Tampereen yliopisto

Jere Aho

ZOOM BEHAVIOR IN MULTI-CAMERA MOBILE PHONE VIDEOS

Faculty of Information Technology and Communication Sciences
Bachelor's thesis
June 2021

TIIVISTELMÄ

Jere Aho: Zoomauksen käyttäytyminen usean kameran älypuhelimen videoissa
Kandidaatintyö
Tampereen yliopisto
Tietotekniikka
Kesäkuu 2021

Älypuhelimien kameroilla valokuvaaminen on lisääntynyt huomattavasti viime vuosien aikana, joten on tärkeää arvioida älypuhelimien kameroiden kuva- ja videolaatua. Uusimmissa älypuhelimissa on usein useampi kamera, mikä on tuonut uusia haasteita puhelimien kuvankäsittelyyn. Yksi haasteista on, miten älypuhelin pystyy vaihtamaan tasaisesti yhdestä kamerasta toiseen, käyttäen laitteen zoomaus ominaisuutta, ilman havaittavia häiritseviä muutoksia kuva- ja videolaadussa. Tätä haastetta nimitetään zoomaustasaisuudeksi.

Huawei halusi tutkia zoomaustasaisuutta parantaakseen sitä puhelimissaan, joissa on useampi kamera. Tämä työ tutkii, mitkä kuvanlaatutekijät vaikuttivat negatiivisesti zoomaustasaisuuteen subjektiivisesta ja objektiivisesta näkökulmasta. Tutkimuksessa myös pohditaan, mitkä kuvankäsittelyalgoritmit säätelevät kuvanlaatutekijöitä, jotka toimivat epämiellyttävästi. Täten saadaan tietoa kuvanlaatualgoritmeista, joita kehittämällä zoomaustasaisuutta voitaisiin parantaa.

Kuvanlaatutekijöiden mittaamiseksi suunniteltiin toistettava video kaappaus ympäristö, jossa voidaan ottaa kameranvaihtovideoita. Sitten kameranvaihtovideoita analysoitiin tarkemmin kohdissa, joissa kuvanlaatu muuttuu huomattavasti subjektiivisesti ja objektiivisesti. Työ myös tarkastelee, ovatko kuvanlaatutekijöiden objektiivisten mittareiden muutokset havaittavissa vastaavina subjektiivisina muutoksina.

Tässä työssä osoittautui, että selkeät muutokset kirkkaudessa ja värilämpötilassa vaikuttavat negatiivisesti zoomaustasaisuuteen. Älypuhelimessa näitä kuvanlaatutekijöitä kontrolloivat erilliset algoritmit: automaattivalkotasapaino ja automaattivalotus. Tämä tutkimus näyttää myös, että kirkkauden ja värilämpötilan objektiivisissa mittareissa havaitut muutokset vastaavat subjektiivisia havaittuja muutoksia.

Avainsanat: Zoomaus, kamera, älypuhelin, kameran testaus, kuvanlaatu, videolaatu, zoomaustasaisuus

ABSTRACT

Jere Aho: Zoom behavior in multi-camera mobile phone videos
Bachelor’s thesis
Tampere University
Computing Sciences
June 2021

Due to the growing demand for mobile phone consumer photography, it is important to evaluate mobile phones’ image and video quality performance. Modern multi-camera mobile phones have brought along new image processing challenges. One of the challenges is how a multi-camera mobile phone can switch from one camera to another camera smoothly, when using its zoom capability, without unpleasant perceptual changes. This challenge is referred to as zoom smoothness.

Huawei wanted to research zoom smoothness since it wants to improve the zoom smoothness quality of its multi-camera mobile phones. To help Huawei accomplish its goal, this research considers what image quality factors subjectively and objectively hurt zoom smoothness. The research also identifies which image processing algorithms control those image quality factors. These algorithms can then be improved to enhance the quality of zoom smoothness.

To measure the image quality factors, this research presents a repeatable video capturing test environment for capturing camera switch videos. The captured videos were analyzed in parts where image quality significantly changed subjectively and objectively. The research also considers whether any image quality factors’ objective metric changes correspond to subjectively perceived changes.

This research found that the brightness and white balance image quality factors hurt zoom smoothness. These image quality factors are controlled by the auto white balance and the auto-exposure image processing algorithms, respectively. This research also shows that changes in the brightness and the white balance objective metrics correspond with the subjectively perceived changes.

Keywords: Zoom, camera, mobile phone, camera testing, image quality, video quality, zoom smoothness

PREFACE

This paper was created as a part of the Tampere University signal processing bachelor's thesis seminar. I am doing this research and development for zoom smoothness in the image quality department at Huawei Technologies Finland. The more I researched zoom smoothness, the more I got excited about it. It was also a unique opportunity to gain more knowledge on image and video quality which I'm thankful for. For this opportunity, I would like to thank my supervisor and colleagues for their encouragement and help along the way. I would also like to thank my family and friends for supporting me in this endeavor.

