represent the corresponding speeds of 2 cm/s, 5 cm/s, 7 cm/s and 11 cm/s, respectively.

As mentioned above, the most effective parameter that can be adjusted by the camera is the exposure time in order to decrease or increase the motion blur in the images or the frames of the video. If the exposure time is adjusted shorter, then motion blur in the image decreases. Similarly, if the exposure time is extended, then the image would have more motion blur effect and it is desired sometimes in photography [38]. The cars which have almost the same speed are photographed by the different exposure times. In the left side of Figure 6.3, the image acquired with 33 milliseconds exposure time while the one on the right side was taken with 2 milliseconds. In the first one, it can be seen that the car is moving. However, in the second one, the cars look like
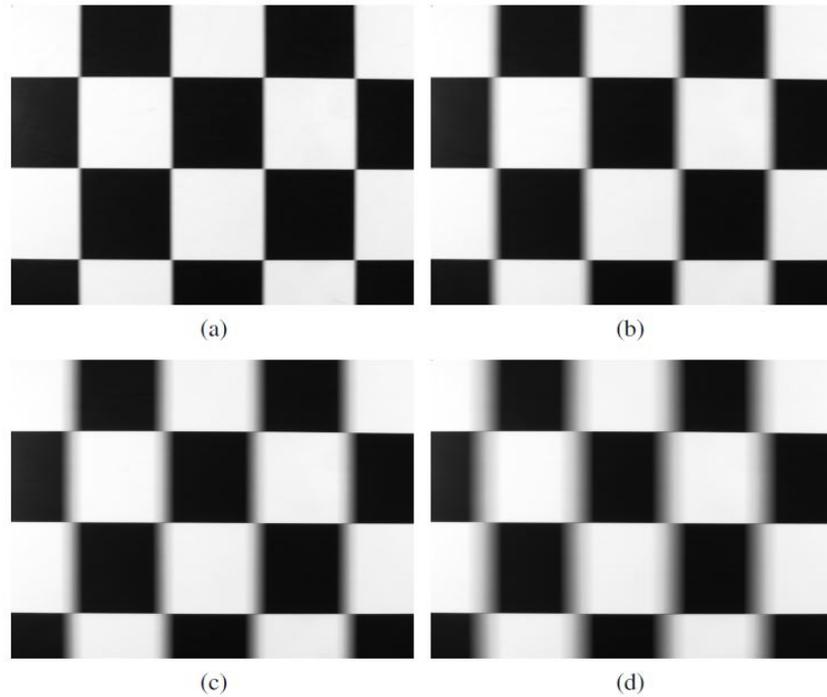
Figure 6.2: Motion Blur Change for Different Test Chart Speeds [37]

they are stationary due to the low motion blur.

Figure 6.3: The Effect of the Exposure Time on Motion Blur [38]

On the other hand, decreasing the exposure time is not always the solution for decreasing the motion blur. First of all, the exposure time of the camera has its physical limits. In other words, it is not possible to decrease it to the certain level. Secondly, as explained before, the exposure time affects the number of photons hitting to the camera sensor and thus, both the voltage of the photosite and the value in the corresponding pixel are proportional to the exposure time. Thus, in some cases, it is

inevitable to increase the exposure time. For example, extending exposure time is the most preferable way to obtain a detailed image in the dark illumination conditions. So, actually, adjusting the exposure time provides a trade-off to the user with the details in the scene and the motion blur.

If the user is not a professional photographer like in the most of the case, in photo or video mode, users prefer to choose the auto adjustment for the exposure time. In auto mode, the camera tries to predict and choose the most suitable value for the exposure time in the current environment. For the cameras which cannot be adjusted manually by a user, like security cameras, this selection for the exposure time is important. Thus, both in this and the previous test, the cameras under test are kept in the auto mode for the exposure time selection.

Until this point, the motion blur caused by the existence of a moving object in the scene has been discussed. However, it is not the only reason for the motion blur. Even if the scene is stationary, the movement of the camera can create the motion blur in the image or the video. Actually, this is one of the biggest problems in the camera industry. Most of the cameras on the market such as handy cameras, action cameras and the cameras on the mobile phones experience a motion while acquired image or video. This motion is called a handshake and it creates motion blur as shown in Figure 6.4. Nowadays, the handshake problem is partially solved by using two different methods. In the first way, the methods based on convolution and/or Fourier Transform are used in order to motion blur reduction [103]. In the other way, the different procedures are applied for image stabilization by using the data taken from the gyroscopes and/or accelerometers on the devices. Generally, it is the case for the smartphones; and they give good performance for reducing motion blur due to the camera shake according to the studies in the literature [104], [105], [106], [107].

It is important to note that using the image stabilization methods instead of algorithms of motion blur reduction is much more effective because the motion blur reduction algorithms increase the noise in the resulting image. This effect can be observed in Figure 6.5.

The only standardized test procedure for the measurement of image stabilization was proposed by Standard of Camera & Imaging Products Association (CIPA) in 2012

120

Figure 6.4: The Motion Blur of Hand Shake [39]

and revised in 2014 [40]. According to the standard, a testing setup was proposed as shown in Figure 6.6. The illuminator type was not specified but in order to reply the tests for different conditions, it was preferred as a variable. The proposed test chart has black and white stripes and natural color images in the middle of the stripes as shown in Figure 6.7.

The most complicated and critical device in the test setup is the vibratory apparatus. This device applies the motion provided by the simulation computer to the camera under test in order to simulate the handshake of the human body. The motion directions can be seen in Figure 6.8.

The measurement of the motion blur is done in the simulation computer frame-by-frame by using the software provided by CIPA. Due to the evaluation software is provided by CIPA, the implementation details are not well-defined. However, according to the works of Uwe Artmann and Philipp Feldker from Image-Engineering GmbH, the evaluation software of CIPA uses the edges near the transitions between black and white stripes [41]. After several processing steps like morphological filtering, the width of the edge is calculated in order to determine the motion blur. An example image from the camera under test and its corresponding edge the width are shown in Figure 6.9.

Even though their work is not the standard for the image stabilization, Artmann and Feldker proposed using slanted edges and MTF calculation of ISO 12233 by using the same vibration apparatus and the motion profile. According to their results, the

121
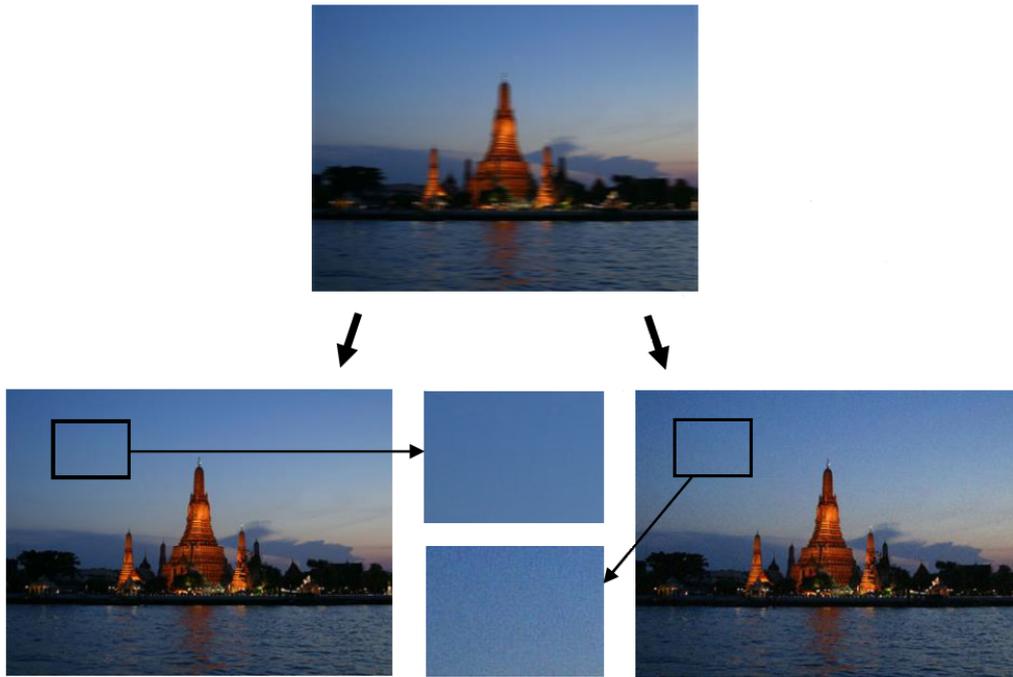
Figure 6.5: The Difference between Image Stabilization (left) and the Motion Blur Reduction (right) [40]

measurement takes less time and provides more stable results than CIPA [41]. The used test chart can be seen in Figure 6.10.

It is important to note that our main aim is to prepare a motion blur test which can be applied to every camera in order to compare their differences. Of course, the cameras are classified under some group and it is important to choose the correct one according to the application. As mentioned in the previous chapters, although our tests are applicable for every kind of camera, the security cameras are our main concern. The two different reasons for the motion blur are explained in detail above. Firstly, a moving object in the scene can create a motion blur in the image or the video. Secondly, the camera movement creates motion blur even the scene is stationary. So, if we think about the security cameras, specifically; the second case is not possible for most of the case because the security cameras are generally mounted on a stationary stuff like wall or pole. Thus, other than cases like earthquakes, strong wind or vandalism, the security cameras record stationary which does not create motion blur. Hence, testing the image stabilization performance of a security camera does not give an important
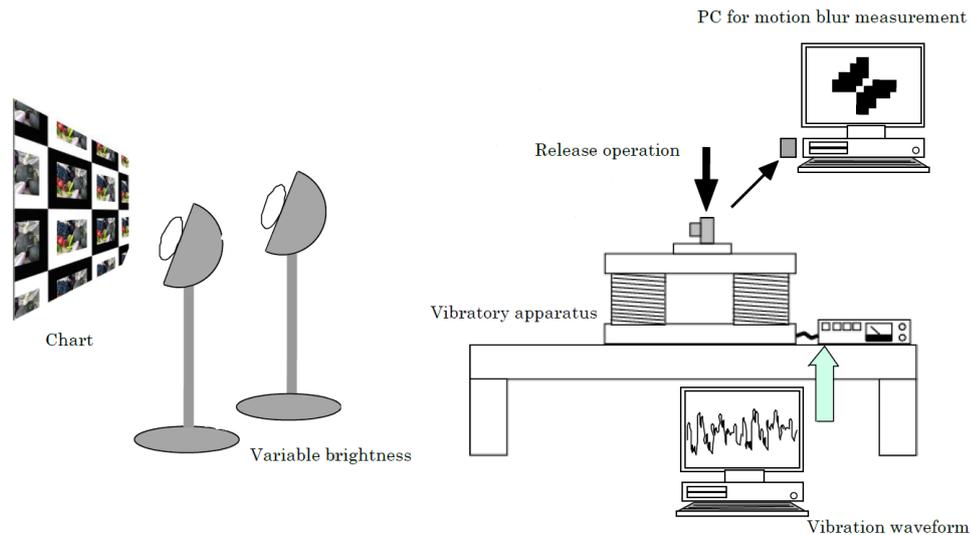
Figure 6.6: The Test Setup Proposed by CIPA for Image Stabilization Measurements [40]

criterion for the end user. This is the reason that we concentrate on the moving objects in the scene instead of the camera movement.

Using the moving test charts in order to measure the motion blur of the cameras is also studied by Image-Engineering GmbH [42].For this purpose, a test setup which has a moveable test chart is prepared. The illumination is provided by a controllable light source such that 1310 lux is provided on the test chart. The test setup can be seen in Figure 6.11. The movable test chart can move from left to right with adjustable speeds from 1 cm/s to 200 cm/s. The proposed test chart includes different testing tools like gray patches, markers or the corners, colored dead leaves structure in the middle and slanted edges around the test chart. Its details are given in Figure 6.12.

In the studies of Artmann, the colored dead leaves structure and slanted edges are examined for the different speeds of the test chart. The movement of the test chart is recorded by the camera under test and the frames of the videos are examined. The usage of the slanted edge for the sharpness measurement was explained in detail before in the third proposed test. Thus, using the colored dead leaves target will be shortly explained here.