At Tampere, 29.06.2020

Jere Aho

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

ABBREVIATIONS AND MARKINGS

ADB	Android Debug Bridge. A command-line tool for controlling mobile phones.
AE	Auto Exposure. An abbreviation for an image processing algorithm.
API	Application Programming Interface. A term used in software engineering.
AWB	Auto white balance. An abbreviation for an image processing algorithm.
CFA	Color filter array. A part of a camera module
CMOS	Complementary metal-oxide-semiconductor. A part of a camera module.
DxoMark	An image and video quality evaluation company.
FFmpeg	Fast Forward MPEG. A video manipulation tool.
HTTP	Hypertext Transfer Protocol. A term used in computer networking.
HVS	Human visual system. A term used in signal processing.
IPP	Image Processing Pipeline. A set of cascaded image processing algorithms that manipulate an image.
JPEG	Joint Photographic Expert Group., An image file format standard.
LED	Light-emitting diode. A part of an illuminant.
MATLAB	Matrix Laboratory. The name of a programming language.
MPEG	Moving Picture Experts Group. A video file format standard.
Python	The name of a programming language.
RAW	A term for an unmodified image which is the initial output of an image sensor.
RGB	Red, green and blue. A color space in signal processing
Lx	Lux
K	Kelvin

1. INTRODUCTION

Mobile phones are ever more prevalent devices in everyday life for capturing images and videos. This means that image and video quality are very important. It is especially relevant for mobile phone manufacturers who want to gain a competitive edge over their competitors by developing the best mobile phone cameras. Research institutions and mobile phone manufacturers have researched and developed different methodologies and standards for measuring the image and video quality of a mobile phone with a single camera. Many modern mobile phones these days often have multiple different cameras. It is important to analyze how multi-camera setups affect video quality when using the zoom capability of a mobile phone.

The purpose of this research was to create objective metrics based on subjective findings for measuring and understanding the zoom smoothness of mobile phones with multiple cameras. Zoom smoothness means how well the mobile phone's image and video quality behaves when changing the zoom ratio of a mobile phone over a specific zoom ratio value. It is important to analyze and identify what negatively impacts zoom smoothness with subjective and objective analysis of image and video quality.

This research first goes through the basics of how a mobile phone camera works. Then, some relevant literature was reviewed which identifies unpleasant camera transition behavior which negatively impacts zoom smoothness. Then these unpleasant behaviors, found in the literature, were confirmed through the subjective analysis of captured videos. Then, to further objectively and subjectively measure zoom smoothness, an automated and repeatable video capturing test environment was introduced. Videos were captured from the test environment and then further subjectively analyzed to validate if the initial subjective findings were still present. Finally, based on the subjective analysis findings, some objective metrics were proposed as well as analyzed to see if they correlate with the subjective findings.

2. MOBILE PHONE CAMERA BASICS

A modern mobile phone camera consists of the hardware and software: the camera module and the image processing pipeline (IPP), respectively [1, p. 9]. Together these parts attempt to create an accurate and visually pleasing image. This chapter goes through how a generic camera module and the IPP work. Also, the differences between camera modules and the zoom capabilities of a mobile phone are discussed.

2.1 The camera module

This chapter covers how the different components of a camera module work together to create an image. Differences between camera modules and their different use cases will be covered as well.

2.1.1 Camera module parts

The purpose of a camera module is to capture information of light to form an image. Light consists of photons that have different characteristics such as intensity and wavelength [2]. Figure 1 illustrates how an image is captured with a simplified structure of a camera module. The structure consists of the following components: lenses, a color filter array (CFA) and an image sensor.

To capture an image, light first travels through a lens cover, then lenses that focus the light onto an image sensor which measures the light's intensity. The measurements are then digitized into a digital signal which represents the initial version of the image also known as a RAW image. [1, p. 10] Next, this chapter goes into further detail about the characteristics of each camera module component.

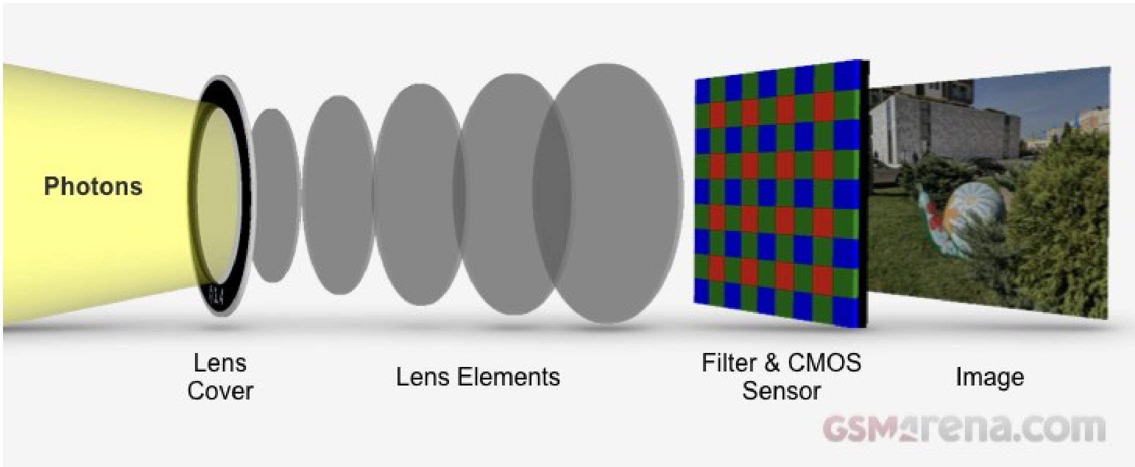


Figure 1: Components of an image sensor [3]

The primary objective of lenses is to focus incoming light onto an image sensor. To achieve this, a modern mobile phone lens system consists of multiple lenses that vary in size, curvature and focal length [1, p. 11]. By using the different properties of the lenses the incoming light can be collected and focused accurately onto the image sensor [4, p. 64].

Before the light reaches the image sensor, the focused light goes through the CFA. The purpose of a CFA is to separate light into different color channels. The CFA forms a matrix plane of squares which represent pixels. Each pixel only lets through one color channel of light by bandpass filtering wavelengths of light that are not in the pixel's color channel band. Figure 2 illustrates different CFA arrangements. A commonly used CFA is the Bayer CFA [5, p. 11] which separates the incoming light into red, green and blue color channels. The Bayer CFA has the property of having the same number of green pixels as the number red and blue pixels combined. This property corresponds well with the human visual system (HVS) which is most color sensitive to green light [1, p. 10] [5, p. 11].

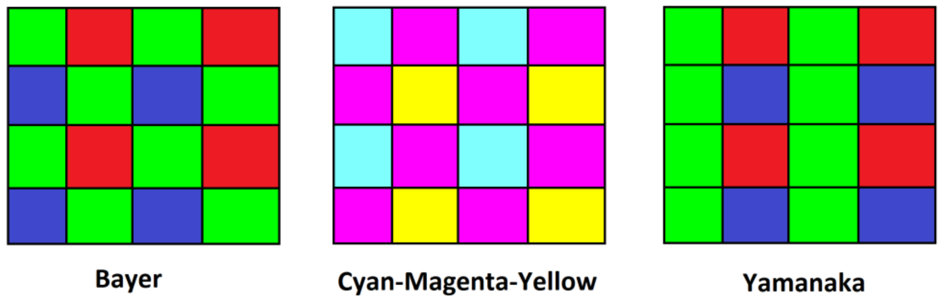


Figure 2: Different color filter array arrangements

Finally, the focused and color-filtered light reaches the final part of the camera, the image sensor. The most commonly used image sensor in mobile phones is the CMOS sensor [1, p. 11]. The CMOS sensor consists of photodiodes that form a pixel matrix like the CFA. Since light consists of photons, each photodiode is exposed to incoming photons for a short period of time. During this short period of time, the incoming photons create a charge on the photodiode which can then be measured as a voltage. The measurements are then digitized and combined into the first version of the image, commonly known as a RAW image.

2.1.2 Camera module types

A modern mobile phone often has multiple camera modules. Different camera modules can be classified by their different properties into the following categories: wide, main, and tele. Table 1 shows the specifications of the rear cameras of the ”Huawei Mate 40 Pro+” mobile phone. From the table, it is evident that each camera module type differs in aperture, focal length, and megapixel count. Each camera module fits a different use case. The wide camera module is great for landscape photography since it has a wide field of view. On the other hand, the tele camera module has a large focal length which is good for taking images of objects far away.

Table 1: The specifications of the Huawei Mate 40 Pro+ rear cameras [6]

Camera module	Focal length (mm)	Aperture	Megapixel count (MP)
Main camera	23	f/1.9	50
Tele camera 1	-	f/2.4	12
Tele camera 2	240	f/4.4	14
Wide camera	14	f/2.4	20

2.2 Image processing pipeline

The IPP attempts to enhance a RAW image, the output of the camera module. This is achieved by applying a combination of different image processing algorithms to the RAW image.