The colored dead leaves target has known power spectrum thanks to the design and the
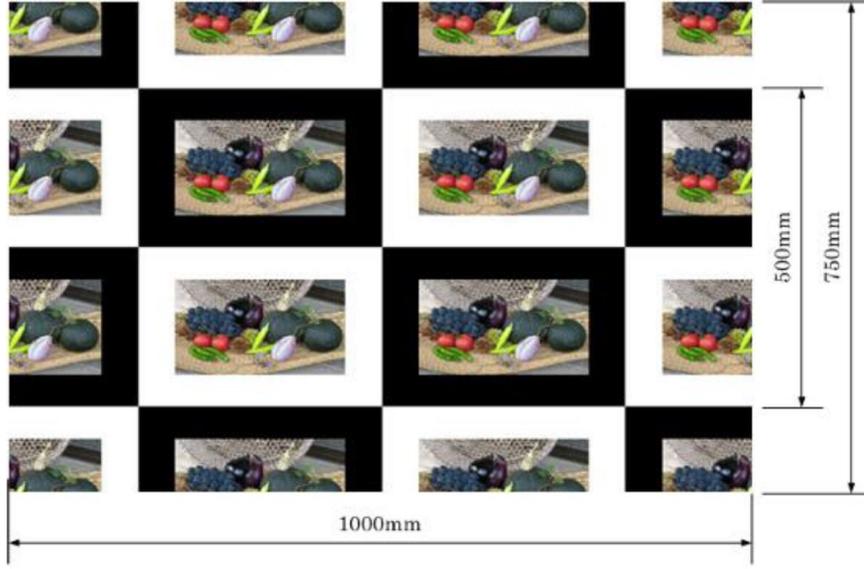
Figure 6.7: The Test Chart Proposed by CIPA for Image Stabilization [40]
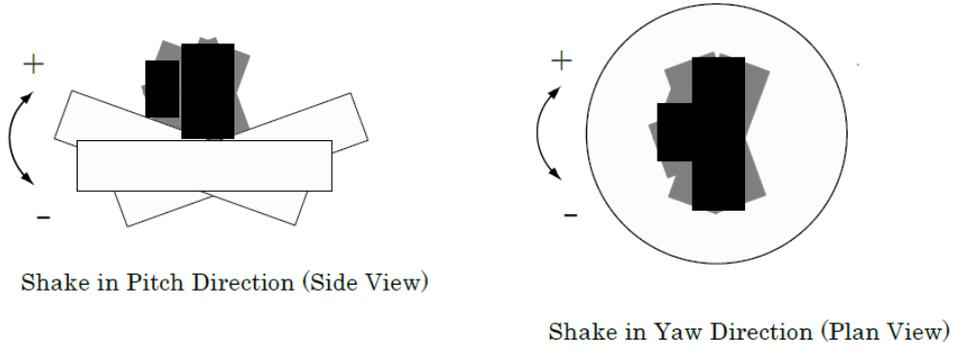
Figure 6.8: The Moving Directions of Vibratory [40]

production technique of the test target. This power spectrum is called as $PS_{target}(f)$. While the test chart and the camera are stationary, the test chart is photographed and the power spectrum of the region of interest, the colored dead leaves target, in this case, is calculated. It is called as $PS_{reference}(f)$. Then, the power spectrum of the noise, $PS_{noise}(f)$, can be calculated as the difference between $PS_{target}(f) - PS_{reference}(f)$. The spatial frequency response (SFR) is then calculated by using the following formula for each frame in the video while the test chart is moving.

$$SFR = \sqrt{\frac{PS_{image}(f) - PS_{noise}(f)}{PS_{target}(f)}} \qquad (6.1)$$
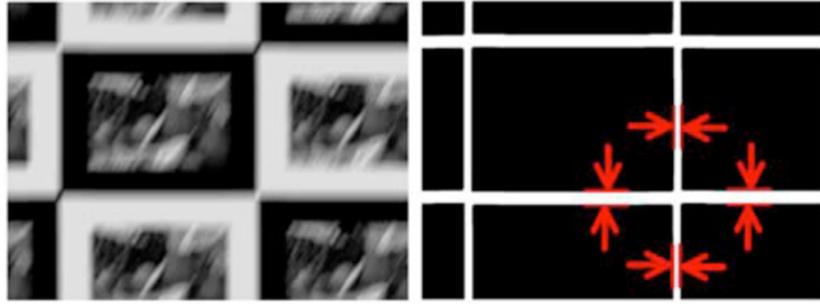
Figure 6.9: An Example Image Acquired by the Camera under Test (left) and Its
Edge Width (right) [41]
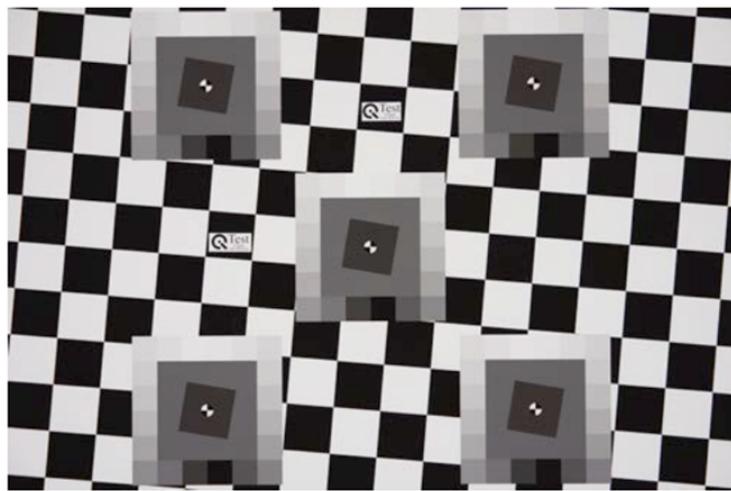
Figure 6.10: The Proposed Test Chart by Image-Engineering GmbH for Image
Stabilization Measurements [41]

Although using the colored dead leaves chart in the proposed way has a theoretical
basis, the results of the study show that the method can be fooled by the noise and the
artifacts due to the existence of the colors. So, the results do not only represent the
acutance change due to the motion blur but also the color formation problems because
of the motion in the scene.

6.2 Test Chart

As mentioned in the previous part, the motion blur occurs for two different cases;
the motion of the camera and the moving objects in the scene. Due to the fact that

Figure 6.11: The Test Setup Proposed by Image-Engineering GmbH [42]

our main concern is forming test procedures for security cameras, the second case is more critical in the security applications. With this idea, the only test chart in the literature is proposed by Image-Engineering GmbH which contains the slanted edge and the structure of colored dead leaves. However, as noted above, the color artifacts and noise dominate the motion blur with the increasing speed of the chart. On the other hand, our results which will be given in the next part prove that an edge in a moving test chart can experience different motion blur according to its orientation. The proposed method in ISO 12233 which provides MTF calculation by using slanted edge for digital still images states that the procedure is applicable for each slanted edge in the test chart. However, as noted before, the characteristics of the edges differ for moving targets. Thus, unlike the still images, using slanted edge does not provide the same graph for the different edges and it is hard to give meaning to these graphs. Because of these reasons, the colored dead leaves and slanted edge structures are not used in our test.

The only standardized test for motion blur is proposed by CIPA [40]. However, its main aim is to measure the image stabilization performance of the camera instead of motion blur directly. However, the idea behind the measuring software is also applicable for the tests with moving objects. Using the black and white stripes in the
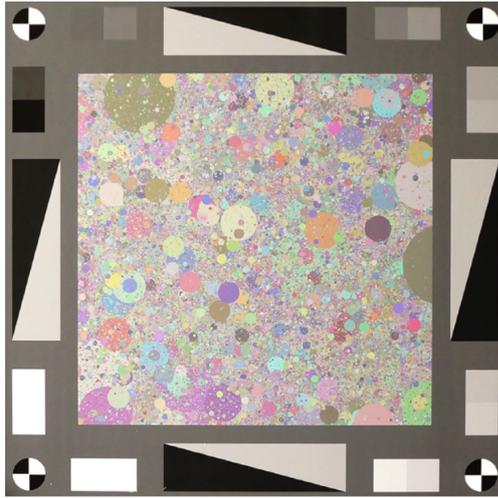
Figure 6.12: The Test Chart Proposed by Image-Engineering GmbH [42]

test chart, and observing the transition between these regions when there is motion gives a good idea about the motion blur performance of the camera. Thus, with the same idea, a test chart shown in Figure 6.13 with black and white stripes is used in our motion blur measurement tests. Due to the problems with color encountered by Image-Engineering GmbH, we decided not to add the "natural color image" inside the stripes unlike the test chart of CIPA. For future use, the circles are located to the corners of the black-white stripe patches.

Actually, it is called as checkerboard chart and widely used for different purposes in the camera industry. The effect of lens distortion and chromatic aberration can be measured by using a checkerboard chart [108]. Also, it is used on the movable mechanism in order to determine the intrinsic and extrinsic camera constraints [109].

6.3 Test Setup

Unlike the previous tests which are conducted in the dark room of Multimedia Research Laboratory, this test is performed in Laboratory of Robotics and Autonomous Systems of Middle East Technical University. It is obvious that a fully controlled moving mechanism is needed in order to provide all the conditions equal for all the cameras under test. This requirement is satisfied by using the 1D motion platform developed by Quanser which is composed of a cart and a rail which the cart can move
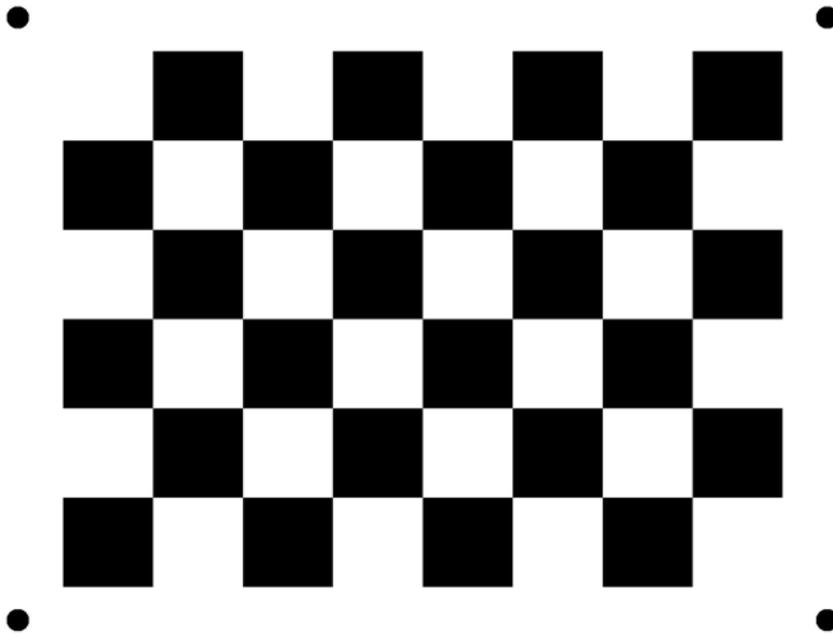
Figure 6.13: The Proposed Test Chart for Motion Blur Measurement

through it [110]. The movement of the cart is performed by using a DC motor and a high-resolution optical quadrature encoder is used in order to obtain the location information. A control loop implemented in Simulink determines PWM (pulse-width modulation) signal of the DC motor in order to keep the speed of the cart in the desired level by using the encoder data which is obtained by a data acquisition card. The described system above is already implemented in Laboratory of Robotics and Autonomous Systems of METU in order to use it for the studies which require the same conditions. It is vital to note that unfortunately, turning this common room into a dark room like the other one is not possible. However, thanks to its location inside the building, this room is quite dark such that there is almost no light when the small windows are covered with cardboard. So, fully controlled experimental conditions are provided for all cameras under test. Due to the same reasons which we cannot use our light sources in this laboratory, the illumination is provided by using a LED projector. The projector is placed behind the camera such that the shadow of the camera does not disturb the scene. On the hand, the evenness of the illumination on the test chart cannot be perfectly satisfied by using one projector. This condition shows its effect on the results as gradient change between the first and last position of the camera; however, thanks to the proposed performance metric, it does not affect

the test results. It is crucial to note that the non-evenness of the illumination exists for all of the cameras, so the controlled experimental conditions are still valid. In order to keep the consistency between the camera tests, the distance to the test chart is again adjusted to 45 cm as shown in Figure 6.14. In the same figure, there exists a metronome behind the test chart. The reason for its existence will be given in the next parts. Two images of the test setup can be seen in Figure 6.15.
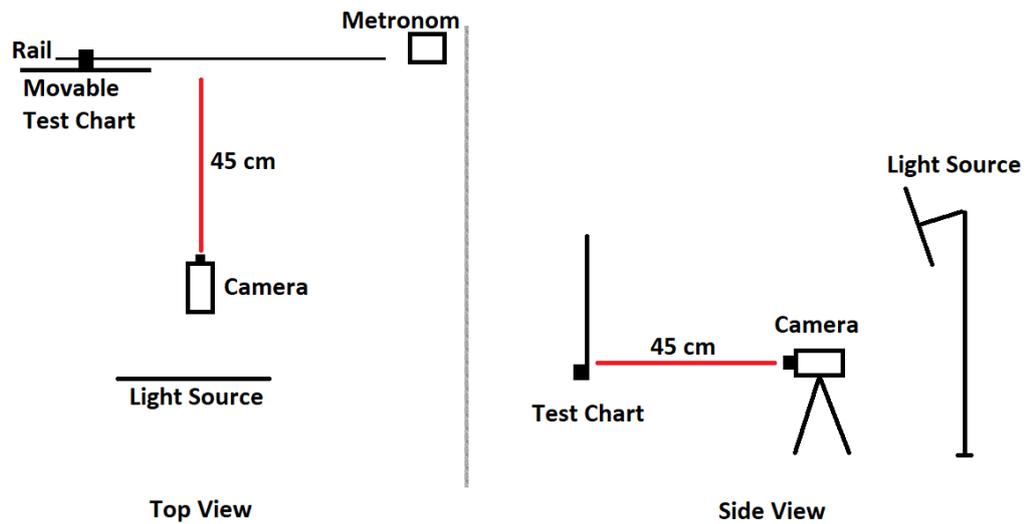
Figure 6.14: Test setup (not to scale)

Figure 6.15: Real test setup

6.4 Test Details

The proposed test setup is explained in the previous chapter but there is still a variable which is not determined, the speed of the test chart. At the first stage, it should be noted that the test chart speed is chosen as 15 cm/s for our test. The reason for this

will be explained by the details of the procedure. However, before going into the details of the testing procedure, it is vital to note that the frame rate of the cameras is adjusted to 25 or 30 fps like the previous test. As mentioned in the background information part, the exposure time is very important about the motion blur in the image. If the exposure time of the camera is long, then motion blur will affect the video more. On the other hand, if the scene is dark, increasing the exposure time is inevitable in order to catch the details. So, it can be called as a trade-off that should be chosen by the camera. Thus, we consider that it is the characteristic of the camera in the auto-exposure mode and should be part of the performance comparison. As a result, like the previous test, the exposure time is chosen as the auto-exposure mode. Also, for keeping the consistency throughout the work, H.264 is used as the codec like the previous test. Again, the video is recorded by resizing the input frames into a new base size which has 720 pixels in height without distorting the width/height ratio.