The IPP consists of adjustable algorithms. The algorithms within an IPP form a chain where each algorithm takes an input image that has been adjusted by a previous algorithm. This cascading property makes algorithms at the beginning of the IPP harder to adjust since adjustments made to an algorithm will affect all the following algorithms in the IPP. [1, p. 13]

The IPP is also a computationally heavy task [1, p. 12] and mobile phones are limited in computational power and battery capacity, so the algorithms need to be computationally efficient. Mobile phone cameras have small pixel sizes which limit their photosensitivity and therefore decrease the image quality [1, p. 12]. These facts make designing and adjustment of the IPP very difficult [1, p. 13].

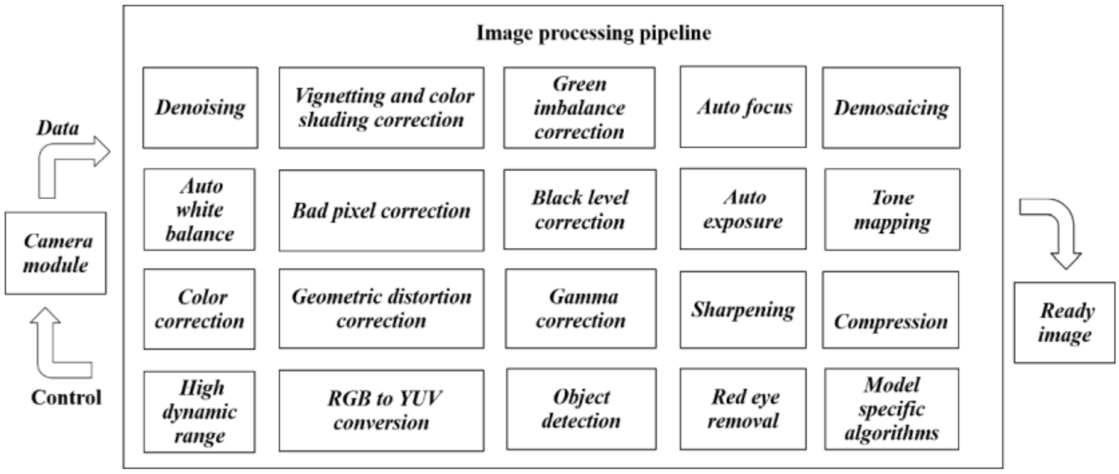


Figure 3: Generic image processing pipeline [1, p. 13]

There is no standard way of creating an IPP. Different corporations and institutions do image processing slightly differently. Generally, the IPP can be sliced into 3 tasks: correction, conversion and controlling [1, p. 13]. Each of these tasks is usually implemented via a combination of image processing algorithms. Some common image processing algorithms are highlighted in Figure 3. Next, some of the conversion and controlling algorithms are introduced in more detail.

The initial RAW image data consists of pixel values that have one color channel value. Since the Bayer CFA is commonly used, each pixel's color channel is either red, green or blue. The RAW image is converted into a more presentable RGB image by applying a demosaicing algorithm. When the Bayer CFA is used, the demosaicing algorithm interpolates the two missing color channel values for each pixel. [5, p. 11] Many different demosaicing algorithms do the interpolation differently. Each demosaicing algorithm may produce artifacts in the image of varying severity [4, pp. 52-53].

The auto-exposure (AE) and the auto white balance (AWB) are both controlling algorithms. [4, p. 101]. The AE algorithm mimics the human visual system's dark adaptation behavior [7]. Since the mobile phone cameras have fixed aperture, the AE algorithm controls the brightness of an image by adjusting the exposure time and gain camera parameters according to the light level of a scene [4, p. 101] [8, p. 12]. The AWB algorithm mimics the human visual system's (HVS) chromatic adaptation behavior [5, p. 275]. The AWB algorithm tries to identify the correct white balance setting from the illumination of the captured image. Then the image's red, green and blue color channel signal gains are modified according to the white balance setting. Figure 4 shows some common white balance settings for a camera: tungsten, fluorescent, flash, cloudy, shade and daylight. [8, p. 9] Each of these is meant to be used in the lighting conditions indicated by the white balance setting name to achieve the correct color cast for the image. Therefore, in Figure 4, the correct white balance setting would be cloudy. [9]

Figure 4: Common white balances at different color temperatures [9]

Finally, once the image or video has been processed by the IPP, it needs to be stored. The processed file will be compressed to save space in file storage. The image or video will be commonly stored in a lossy file format such as JPEG or MPEG, respectively. A lossy image file format does not retain information that is redundant to the human visual system (HVS) [5, pp. 319-320]. This means that the information loss in the image or video is not usually perceivable to the human eye.

2.3 Zoom

Zoom is the capability of a camera to capture sharp images of objects at varying distances. There are two types of zoom, digital and optical zoom. The latter is much better at retaining the image quality.

Digital zoom is a software-based solution that is applied to an image after an image is captured. In it, a certain area of the image is magnified and upscaled to the original resolution of the image. The upscaling is done by interpolating new pixel values from existing ones. Digital zoom results may look blurred or may even contain some interpolation artifacts. [5, p. 276]

Optical zoom is an optics-based solution for zooming. In it, the focal length of the camera is changed by adjusting the distance between the lenses by moving the lenses. Optical zoom retains image quality much better than digital zoom since no interpolation is needed [5, p. 276]. In the past, mobile phone cameras have mostly utilized digital zoom since larger zoom lenses have been harder to fit within a mobile phone due to physical space constraints. The larger zoom lenses would enable a larger focal length, which helps with capturing sharper images with larger zoom ratios.

The zoom lens technology in mobile phones has advanced quite a bit in recent years. Figure 5 showcases one advancement, the periscope telephoto camera. In it the zoom lenses are fitted horizontally into a mobile phone and light is guided via a prism to the lenses and image sensor. The periscope telephoto camera enables a longer and adjustable optical zoom range capability for a mobile phone. [10]

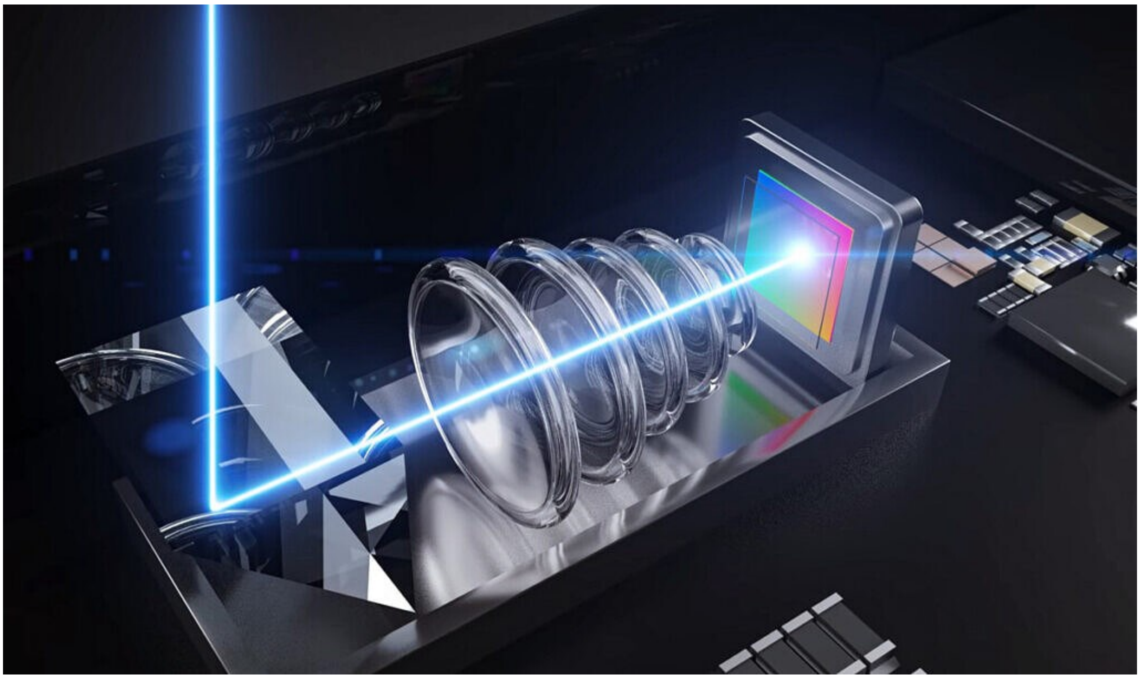


Figure 5: A periscope telephoto camera module [10]

3. PROBLEM FORMULATION

Mobile phones with multiple camera modules have not existed for that long, so there is not much research that has studied zoom smoothness. This chapter will describe how camera transitions can negatively impact zoom smoothness by reviewing relevant work relating to the topic. Then a testing procedure to further research zoom smoothness will be introduced.

3.1 Problem formulation and relevant related work

One of the first mobile phones that debuted with multiple cameras was the HTC One M8 back in 2014 [11]. Since then, many other mobile phone manufacturers have gradually started adopting multiple camera setups as Apple did in 2016 in the iPhone 7 Plus. [12] This multi-camera trend has brought along new image processing challenges, one of which is zoom smoothness. Zoom smoothness considers how the mobile phone’s zoom capability should works with multiple cameras, especially when transitioning from one camera to another. A camera transition happens when a mobile phone switches from one camera to another camera while the zoom ratio is incremented or decremented over a specific zoom ratio value.