For the measuring of motion blur performance of the camera, the video recorded while the test chart is moving from left to right is examined frame-by-frame. The following procedure is repeated for each frame and the characteristic of the camera is obtained. As the last step, the difference between the ideal and the real values are calculated and this metric shows the performance of the camera.

As mentioned above, the main idea when the test chart is proposed for motion blur measurement is to observe the change in the gradient values near the edges like used in the standard of CIPA. Thus, the main problem to make the procedure automatically is to locate the black boxes in the test chart in general. For this purpose, in the first step, the location of the chart is determined by using Bradley's adaptive thresholding method. In order to discard the other details in the background, the regions which are smaller than the possible test chart size discarded. An example can be seen in Figure 6.16.

Then, Bradley's adaptive thresholding method is again used in order to find the location of the black boxes in the test chart. Finding the test chart and discarding the background are crucial because there can be other objects which have the similar size with the black boxes. Thus, it becomes harder to filter them and determining the
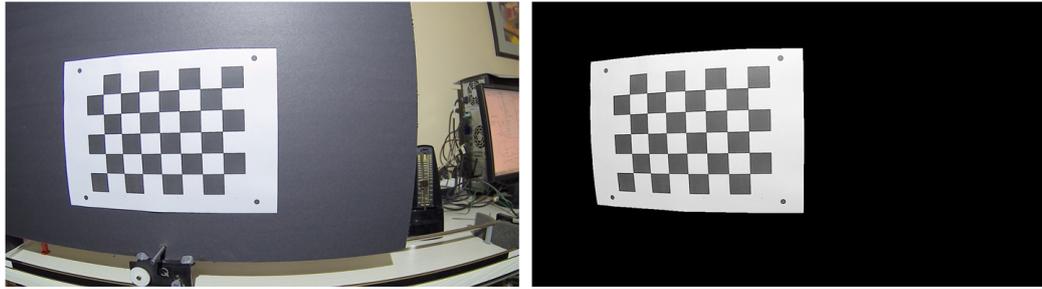
Figure 6.16: A frame (left) and Detected Test Chart (right)

correct locations. After thresholding, opening morphological operation is performed in order to guarantee that the boxes are differentiated from each other. The resulting binary mask can be seen from Figure 6.17.

There exist 24 black boxes in the test chart and each black box has 2 edges on its left and right side in the direction of the motion. Due to the fact that these are the only edges in the direction of the motion, the effect of the motion blur is only observable there. In other words, the edges on the top and the bottom of the boxes do not experience motion blur. Thus, only right and left edges of the black boxes are examined. However, our observations show that for some cameras, the edges on the right and the left can experience different amount of motion blur during the motion of the test chart. Thus, two different characteristics are obtained in order to represent the motion blur of the camera. While forming these two characteristics, the averages of 24 boxes are calculated for right and left edges separately.

The right and the left edges of a black box can be easily found by using the obtained binary mask shown on the left side of Figure 6.17. Each mask is cropped as shown on the right side of Figure 6.17. Then, the gradient on the horizontal direction is calculated and for each row, the maximum value is taken. The average of these maximum values is determined as the gradient of that edge. The average of the same edge for all the boxes is recorded as the gradient value of that frame. The record of these values is used as the characteristic of the camera. As shown in Figure 6.18, a camera has two characteristics for the left and the right edge.

The graphs can be read like as follows; the test chart is still in the first 280 frames, approximately. Then, the motion of the test chart starts and the absolute value of

131
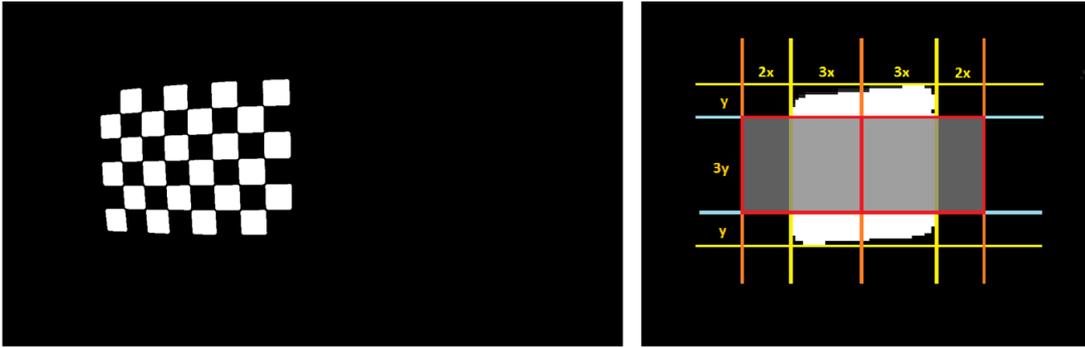
Figure 6.17: The Mask of the Black Boxes (left) and Cropping Details of ROI (left)
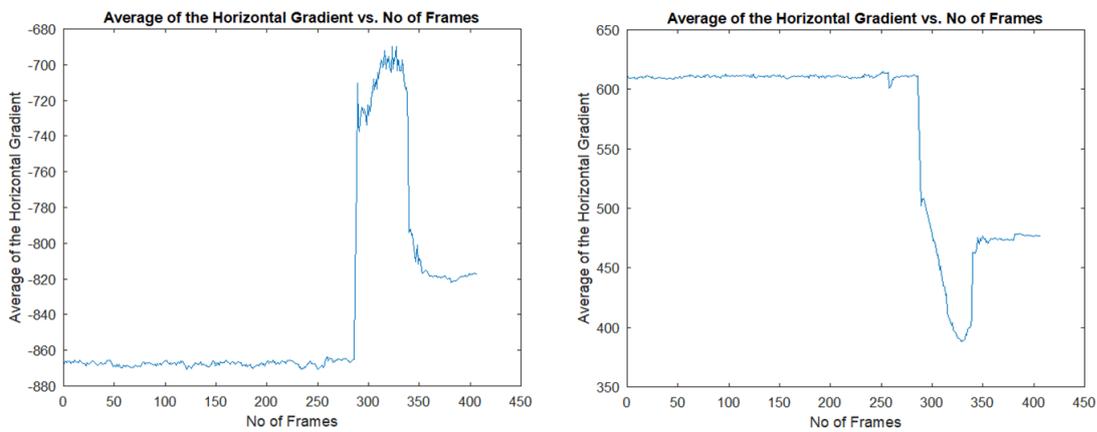
Figure 6.18: The Left Edge Graph and the Right Edge Graph

the gradient decreases for both right and the left edge. The motion ends around 350th frame and at the rest of the frames, the chart stands still. As mentioned in the previous part, the illumination on the test chart does not stay the same between the first and the last position of the test chart due to the fact that only one projector is used in the test setup. The gradient change caused by the change in the illumination can be observed clearer when the speed of the test chart is decreased. For example, when the test chart speed is reduced to 2 cm/s, the result shown in Figure 6.19 is obtained. Hence, the speed of the test chart is slow in this case, the camera does not experience motion blur and the only change in the gradient occurs because of the illumination change.

As mentioned in the background information part, studies show that the motion blur increases when the speed of the moving object increases as expected [37]. In the light of this information, the same effect should be observable from the proposed test. For

132

this purpose, the test procedure which is explained above is repeated by using the same camera for the different test chart speeds. The results are given in Figure 6.19 and Figure 6.20. As mentioned before, for slow motions like 2 cm/s, the motion blur does not affect the frames. If it is examined carefully, the gradient changes from the starting point to the endpoint linearly because of the illumination change. However, if the speed of the test chart is increased to 5, 10 and 15 cm/s, then the change in the gradient values becomes more and more. This behavior proves that our approach to the motion blur is consistent with the previous works in the literature [111].
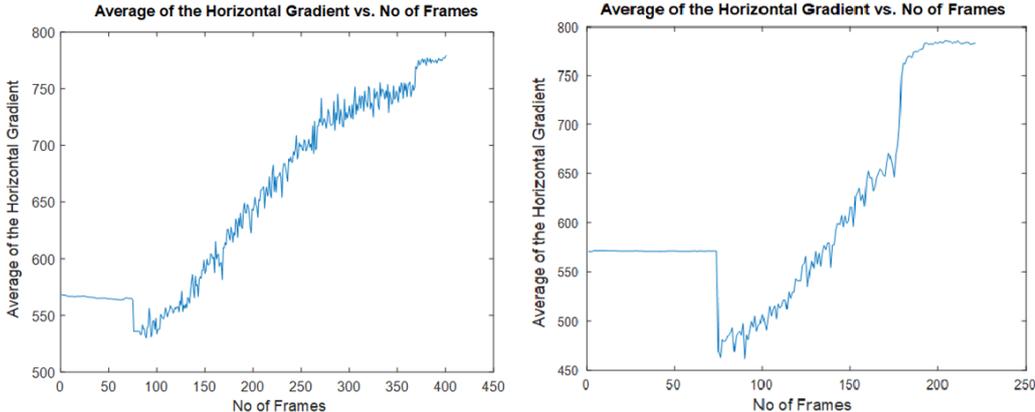
Figure 6.19: The Right Edge Characteristics for 2 cm/s (left) and 5 cm/s (right)
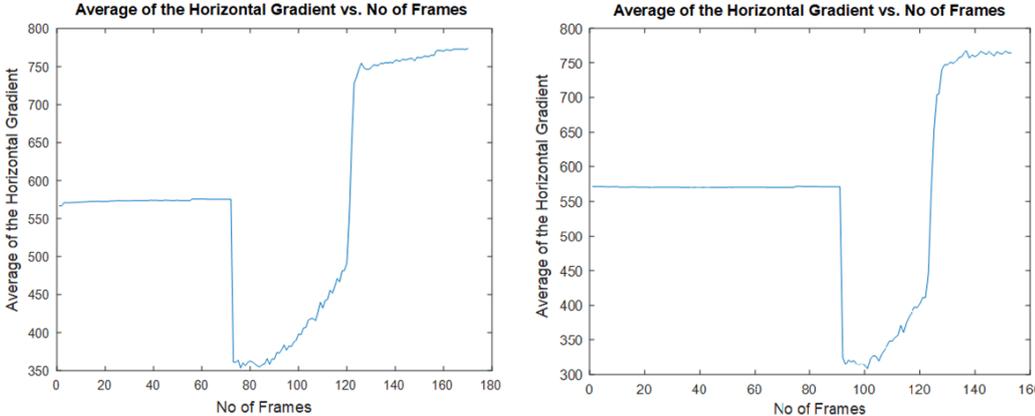
Figure 6.20: The Right Edge Characteristics for 10 cm/s (left) and 15 cm/s (right)

As explained above, the increasing speed of the test chart increases also the motion blur in the video. On the other hand, for slow test chart speed, it is impossible to observe the motion blur. In order to observe the effect of the motion blur clearly, the test chart speed is chosen as 15 cm/s.

As the last step of the test details, the right and left edge characteristics of the cameras should be turned into the numerical values such that the performances of the cameras can be compared. As mentioned above, it is observed that the gradient changes linearly from the starting point to the end point for the slow test chart speeds. Thus, this response is considered as the ideal characteristic. For this purpose, the numbers of frames in which the motion of the test chart starts and ends are detected from the peaks of the edge graphs. Then, the linear equation which passes through these points is determined. Only the frames in which the test chart is moving are examined and the measured gradient values in these frames are compared with the corresponding values in the ideal characteristic. The average value of the difference between the measured gradient value and the ideal value is calculated. Lastly, the ratio of this average value with the average of the gradient values of the starting and the end frames is determined. This ratio is used as the performance metric. If the ratio has a small value, it indicates that deviation from the ideal characteristic is also small and thus, the effect of the motion blur is less. For the results given in Figure 6.18, the obtained graphs are merged with the ideal characteristics and it is shown in Figure 6.21.
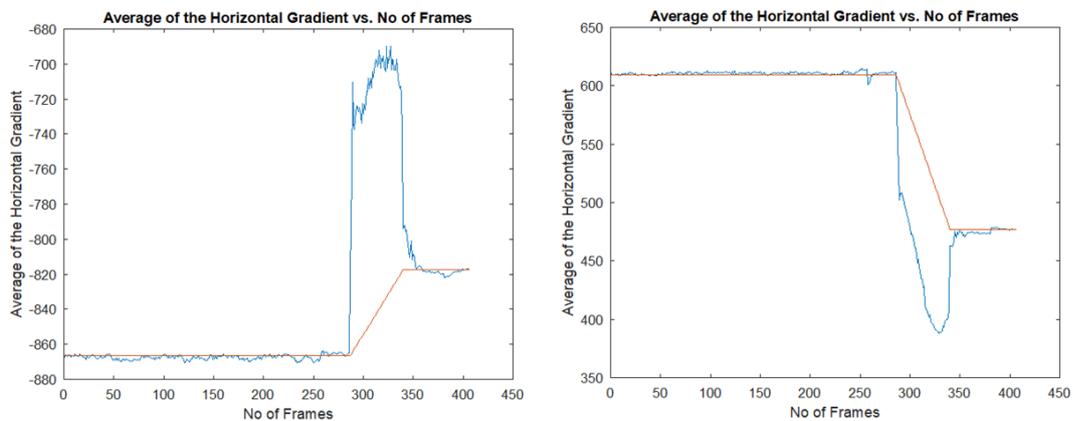
Figure 6.21: The Left Edge Graph and the Right Edge Graph with Ideal Responses

6.5 Test Results

The test software whose details are explained in the previous chapter is used in order to determine the motion blur performance of eight cameras under test. The obtained results will be given in this chapter. However, before giving these results, it is impor-

134

tant to explain the reason for the existence of the metronome on the scene.