There is not much research related to the topic of zoom smoothness. One company called DxoMark has studied zoom smoothness on mobile phones with multiple cameras. DxoMark is a company that benchmarks mobile phone cameras by evaluating their image and video quality. [4, p. 322] According to them, for a camera transition to be subjectively smooth, the perceived brightness and white balance must be similar between two cameras [12].

When a camera transition happens, the IPP must quickly set up the camera parameters for the camera that is being transitioned to. Important camera parameters include white balance and brightness. The AWB and AE algorithms adjust these camera parameters. During a camera transition, if the IPP does not use the previous camera parameters before the camera transition happened, then the IPP will have to set up the camera parameters from scratch. If this is the case, then undesirable white balance, color or brightness change might appear in the video.

3.2 Camera testing procedure for this research

In this research, the main focus is observing the zoom smoothness of transitions between two cameras. One goal is to subjectively and objectively identify what image processing algorithms negatively affect the zoom smoothness during a camera transition. Next, a camera testing procedure is introduced for measuring zoom smoothness.

Camera testing is a crucial part of mobile phone camera research and development. Figure 6 shows how this research was conducted. The first part is capturing videos from different environments with different lighting using the zoom capability of a mobile phone. The captured videos are then subjectively analyzed to see if they have unpleasant behaviors that impact zoom smoothness. An automated and repeatable video capturing environment is then created based on the found undesired behaviors. Once created, videos are captured from it. The captured videos are subjectively analyzed to see if the initially observed unpleasant behaviors are still present. If so, then objective metrics will be proposed to measure subjectively undesired behaviors. Finally, the objective metrics' graphs are analyzed to see if they have features that correspond with the observed undesired behavior.

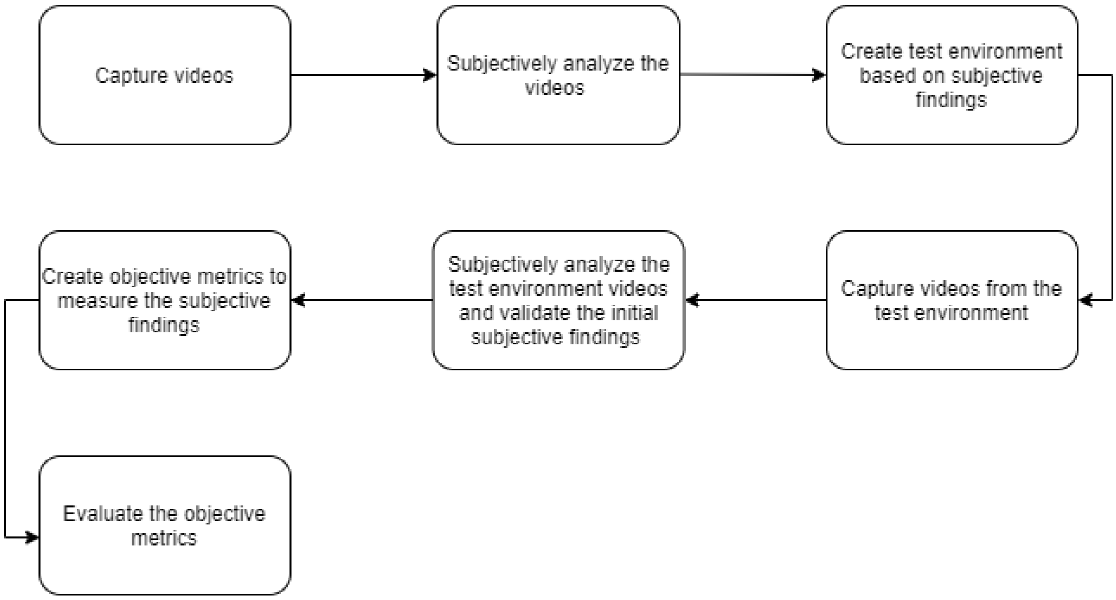


Figure 6: Zoom smoothness camera testing process

4. SUBJECTIVE DATA ANALYSIS AND TEST ENVIRONMENT

In this chapter, we will check if DxoMark's zoom smoothness findings related to camera transition brightness and white balance were present in videos. First, some videos were recorded using the zoom capability of a mobile phone and were analyzed to see if they have any undesired behavior. Based on the subjective analysis, a repeatable and automated video capturing test environment was created. Finally, more videos were captured from the video capturing test environment and subjectively analyzed in more detail.

4.1 Initial observations of subjectively unpleasing behaviors

When considering zoom smoothness, we wanted to understand what degrades the quality of the zooming experience the most. To better understand this, some videos were recorded utilizing the zoom capabilities of the mobile phone. The videos were captured from different lighting conditions, indoors and outdoors. Then, the videos were subjectively analyzed to find the most subjectively unpleasant zoom behaviors.

The recorded videos utilized the full range of the zoom. When analyzing the videos, for the most part, they did not have many unpleasant behaviors except during camera transitions. According to DxoMark, the white balance should be the same between two cameras before and immediately after a camera transition. This is not the case between two consecutive frames of a camera transition shown in Figure 7. The white balance between them differs perceptually. This perceptual change in white balance is perceived as unpleasant when viewing it in the video. These findings confirm the findings made by DxoMark on what negatively impacts zoom smoothness.

Figure 7: Camera transition between two frames

4.2 Video capturing test environment

One of the objectives of this research was to make the video capturing test environment as repeatable and automated as possible. The repeatability ensures that experiments can be conducted in the future again for new devices. The automation ensures that the data collected from the test environment is as easy as possible and reduces the risk of user errors. The following chapters will go through the video capturing test environment and software automation solutions that were designed with the repeatability and automation objective in mind.

4.2.1 Automated software and test environment details

Mobile phones have different user interfaces and software implementations for controlling the cameras. To make the testing repeatable, an automated touch simulation control software was created to control the phones. The software tools that were used to implement this include the programming language, Python, and the command-line tool, android debug bridge (ADB). The android debug bridge enables communicating and sending individual commands to control the phone. [13] The ADB commands can simulate touch, unlock the screen and open applications, like the camera app on the mobile phone. The different commands were then chained into a sequence of commands with Python scripts which were used to control the mobile phone.

Figure 9 shows the test environment diagram. In this research, the test environment consists of Imatest's SFRplus chart [14], light-emitting diode (LED) illuminants, and a robot that can change the distance between the phone and the chart. The programmatically controlled LED illuminants and robot are both adjustable by sending hypertext transfer protocol (HTTP) requests through an application programming interface (API). The robot position is adjustable vertically and horizontally. It is also rotatability of the two axes helps to align the phone parallel to the chart. The color temperature and intensity of the lights are adjustable. The color temperature range of the lights is from 3000 to 6000 (K) and the intensity range is from 1 to 2000 (lx). Figure 8 showcases the Imatest SFRplus chart. It was chosen since it has useful features that can be measured objectively with software and evaluated subjectively by humans.

Figure 8: Imatest SFRPlus chart

4.2.2 Data collection from the test environment

The data collection procedure consists of a few phases: adjusting illuminants, robot positioning, capturing videos and extracting data from captured videos.

Figure 9 illustrates a diagram of how the illuminants were set up in the test environment. The intensity of the illuminants is set to 1000 (lx) and color temperatures to 6000 (K). In this research, this is the only tested illumination condition. The illuminants were positioned between a 20 to 45 degrees angle from the normal of the chart to ensure uniform illumination [15]. Uniform lighting was verified with lux meter intensity measurements at each corner of the chart and center of the chart. The corner measurements should be within ±5% of the illuminance of the center point to attain accurate color measurements. [16]