At the beginning of our studies about the motion blur, we proposed the same test setup and test software as explained above but without the metronome. When the frames are examined carefully, for some test cameras, the extreme distortion is observed in the first few frames after the motion of the test chart is started. In the rest of the motion, the distortion decreases as shown in Figure 6.22. The shape of the distortion indicates that the processor inside the camera performs some image processing operations while forming the video. Just as a guess, when there is no motion in the scene, the camera uses averaging or median filtering between the frames in order to decrease the noise. Thus, the first frames of the motion carry the effect of the previous frames which distorts the video. However, when the camera detects the motion, this averaging operation is interrupted and the distortion decreases. It can be observed more clearly in Figure 6.23.

Figure 6.22: The First Frame, The Most Distorted Frame, Normally Distorted Frame and The Last Frame (respectively, from left to right)
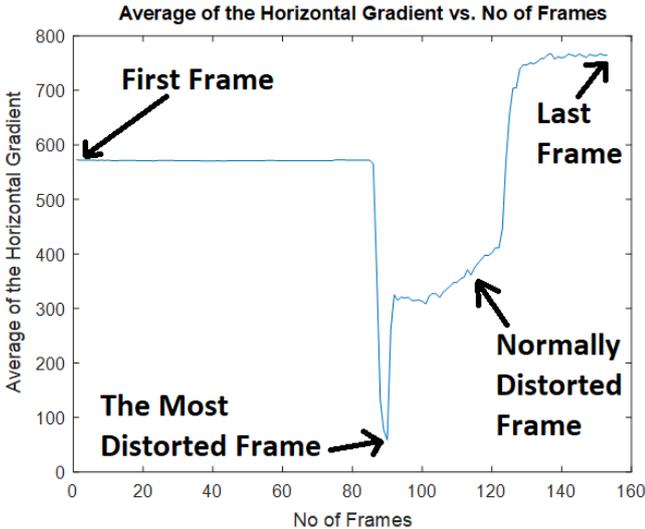
Figure 6.23: The Gradient Change on the Right Edge

135

The observation explained above leads us to add a metronome to the proposed test setup in order to assure that there would be no time without motion during the recording. Thanks to the existence of the metronome, the cameras which perform post-processing when there is no motion do not experience extreme distortions like shown above. Hence, these peaks vanish in the results which will be given below.

In the rest of this chapter, the results which are obtained by the explained procedure will be given for all the cameras under test. The blue plots represent the mean gradient value for the left and right edges of the black boxes in the test chart. The red lines represent the ideal characteristics for the average of left or right edge like there is no motion but only the gradient change due to the illumination change.

6.5.1 The results of Camera-1

According to our motion blur test, for the left edge and right edge characteristics can be observed in Figure 6.24 for Camera-1.
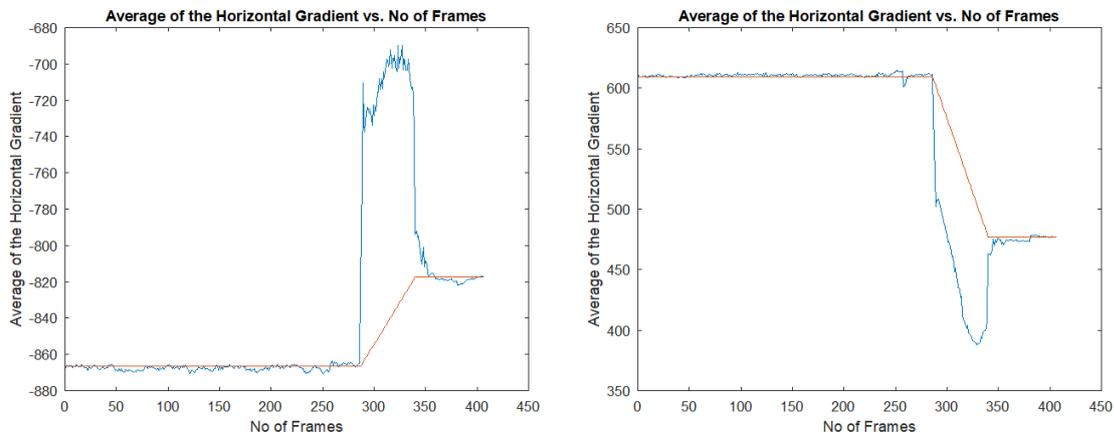
Figure 6.24: The Gradient Characteristics of Left (left) and Right Edge (right) of Camera-1

6.5.2 The results of Camera-2

The obtained left and right edge characteristics can be seen in Figure 6.25. This example proves that the different edges in the image can experience different motion

blur.
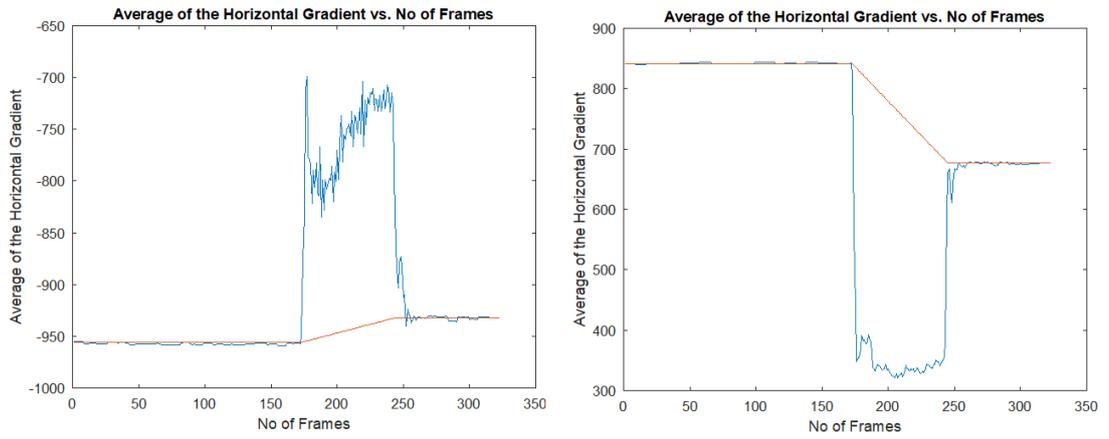
Figure 6.25: The Gradient Characteristics of Left (left) and Right Edge (right) of Camera-2

6.5.3   The results of Camera-3

The left and right edge characteristics of Camera-3 are given in Figure 6.26.
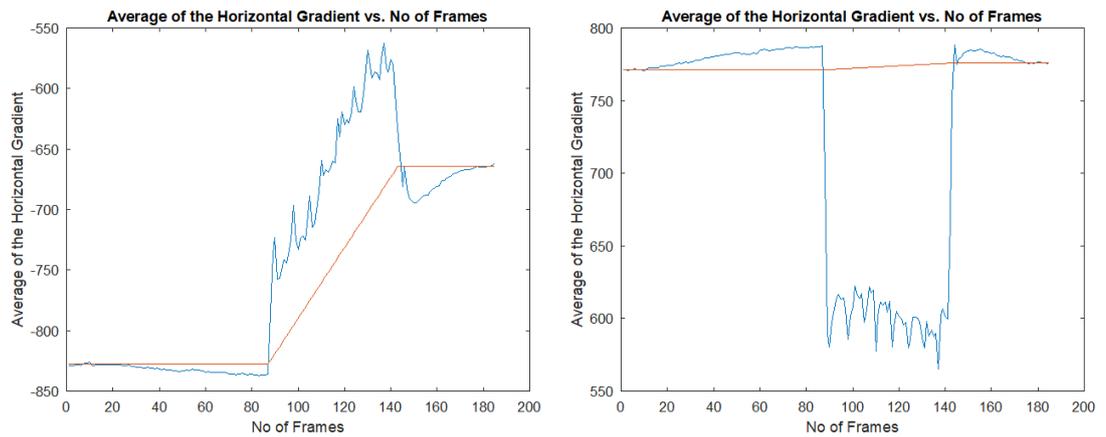
Figure 6.26: The Gradient Characteristics of Left (left) and Right Edge (right) of Camera-3

6.5.4   The results of Camera-4

The obtained test results for Camera-4 are shown in Figure 6.27.
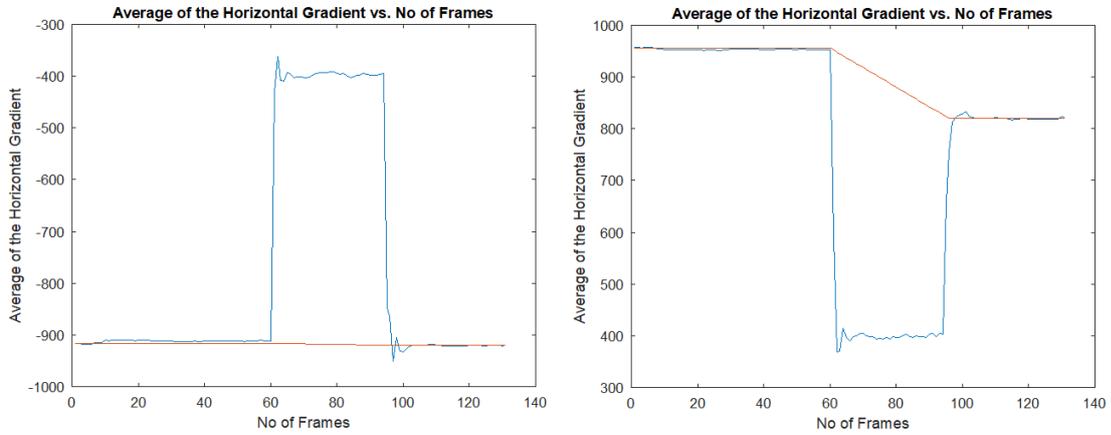
Figure 6.27: The Gradient Characteristics of Left (left) and Right Edge (right) of Camera-4

6.5.5   The results of Camera-5

The test results of Camera-5 which are given in Figure 6.28 show that it has the worst performance when there is motion in the scene.
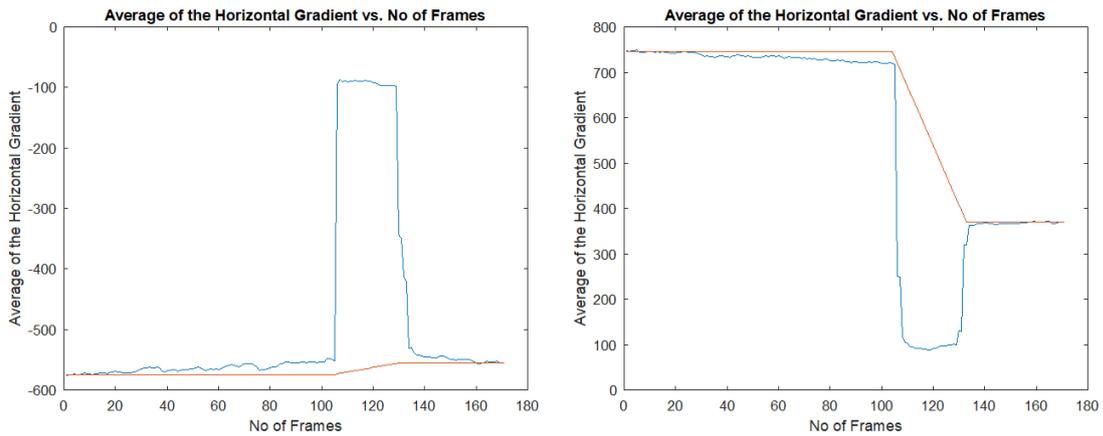
Figure 6.28: The Gradient Characteristics of Left (left) and Right Edge (right) of Camera-5

6.5.6 The results of Camera-6

The obtained results by using the Camera-6 are given in Figure 6.29.
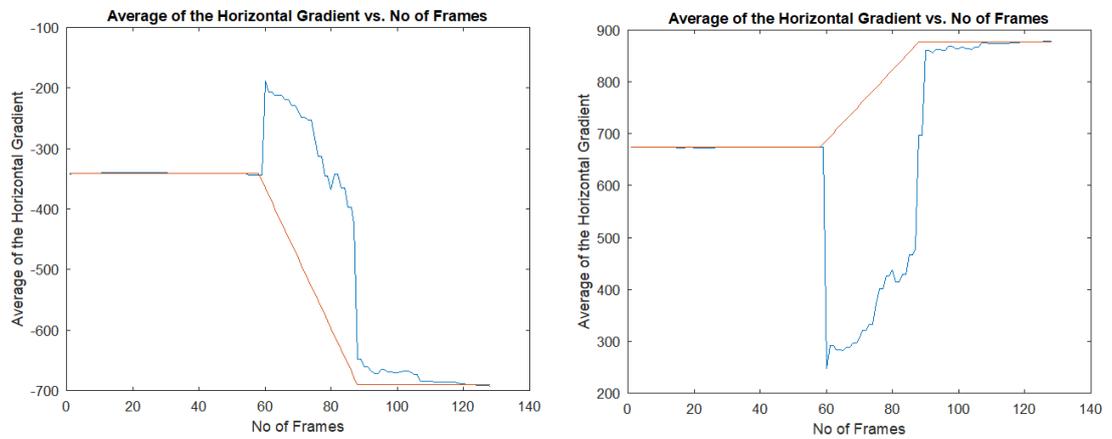
Figure 6.29: The Gradient Characteristics of Left (left) and Right Edge (right) of Camera-6

6.5.7 The results of Camera-7

The left and the right edge characteristics of Camera-7 in the existence of the motion can be seen in Figure 6.30..
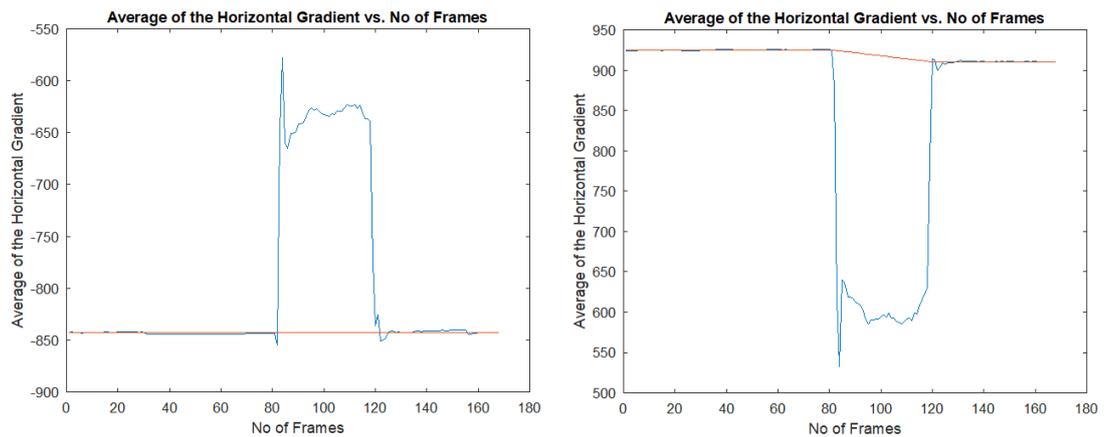
Figure 6.30: The Gradient Characteristics of Left (left) and Right Edge (right) of Camera-7

6.5.8 The results of Camera-8

The test results of Camera-8 are shown in Figure 6.31.
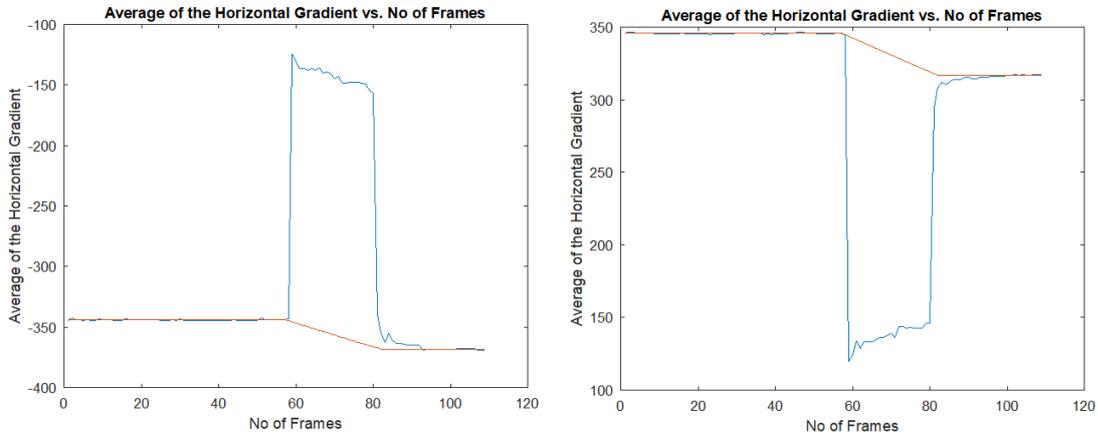
Figure 6.31: The Gradient Characteristics of Left (left) and Right Edge (right) of
Camera-8

The given results for all the test cameras given above are gathered into a single table
which is shown in Figure 6.32 in order to increase the readability. According to the
results, while Camera-3 gives the best motion blur performance, Camera-5 experi-
ences the motion blur, extremely.
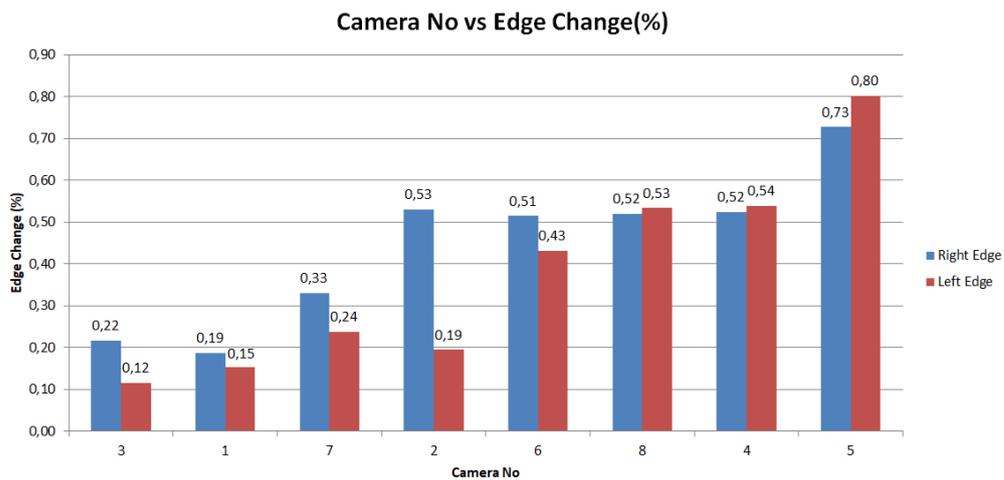
Figure 6.32: The Motion Blur Performance Comparison of the Test Cameras

CHAPTER 7

CONCLUSION AND FUTURE WORK

The main purpose of this thesis is to prepare automatic camera performance tests which can be easily used by any user without almost any interaction. The followed method can be summarized through two paths. If there is a common consensus in the literature about a measurement procedure of the camera performance property such as a standard, then the implementation of the previously proposed method is followed by minimizing the user interaction in order to prevent any user-related errors and offer user-friendly software. The consistency of these test results with the other software available in the market is important to prove the fidelity of the methods. These results are given in the corresponding chapters for the color accuracy and sharpness tests. On the other hand, if there is no totally accepted method which is followed by the most of the society, then a test procedure is proposed in the light of the information in the literature. The proposed test procedure includes all the details about the test setup, test chart, and evaluation process. The evaluation algorithm and software is presented again by minimizing the user interaction. The proposed tests for dynamic range, color constancy, change sensitivity and motion blur can be found in Chapter Two, Five and Six. Due to the fact that there is almost no performance comparison method by using the video acquired by the test cameras, proposing the new procedures was inevitable. It is important to note that the performance measurement tests are performed in the prepared dark room in Multimedia Research Laboratory of METU in order to ensure the fixed environmental conditions. Only the motion blur tests are done in the Laboratory of Robotics and Autonomous Systems at METU. These tests can be completed in a day for a test camera.

To conclude, the results of the dynamic range tests show that the proposed method

determines the dynamic range correctly even under the different environmental conditions. The obtained result of the color accuracy tests matches exactly same with the commercial software, Imatest. Furthermore, a typical method to obtain the color correction matrix is implemented and the color accuracy results of the color corrected images prove that the color correction matrix works fine. In the last test about the still image performance of the cameras, the sharpness of the cameras is measured and the consistency between the acquired results and the commercial software of Imatest GmbH verifies the accuracy of the tests. Lastly, in order to measure the video performances of the cameras, two well-defined test procedures are proposed which are based on change sensitivity and motion blur. The stability of the test results indicates the reliability of our work.

It is important to note that the dynamic range tests in the literature are also performed by using a transparent chart. However, in this method, acquiring the raw data from the camera is necessary. Due to the fact that all the test cameras used in this thesis are security cameras that we cannot reach their raw data, we proposed a test method based on the comparison of the dynamic range values of the cameras by using a reflective chart. However, in future, we plan to prepare a test setup by using suitable lightbox in order to perform dynamic range measurement with transparent test chart. Employing another test chart which has more patches can be noted as the future work on the color tests. Also, observing the relatioship with the color accuracy and the opto-electrical conversion function (OECF) could be interesting if the raw data obtained by the camera is reachable. Utilization of the new version of the test charts of ISO 12233:2017 standard can be considered as the future work on camera sharpness tests. In the motion blur tests, due to the usage of only one light source, the gradient change between the first and the last position of the test chart is assumed as linear as indicated in the related chapter. As the last future work, we plan to perform these tests in the dark room in MMRG Laboratory by using two light sources. In such case, the only gradient change is expected in the frames that the test chart moves.

As noted before, the scope of this thesis is the performance comparison of the cameras. Thus, the different evaluation properties of the cameras are the main concern throughout the studies. On the other hand, the end user should consider the price of the camera besides its performance while deciding the best choice according to the

Camera No	Manufacturer	Model	Price($)
1	Bosch	NBN-50022-V3	395
2	Bosch	NBN-733V-IP	765
3	Bosch	NBN-80052-BA	1020
4	LG	LW352-F	250
5	Pelco	IXE20DN	955
6	Pelco	IME319-1EP	590
7	Samsung	SNB-5004P	410
8	Vivotek	IP816A-HP	675

Table 7.1: Camera Models and Prices

application. Due to the fact that this decision between the performance and the price is subjective, any comment will be given about the prices of the cameras. However, it is important to give the prices of the test cameras whose test results are shared in this thesis in order to provide insight to the readers. The camera numbers used in the thesis and the prices of the cameras taken from www.surveillance-video.com in May 2018 can be found in Table 7.1.

REFERENCES

[1] "The Imatest Test Lab." `http://www.imatest.com/docs/lab/`. Accessed: 2018-05-06.

[2] D. Wueller, "Image-Engineering Digital Camera Tests," tech. rep., Image-Engineering GmbH, 2006.

[3] K. Schurman, "How Does a Digital Camera Work?." `http://www.gadgetreview.com/how-does-a-digital-camera-wor`. Accessed: 2018-05-06.

[4] "Dynamic Range in Digital Photography." `https://www.cambridgeincolour.com/tutorials/dynamic-range.htm`. Accessed: 2018-05-06.

[5] "KODAK Color Separation Guides and Gray Scales." `https://www.kodak.com/it/it/motion/products/lab_and_post_production/control_tools/kodak_color_separation_guides_and_gray_scales/default.htm`. Accessed: 2018-05-06.

[6] "Stouffer Step Wedge R2110." `https://www.asmetec-shop.de/Stouffer-R2110.html`. Accessed: 2018-05-06.

[7] "Using Stepchart." `http://www.imatest.com/docs/stepchart/#aicharts`. Accessed: 2018-05-06.

[8] "Taksim Meydan Tüneli, Istanbul." `http://www.kitokogroup.com/?s=project&id=7`. Accessed: 2018-05-06.

[9] "Spectroscopy." `https://2012grade10.wikispaces.com/Spectroscopy`. Accessed: 2018-05-06.

[10] B. Wandell, *Foundations of Vision*. Sinauer Associates, 1995.

[11] "What is color constancy?." `http://colorconstancy.com/?page_id=9`. Accessed: 2018-05-06.

[12] D. Scoles, Y. Sulai, C. Langlo, G. Fishman, C. Curcio, J. Carroll, and A. Dubra, "In vivo imaging of human cone photoreceptor inner segments," *Investigative Opthalmology and Visual Science*, vol. 55, June 2014.

[13] "Light and the eye." `https://www.handprint.com/HP/WCL/color1.html`. Accessed: 2018-05-09.

[14] "How Digital Caemras Work." `https://electronics.howstuffworks.com/cameras-photography/digital/digital-camera4.htm`. Accessed: 2018-05-09.

[15] T. Darrel, "Computer vision and applications."

[16] "CIE 1931 color space." `https://en.wikipedia.org/wiki/CIE_1931_color_space`. Accessed: 2018-05-09.

[17] "HDR and Color Spaces." `https://blogs.msdn.microsoft.com/billcrow/2007/10/25/hdr-and-color-spaces/`. Accessed: 2018-05-09.

[18] "Enter the L*a*b*oratory: Making Color Spatial." `https://www.nazdar.com/en-us/News-Press-Releases/ArtMID/4163/ArticleID/224`. Accessed: 2018-05-09.

[19] "X-Rite ColorChecker Classic." `https://www.amazon.com/X-Rite-MSCCC-ColorChecker-Classic/dp/B000JLO31C`. Accessed: 2018-05-09.

[20] "Sharpness, Acutance and Resolution." `http://www.photoreview.com.au/tips/shooting/sharpness,-acutance-and-resolution`. Accessed: 2018-05-12.

[21] "Acutance." `https://en.wikipedia.org/wiki/Acutance`. Accessed: 2018-05-12.

[22] "TUTORIALS: SHARPNESS." `https://www.cambridgeincolour.com/tutorials/sharpness.htm`. Accessed: 2018-05-12.

[23] "Limiting resolution and MTF." http://www.vision-doctor.com/en/optic-quality/limiting-resolution-and-mtf.html. Accessed: 2018-05-12.

[24] "Airy Disk." https://en.wikipedia.org/wiki/Airy_disk. Accessed: 2018-06-01.

[25] "The Rayleigh Criterion." http://hyperphysics.phy-astr.gsu.edu/hbase/phyopt/Raylei.html. Accessed: 2018-06-01.

[26] "How to measure mtf and other properties of lenses," tech. rep., Optikos Corporation, July 1999. Rev. 2.

[27] "THE MTF TEST AND MEASUREMENT COMPANY." http://www.image-science.com/. Accessed: 2018-05-12.

[28] B. Atkins, " Modulation Transfer Function (MTF) and Subjective Quality Factor (SQF)." http://www.bobatkins.com/photography/technical/mtf/mtf1.html. Accessed: 2018-05-12.

[29] "Sharpness: What is it and how is it measured?." http://www.imatest.com/docs/sharpness/#imatest_calc. Accessed: 2018-05-12.

[30] "Acutance and SQF (Subjective Quality Factor)." http://www.imatest.com/docs/sqf/. Accessed: 2018-05-12.

[31] "ISO 12233 — Resolution and spatial frequency responses." http://www.imatest.com/solutions/iso-12233/. Accessed: 2018-05-12.

[32] "Digital Capture Resolution Measurements." http://www.imaging.org/site/IST/Standards/Digital_Camera_Resolution_Tools/IST/Standards/Digital_Camera_Resolution_Tools.aspx?hkey=f9040928-44d3-411a-9594-98ee6fd81e69. Accessed: 2018-05-12.

[33] "How to Measure Modulation Transfer Function." http://harvestimaging.com/blog/?p=1294. Accessed: 2018-05-12.

[34] P. D. Burns, "Slanted-edge mtf for digital camera and scanner analysis," *Proc. IS & T 2000 PICS Conference*, pp. 135–138, 2000.

[35] S. C. Huang, "An advanced motion detection algorithm with video quality analysis for video surveillance systems," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 21, January 2011.

[36] "The Illusion of Motion." `https://paulbakaus.com/tutorials/performance/the-illusion-of-motion/`. Accessed: 2018-05-13.

[37] G. Gultekin, "Single and multi-frame motion deblurring for legged robots: Characterization using a novel fd-aroc performance metric and a comprehensive motion-blur dataset," September 2016.

[38] "Using Shutter Speed to Freeze or Blur Motion." `https://www.digitalphotomentor.com/motion-shutter-speed-freeze-blur/`. Accessed: 2018-05-13.

[39] "Shutter Speed: A Beginner's Guide." `https://www.photographymad.com/pages/view/shutter-speed-a-beginners-guide`. Accessed: 2018-05-13.

[40] Standard of Camera & Imaging Products Association, "Measurement and description method for image stabilization performance of digital cameras," *CIPA DC-X011*, 2014.

[41] Uwe Artmann and Philipp Feldker;, "Image stabilization performance: Existing standards and the challenges for mobile imaging," *Electronic Imaging Conference*, 2016.

[42] U. Artmann, "Image quality evaluation using moving targets," in *Proceedings of SPIE*, vol. 8667, 2013.

[43] "Sobel Edge Detector." `http://homepages.inf.ed.ac.uk/rbf/HIPR2/sobel.htm`. Accessed: 2018-05-06.

[44] "imfindcircles Function." `https://de.mathworks.com/help/images/ref/imfindcircles.html`. Accessed: 2018-05-06.

[45] T. Atherton and D. Kerbyson, "Size invariant circle detection," *Image and Vision Computing*, vol. 17, pp. 795–803, August 1998.

[46] "imdilate Function." `https://de.mathworks.com/help/images/ref/imdilate.html`. Accessed: 2018-05-11.

[47] "Morphological Image Processing." `https://www.cs.auckland.ac.nz/courses/compsci773s1c/lectures/ImageProcessing-html/topic4.htm#erosion`. Accessed: 2018-05-11.

[48] "Histogram Eqaualization." `https://en.wikipedia.org/wiki/Histogram_equalization`. Accessed: 2018-05-11.

[49] "Unsharp Mask." `https://docs.gimp.org/2.2/en/plug-in-unsharp-mask.html`. Accessed: 2018-05-12.

[50] "Unsharp Masking." `https://commons.wikimedia.org/wiki/File:Usm-unsharp-mask.png`. Accessed: 2018-05-12.

[51] H. E. Jonathan B. Phillips, *Camera Image Quality Benchmarking*. John Wiley & Sons, 2018.

[52] M. Pagnutti, R. E. Ryan, G. Cazenavette, M. Gold, R. Harlan, E. Leggett, and J. Pagnutti, "Laying the foundation to use raspberry pi 3 v2 camera module imagery for scientific and engineering purposes," *Journal of Electronic Imaging*, vol. 26, February 2016.

[53] C. Man, "Quantization noise: An expanded derivation of the equation, snr=6.02n+1.76 db," tech. rep., Analog Devices Inc., August 2012. MT-229.

[54] P. W. Wong and Y. H. Lu, "A method for the evaluation of wide dynamic range cameras," *Proceedings of SPIE*, vol. 8299, 2012. Digital Photography 8.

[55] "Dynamic Range." `http://www.imatest.com/solutions/dynamic-range/`. Accessed: 2018-05-06.

[56] International Telecommunication Union, "Studio encoding parameters of digital television for standard 4:3 and wide-screen 16:9 aspect ratios," *Recommendation ITU-R BT.601-7*, March 2011.

[57] "Güvenlik kamera sistemi kurulumu." `https://www.ilan.gov.tr/detay-i%CC%87hale-duyurulari-hizmet-guvenlik-kamera-sistemi-kurulumu.html?archived=true`. Accessed: 2018-05-06.

[58] "NAV-T 70 W SUPER 4Y VIALOX NAV-T SUPER 4Y," tech. rep., OSRAM, 2017. High-pressure sodium vapor lamps for open and enclosed luminaires.

[59] "What is Color?." `http://www.devx.com/projectcool/Article/19954`. Accessed: 2018-05-06.

[60] "What is Light?." `http://www.andor.com/learning-academy/what-is-light-an-overview-of-the-properties-of-light`. Accessed: 2018-05-06.

[61] "Rods and Cones." `https://www.cis.rit.edu/people/faculty/montag/vandplite/pages/chap_9/ch9p1.html`. Accessed: 2018-05-06.

[62] "Rods and Cones." `http://hyperphysics.phy-astr.gsu.edu/hbase/vision/rodcone.html`. Accessed: 2018-05-06.

[63] A. Stockman, D. I. A. MacLeod, and N. E. Johnson, "Spectral sensitivities of the human sensitive cones," *Journal of the Optical Society of America*, vol. 10, pp. 2491–2521, 1993.

[64] "Tristimulus system." `https://www.britannica.com/science/tristimulus-system`. Accessed: 2018-05-09.

[65] "Color constancy." `http://psychology.wikia.com/wiki/Color_constancy`. Accessed: 2018-05-09.

[66] E. H. Land and J. J. McCann, "Lightness and retinex theory," *Journal of the Optical Society of America*, vol. 61, pp. 1–11, 1971.

[67] S. Bianco, C. Cusano, and R. Schettini, "Color constancy using cnns," *DeepVision: Deep Learning in Computer Vision 2015*, April 2015. arXiv:1504.04548.

[68] H. Lei, G. Jiang, and L. Quan, "Color constancy with derivative colors," November 2016. arXiv:1611.08389.

[69] J. T. Barron and Y.-T. Tsai, "Fast fourier color constancy," November 2016. arXiv:1611.07596.

[70] J. Guild, "The colorimetric properties of the spectrum," *Philosophical Transactions of the Royal Society of London*, vol. Series A. 230, pp. 149–187, 1932.

[71] W. D. Wright, "A re-determination of the trichromatic coefficients of the spectral colours," *Transactions of the Optical Society*, vol. 30, pp. 141–164, 1929.

[72] W. D. Wright, "A re-determination of the mixture curves of the spectrum," *Transactions of the Optical Society*, vol. 31, pp. 201–218, 1930.

[73] R. Davis and K.S.Gibson, "The relative spectral energy distribution and correlated color temperature of the n.p.l white-light standard," *Bureau of Standards journal of research*, vol. 7, pp. 791–797, 1931.

[74] A. Frich, "Color spaces." `https://www.color-management-guide.com/color-spaces.html`. Accessed: 2018-05-09.

[75] "Device-Independent Color Spaces." `https://msdn.microsoft.com/en-us/library/windows/desktop/dd316870(v=vs.85).aspx`. Accessed: 2018-05-09.

[76] D. Dias, "What is YCbCr ? (Color Spaces)." `https://medium.com/@danojadias/what-is-ycbcr-964fde85eeb3`. Accessed: 2018-05-09.

[77] Rec. ITU-R BT.601-5, "Studio encoding parameters of digital television for standard 4:3 and wide-screen 16:9 aspect ratios," 1982-1986-1990-1992-1994-1995. Section 3.5.

[78] A. R. Robertson, "Historical development of cie recommended color difference equations," *Color Research and Application*, vol. 15, no. 3, p. 167–170, 1990.

[79] "How to convert between sRGB and CIEXYZ." `https://www.image-engineering.de/library/technotes/958-how-to-convert-between-srgb-and-ciexyz`. Accessed: 2018-05-09.

[80] C. S. McCamy, H. Marcus, and J. G. Davidson, "A color-rendition chart," vol. 2, no. 3, pp. 95–99.

[81] "New color specifications for ColorChecker SG and Classic Charts." `https://www.xrite.com/service-support/new_color_specifications_for_colorchecker_sg_and_classic_charts`. Accessed: 2018-05-09.

[82] "Tips for Capturing the ColorChecker Classic Target." `http://xritephoto.com/ph_product_overview.aspx?ID=1192&Action=Support&SupportID=4991`, April 2012. Accessed: 2018-05-09.

[83] S. Wolf, "Color correction matrix for digital still and video imaging systems," *National Telecommunications and Information Administration Technical Memorandum*, December 2003. TM-04-406.

[84] "Color Correction Matrix (CCM)." `http://www.imatest.com/docs/colormatrix/`. Accessed: 2018-05-11.

[85] "Introduction to Colour Spaces." `https://www.colourphil.co.uk/lab_lch_colour_space.shtml`. Accessed: 2018-05-11.

[86] Airy, G. B., "On the diffraction of an object-glass with circular aperture," *Transactions of the Cambridge Philosophical Society*, vol. 5, 1835.

[87] Lord Rayleigh, F.R.S., "Investigations in optics, with special reference to the spectroscope," *Philosophical Magazine*, vol. 8, no. 49, 18379.

[88] "Angular Resolution." `https://en.wikipedia.org/wiki/Angular_resolution#Explanation`. Accessed: 2018-06-01.

[89] "Trioptics." `https://www.trioptics.com/`. Accessed: 2018-06-01.

[90] P. D. Burns, "SFR evaluation for digital cameras and scanners." `http://losburns.com/imaging/software/SFRedge/sfrmat3_post/index.html`. Accessed: 2018-05-12.

[91] E.M.Grainger and K.N.Cupery, "An optical merit function (sqf) which correlates with subjective image judgments," *Photographic Science and Engineering*, vol. 16, no. 3, pp. 221–230, 1972.

[92] J. L. Mannos and D. J. Sakrison, "The effects of a visual fidelity criterion on the encoding of images," *IEEE Transactions on Information Theory*, vol. 20, no. 4, pp. 525–535, 1974.

[93] O. Kurtsev, "Basic steps in calculating the MTF (SFR), and comparing it with the open-source tools sfrmat3 and Mitre SFR 1.4." `http://www.quickmtf.com/slantededge.html`. Accessed: 2018-05-12.

[94] "A Video Database for Testing Change Detection Algorithms." `http://changedetection.net/`. Accessed: 2018-05-12.

[95] "How Motion Detection Cameras Work." `https://www.networkworld.com/article/2761095/it-management/how-motion-detection-cameras-work.html`. Accessed: 2018-05-12.

[96] "Intelligent video technology," tech. rep., Panasonic System Networks Co., Ltd., 2014. Rev. 2.

[97] "Advanced Motion Detection." `https://www.vivotek.com/advanced-motion-detection/`. Accessed: 2018-05-12.

[98] "Axis video motion detection 4-user manual," tech. rep., Axis Communications AB, April 2017. Ver.M2.10.

[99] S. Ching, S. Cheung, and C. Kamath, "Robust techniques for background subtraction in urban traffic video," *EURASIP Journal on Applied Signal Processing*, vol. 14, pp. 2330–2340, 2015.

[100] C. R. Wren, A. Azarbayehani, T. Darrell, and A. P. Pentland, "Pfinder: Real-time tracking of the human body," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 19, p. 780–785, July 1997.

[101] M. Oral and U. Deniz, "Center of mass model: A novel approach to background modeling for segmentation of moving objects," *Image Vis. Comput.*, vol. 25, p. 1365–1376, August 2007.

[102] B. Shoushtarian and N. Ghasem-Aghaee, "A practical approach to real time dynamic background generation based on a temporal median filter," *J. Sci. Islamic Republic Iran*, vol. 14, no. 4, p. 351–362, 2003.

[103] O. Sindelar, "Using gyroscopes and accelerometers for removing blur in mobile phone photography," 2012.

[104] R. Safaee-Rad and M. Aleksic, "Handshake characterization and image stabilization for cell-phone cameras," *Proc. of SPIE-IS& T Electronic Imaging*, vol. 7241, 2009.

[105] S. H. Park and M. Levoy, "Gyro-based multi-image deconvolution for removing handshake blur," in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3366–3373, June 2014.

[106] Z. Hu, L. Yuan, S. Lin, and M. H. Yang, "Image deblurring using smartphone inertial sensors," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1855–1864, June 2016.

[107] O. Šindelář, F. Šroubek, and P. Milanfar, "A smartphone application for removing handshake blur and compensating rolling shutter," in *2014 IEEE International Conference on Image Processing (ICIP)*, pp. 2160–2162, October 2014.

[108] "Checkerboard Test Chart." http://store.imatest.com/checkerboard-test-chart.html. Accessed: 2018-05-13.

[109] "SensorSpace 3D Calibrator CD100." http://store.imatest.com/equipment/motion-generators/ss-3dcalibrator-cd100.html. Accessed: 2018-05-13.

[110] "Linear Servo Base Unit with Inverted Pendulum." https://www.quanser.com/products/linear-servo-base-unit-inverted-pendulum/. Accessed: 2018-05-13.

[111] N. T. Y. M. Shinji Nakagawa, Toshiya Nakaguchi, "Motion blur perception in various conditions of presented edge," 2008.

[112] E. H. Bareiss, "Multistep integer-preserving gaussian elimination," tech. rep., Argonne National Laboratory , May 1966. ANL-7213.

[113] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[114] D. Bradley and G. Roth, "Adapting thresholding using the integral image," *Journal of Graphics Tools*, vol. 12, no. 2, pp. 13–21, 2007.

[115] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Pearson, 2008. Third Edition.

[116] "GUIDE TO IMAGE SHARPENING." `https://www.cambridgeincolour.com/tutorials/image-sharpening.htm`. Accessed: 2018-05-12.

APPENDIX A

TOOLS USED FOR PROCESSING

A.1   The Circular Hough Transform

The Circular Hough Transform is used in order to find the location of the circles in a grayscale or a binary image. As the first step, the possible points on the circles are chosen by taking the gradient of the image. The Sobel gradient operators are used in order to find the edges by taking the convolution of the image with the following kernels in Figure A.1.

Figure A.1: The Sobel Gradient Operators [43]

The thresholding with a fixed value provides the discarding the edge points due to the noise. 0.05 is the thresholding value in order to find the location of the circles in the test chart of dynamic range test. Moreover, in the test chart, the circles are drawn by black on the white paper which is the information that can be used for finding the center points. The edge points which are brighter than the background are also discarded from the candidate edge points.

The success of the Circular Hough Transform depends on the estimation of the radius of the circles. The algorithm searches the given radius range for each edge points. For

a fix radius and a single edge point, the accumulator array is filled with the "votes" as shown in Figure A.2 [44]. If the procedure is performed for each edge pixel, the accumulator array would have a peak at the center of the circle as shown on the right side of Figure A.2.
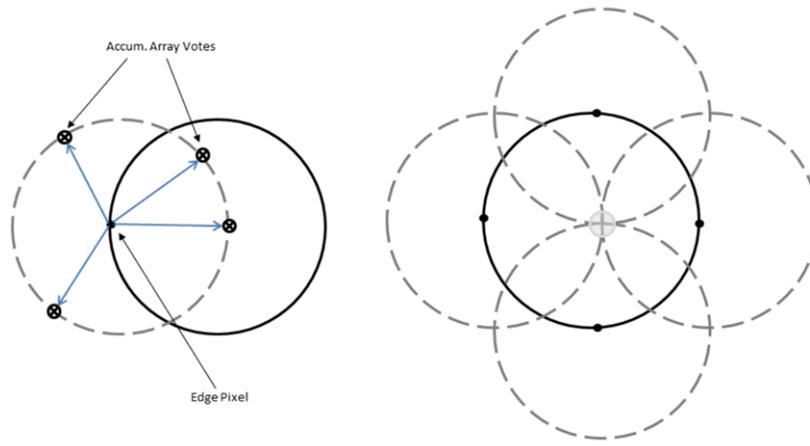
Figure A.2: Classical Circular Hough Transform Algorithm [44]

For the classical circular transform, the same calculations are done for each diameter value in the search range. Thus, it requires more memory, a three dimensional array, and the long processing time. Thanks to the work of Atherton and Kerbyson, the convolution of the thresholded binary image with a filter named "Phase Coded Annulus" makes possible to find the center of the circles by using less memory, a two dimensional array, and less time. Although, in the classical method, the convolution is done by using a single circle in many times, in the phase-coding method, the convolution is done at once. On the other hand, the filter used in the phase-coding method is a complex valued filter which can be computed by the following formula and the resulting filter can be seen in Figure A.3 [45].

$$Q_{PCA}(m,n) = \begin{cases} e^{-i\phi mn} & \text{,if } R_{min}^2 < m^2 + n^2 < R_{max}^2 \\ 0 & \text{,otherwise} \end{cases} \quad (A.1)$$

where

$$\phi mn = 2\pi \frac{(\log(\sqrt{(m^2 + n^2)}) - \log R_{min})}{(\log R_{max} - \log R_{min})} \quad (A.2)$$
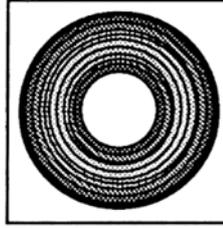
Figure A.3: The Phase-Coding Annulus Filter [45]

A.2 The Projective Spatial Transform

The projective transformation is one of the most common used transformations in the image processing area. Using projective transformation makes possible to handle the differences caused by the tilt of the image plane according to the object plane. In general, the projective transform does not preserve length but preserve the incidents and the cross ratio. The destination of a pixel can be find by the multiplication of the transformation matrix, A, with the pixel location in the original image.

$$
\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{pmatrix} * \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}
$$

where (u,v) is the destination of the pixel at the location (x,y) in the original image and a, b, c, d, e, f, g, h are the constants which will be determined for the transformation.

Finding the constants in the transformation matrix requires solving 8 equations which can be formed by using 4 points in the original image and their locations in the transformed version as shown in the formula below.

$$\begin{pmatrix} x_0 & y_0 & 1 & 0 & 0 & 0 & (-x_0*u_0) & (-y_0*u_0) \\ x_1 & y_1 & 1 & 0 & 0 & 0 & (-x_1*u_1) & (-y_1*u_1) \\ x_2 & y_2 & 1 & 0 & 0 & 0 & (-x_2*u_2) & (-y_2*u_2) \\ x_3 & y_3 & 1 & 0 & 0 & 0 & (-x_3*u_3) & (-y_3*u_3) \\ 0 & 0 & 0 & x_0 & y_0 & 1 & (-x_0*v_0) & (-y_0*v_0) \\ 0 & 0 & 0 & x_1 & y_1 & 1 & (-x_1*v_1) & (-y_1*v_1) \\ 0 & 0 & 0 & x_2 & y_2 & 1 & (-x_2*v_2) & (-y_2*v_2) \\ 0 & 0 & 0 & x_3 & y_3 & 1 & (-x_3*v_3) & (-y_3*v_3) \end{pmatrix} * \begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \end{pmatrix} = \begin{pmatrix} u_0 \\ u_1 \\ u_2 \\ u_3 \\ v_0 \\ v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

where

$$(x_0, y_0), (x_1, y_1), (x_2, y_2), (x_3, y_3)$$

are the pixel locations in the original image and

$$(u_0, v_0), (u_1, v_1), (u_2, v_2), (u_3, v_3)$$

are the new pixel locations in the transformed image, respectively.

The equations are in the form of Ax=b which can be easily solved by Gaussian Elimination Method [112]. The center points of the circles which are placed on the test chart of dynamic range test is used in order to find the transformation matrix such that the test chart is on the image plane without any tilt. This makes possible to crop the region of interest and to get rid of the rest. Furthermore, the transformation eliminates the depth information so; the calculations and the processing can be done by simply using Cartesian coordinates.

The every pixel in the original image is transferred to the new location in the transformed image and the rest of the pixels without any match are filled with a fixed value, 128 in 8 bit color space. On the other hand, this value does not affect the results of test because these filled pixels are out of the region of interest.

A.3 Otsu's Thresholding Method

Thresholding is one of the most commonly used methods in the literature in order to provide binarization. According to this method, the pixel values which have greater value than a certain level which is called threshold are marked as logic true (1) while the others are marked as logic false (0) [113]. Generally, it is used in order to make the differentiation between the foreground and the background in the image. Unlike the simple idea behind the thresholding, the problematic part is the determination of the threshold value. Although there are other methods in the literature, Otsu's method is one of the most popular ones. According to Otsu's method, the separation between the foreground and the background can be best performed by minimizing the intra-class variance. He proved that this minimization problem is the same problem with the maximizing the inter-class variance. The algorithm uses the histogram of the grayscale image, and the class probabilities can be calculated according to the following formulas where L is the range of the histogram, 255 for 8 bit representation.

$$w_0(t) = \sum_{i=0}^{t-1} p(i) \tag{A.3}$$

$$w_1(t) = \sum_{i=t}^{L-1} p(i) \tag{A.4}$$

The intra-class variance, σ_w^2, and the inter-variance σ_b^2, can be calculated by using the variance and the mean values of the classes as shown below.

$$\sigma_w^2(t) = w_0(t)\sigma_0^2(t) + w_1(t)\sigma_1^2(t) \tag{A.5}$$

$$\sigma_b^2(t) = w_0(t)w_1(t)[\mu_0(t) - \mu_1(t)]^2 \tag{A.6}$$

Due to the fact that finding the thresholding value turns into a maximization problem of σ_b^2, the formulas of the mean values are also necessary.

$$\mu_0(t) = \sum_{i=0}^{t-1} i * \frac{p(i)}{w_0} \tag{A.7}$$

$$\mu_1(t) = \sum_{i=t}^{L} i * \frac{p(i)}{w_1} \tag{A.8}$$

So, for all t values from 0 to 255, the class probabilities, w_i, and class means, μ_i, are calculated in order to determine σ_b^2. The chosen threshold value would have the maximum of the inter-class variance, σ_b^2.

It is important to note that Otsu's method is assumes that the image has only two classes which can be separated by a single threshold. However, using a single threshold value does not always perform well in order to differentiate the background and the foreground. Hence, there are other methods in order to improve the performance.

A.4   Bradley's Adaptive Thresholding Method

Unlike Otsu's method which uses a single threshold value throughout the whole image, Bradley proposed an adaptive thresholding method in 2007 [114]. According to this method, a window which has "s" pixels is used on the grayscale image and the threshold value for that pixel is determined according to its neighborhood. On the other hand, calculating the mean value of the neighborhood of each pixel increases the calculation time, significantly. Thus, the integral image concept is used for this purpose. The integral image can be prepared by the following pseudocode where "in" is the input image, "intImg" is the integral image, w is the width and h is the height of the input image.

Algorithm 3 Algorithm for Integral Image Formation

   for $i = 0$ to w **do**

       $sum = 0$

       for $j = 0$ to h **do**

           $sum = sum + in[i,j]$

           if $i = 0$ **then**

               $intImg[i,j] = sum$

           else

               $intImg[i,j] = intImg[i,j] + sum$

           end if

       end for

   end for

When the integral image is obtained at once, then the sum of the neighborhood of each pixel can be obtained by the summation of only 4 elements. If the upper left corner of the window is denoted as (x1,y1) and the lower right corner is denoted as (x2,y2), then the summation of the pixel values can be calculated by the integral image, I, values.

$$\sum_{x=x_1}^{x_2} \sum_{y=y_1}^{y_2} Grayscale(x,y) = I(x_2,y_2) - I(x_2,y_1-1) - I(x_1-1,y_2) + I(x_1-1,y_1-1)$$

(A.9)

The threshold is determined by dividing the summation to s2, and if the pixel value is less than this value than it is marked as logic low (0) and marked as logic (1), otherwise. Also in some cases, "t" percent of the determined threshold can be used as the corresponding threshold for that pixel. In our application image size over 8 is used as the window size, s. Thanks to the Bradley's method, the problems of Otsu's method for the images where illumination changes significantly is solved.

A.5 Morphological Operations

The morphological operations are very important and the most commonly used tools of image processing while working on the binary images. The dilation and the erosion operators are the basics of the others. The opening and the closing operations are the combination of the dilation and the erosion with the reverse order. The more detailed information will be given after the definition of the dilation and the erosion.

A.5.1 Dilation

The dilation operation dilates the outer border of the logic-high (1) region and increases the area covered by the logic high. It takes the union of copies of the structure element centered at every pixel of the logic-high region. The mathematical representation is given below where "I" is the logic high area and "H" is the structure element.

$$I \oplus H = (p+q) \mid p \in I, q \in H$$

(A.10)

Thanks to the dilation operation, the gaps between the different regions become smaller and the holes enclosed by a single region are filled.

A.5.2 Erosion

The erosion operation erodes the outer border of the logic-high (1) region and decreases the area covered by. Practically, the operation keeps the logic-high pixels of the region if the structure element fits inside the logic high area. The mathematical expression is given below where "I" is the logic high area, "H" is the structure element and "Z" is the output of the operation.

$$I \ominus H = p \in Z^2 \mid (p+q) \in I, \forall q \in H \tag{A.11}$$

Thanks to the shrinkage provided by the erosion operation, small regions are eliminated. Thus, it is commonly used in order to prevent the effect of the noise. An original image and its versions with the dilation and the erosion operations are given in Figure A.4.

Figure A.4: An Original Image (left), its Dilated Version (middle) and its Eroded Version (right) [46]

A.5.3 Opening Operation

In the opening operation, first the erosion and then the dilation operations are performed. It separates the regions connected with a thin bridge. The regions which survived after the erosion are restored by the dilation operation. Like all of the mor-

phological operations, choosing the structure element is important. An example opening operation and the effect of using different structure elements can be observed in Figure A.5.
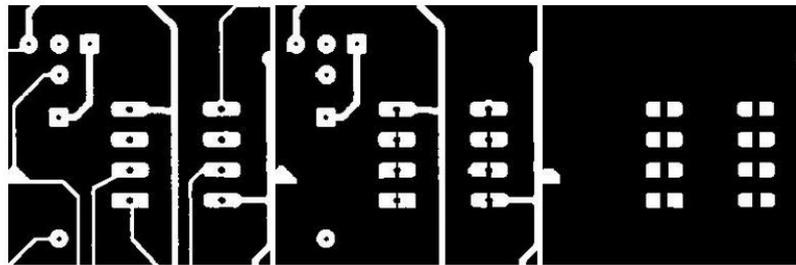

Figure A.5: The Original Image (left), its Opened Version by 5x5 Structure Element (middle) and its Opened Version by 9x9 Structure Element (right) [47]

A.5.4 Closing Operation

In the closing operation, first the dilation and then the erosion operations are performed. Thanks to the closing operation, the areas close to each other can be merged and the holes inside the logic-high (1) areas can be filled without distortion of the shape. An example of the closing operation can be seen in Figure A.6.
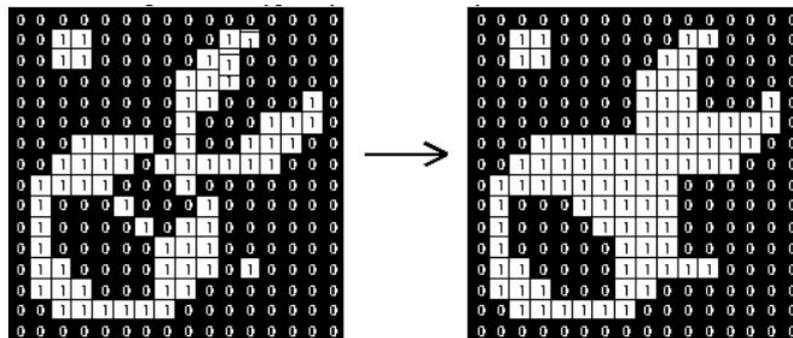

Figure A.6: An Example of Closing Operation [47]

A.6 Histogram Equalization

The histogram equalization is one of the most commonly used techniques in the literature in order to increase the contrast in the image. The idea is to achieve the homoge-

165

nously distribution of the most frequent intensity values in the image. This provides the areas with lower local contrast to gain higher contrast. Also, the histogram equalization is very useful for the images with too dark or too bright background in order to obtain the details in these areas. It uses the histogram of the image and the cumulative distribution function. Before applying the histogram equalization, the original image probably has some peaks in its histogram. The histogram equalization method tries to remap the pixel values into the new values and create a flat histogram. If the cumulative distribution function is considered; the histogram squeezed into a small region is represented with a stepwise cumulative distribution function while the flat histogram is a linear function as shown in Figure A.7.
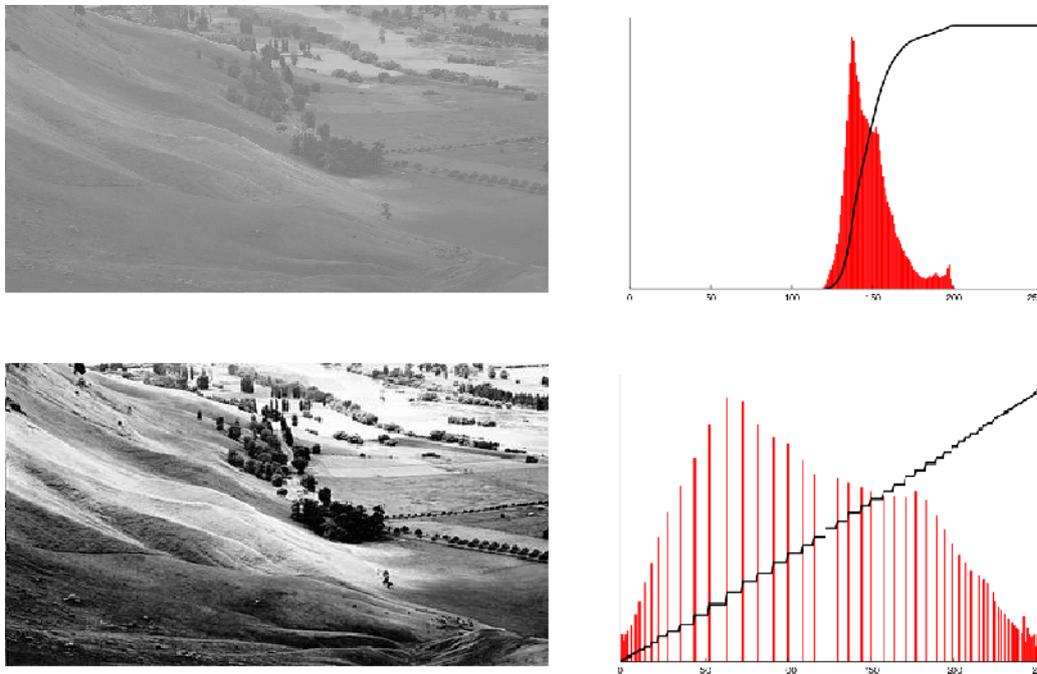

Figure A.7: An Example of Histogram Equalization with Their Histograms [48]

Calculating the cumulative distribution function of the input image is the first step of the histogram equalization [115]. The cumulative distribution function, $cdf_x(i)$, can be calculated by using the formula below by using the probability of occurrence of each pixel values, $p_x(i)$.

$$p_x(i) = \frac{n_i}{n}, 0 \leq i < L \tag{A.12}$$

$$cdf_x(i) = \sum_{j=0}^{i} p_x(j) \tag{A.13}$$

where n_i is the number of the pixels which have the value i, n is the total number of pixels, L is the total number of gray levels, 256 in our case. Once the cumulative distribution function is obtained, then the remapping of the pixel values into the new values are done by using the following formula, where cdf_{min} is the minimum non-zero value of the cumulative distribution function and "h" is the remapped value of pixel value "v".

$$h(v) = round(\frac{cdf_x(v) - cdf_{min}}{n-1}(L-1)) \tag{A.14}$$

A.7 The Unsharp Masking

Although the unsharp masking is not implemented and used in the scope of this thesis, it is necessary to give information in order to explain its effect on the edges of the image. Thus, its effect on the sharpness should be understood better.

The unsharp masking technique has been used in photography since the 1930s in order to increase the acutance of the image. After digital cameras became widespread, the analog technique has been adopted for the digital photography. Today, most of the cameras perform unsharp masking on the images automatically as a post-processing step in order to increase the perceived sharpness of the image. Although there are different methods for performing the unsharp masking, the idea and the main procedure is the same.

Before going into the details, it is important to note that there are 2 main variables in the unsharp masking technique which are radius and amount. Generally, these variables are determined by the manufacturer of the camera and their values changes according to the mode of the camera, like "indoor", "outdoor", "sunset", etc. [116]. As the first step, a Gaussian Blur is added on the copy of the input image. The radius of the Gaussian is one of the variables of the procedure. Please note that, practically the Gaussian Blur would have values different than the original image near the edge points. Then, the blurred image is subtracted from the original image. As stated, the result will have almost zero values near the low frequency parts of the image but have

nonzero values around the edges of the image. The resulting image is multiplied by the amount which is the second variable of the procedure. Lastly, the resulting image is added on the original image. The total formula can be created like as follows:

$$Sharpened = Original + (Original - Blurred)xAmount \qquad (A.15)$$

The effect of the unsharp masking can be shown near an edge easily. Let's assume that the original image has an edge in the image like the black line shown in Figure A.8. Its Gaussian blurred version would have more gradual increase near the edge like the green line in the same figure. When the subtraction of the blurred image from the original one is considered, it has nonzero values near the edges. It is important to note that the difference image would have negative values for the first half of the edge and have positive values for the second half of the edge. Thus, adding this difference image with the original image would result with the red line in Figure A.8. It can be observed that the slope of the sharpened edge is much higher than the edge in the original one and thus, its perceived sharpness is much higher than the original one.
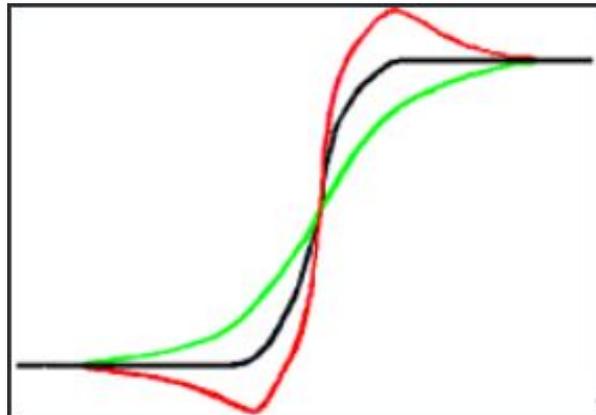

Figure A.8: The Illustration of the Unsharp Masking on an Edge [49]

Even though the idea of the unsharp masking is simple, in the real life applications it results with the increasing of the image noise especially in the low frequency regions. Thus, one more variable which is called threshold is proposed in order to perform "smart sharpen". With this modification, the unsharp masking method is used around the edge points which have greater brightness change than the threshold value.

An example pattern (on the upper) and its version with unsharp masking method (on the lower) are shown in Figure A.9. The method has been used with very high radius size and amplitude value in order to show the effect clearly. The lower part of the image can be perceived as sharper image. However, when it is examined closer, it is obvious that the pixel values near the edges have greater or smaller pixel values than the original one which can create MTF values greater than 1.
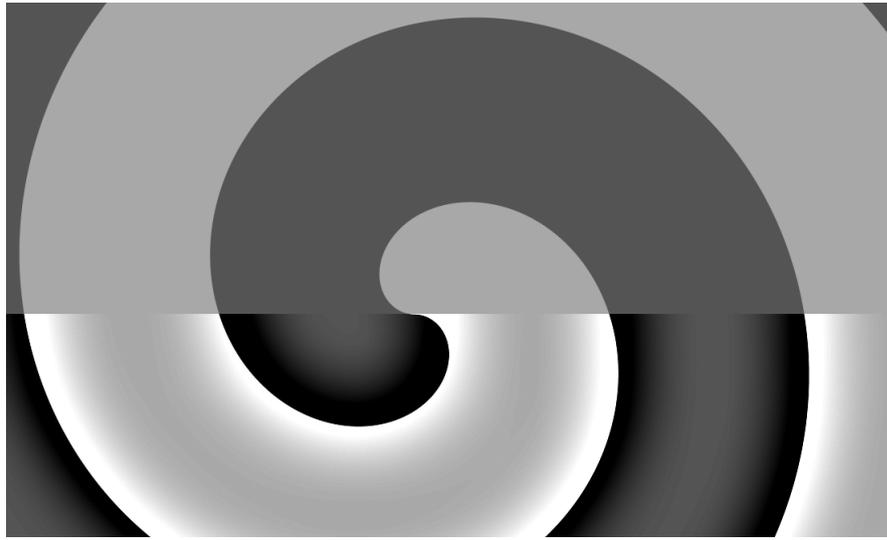

Figure A.9: The Original Image and Its Unsharpened Version [50]

A.8 Eccentricity

The eccentricity of the binary region can be defined as the ratio of the distance between the foci of the ellipse and its major axis length. Thus, it can take values between 0 and 1 while 0 indicates the perfect circle and 1 represents the line. Although there are several ways in the literature to calculate the eccentricity of a binary region; using the moments of the region provides a straightforward method. In order to calculate the moments, firstly the centroid of the region should be found.

$$Centroid = (\bar{x}, \bar{y})$$

$$where \quad \bar{x} = \frac{\sum_x \sum_y x}{\sum_x \sum_y 1} \quad and \quad \bar{y} = \frac{\sum_x \sum_y y}{\sum_x \sum_y 1} \tag{A.16}$$

The necessary central moments for the calculation of the eccentricity can be determined as follows:

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^2 * (y - \bar{y})^2 * f(x, y) \tag{A.17}$$

And lastly, the eccentricity of a binary region can be calculated by using the formula below:

$$Ecc(Region) = \frac{\mu_{20} + \mu_{02} + \sqrt{(\mu_{20} - \mu_{02})^2 + 4 * \mu_{11}^2}}{\mu_{20} + \mu_{02} - \sqrt{(\mu_{20} - \mu_{02})^2 + 4 * \mu_{11}^2}} \tag{A.18}$$