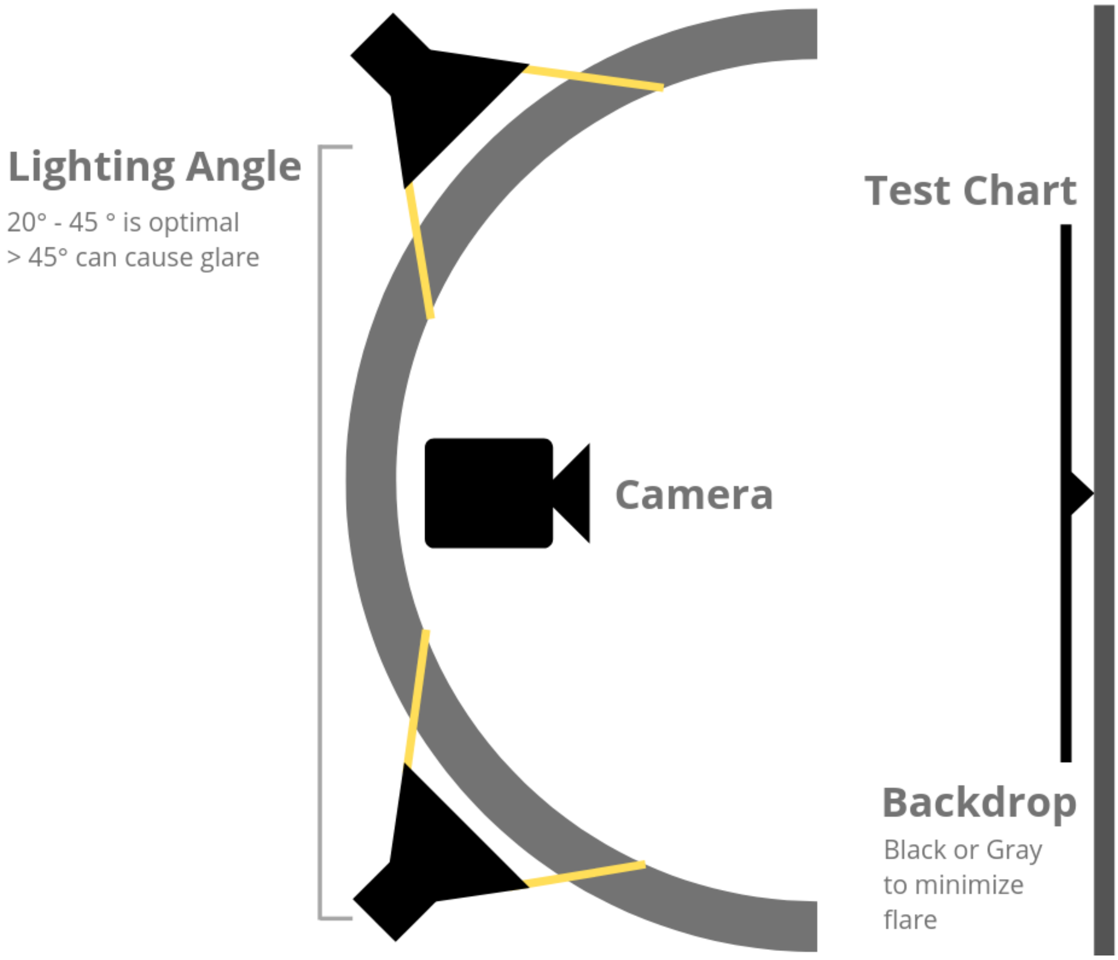


Figure 9: Test environment diagram [16]

Once the illumination has been set up correctly, the mobile phone's optimal position from the chart needs to be figured out according to the camera transition in question. The optimal distance is a distance where the chart stays mostly in full-frame throughout the entire camera transition. Figure 10 illustrates bad framing where the chart is not full-frame and Figure 11 illustrates good framing where the chart is in full-frame. The optimal distance will need to be found through trial and error for each camera transition.

Figure 10: Bad framing

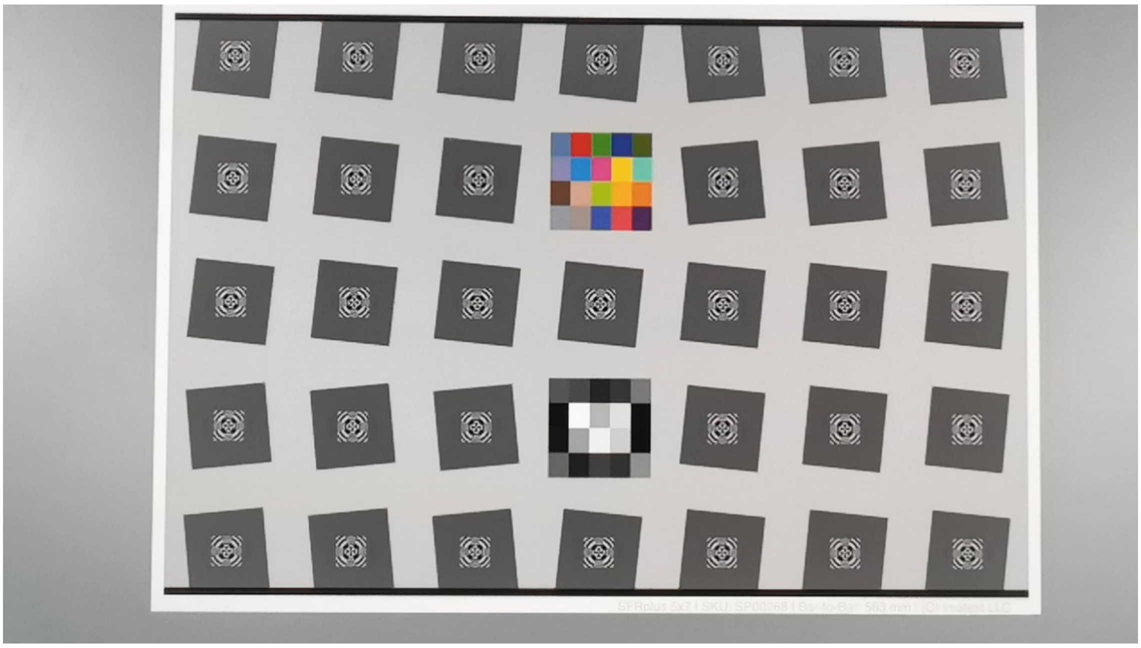


Figure 11: Good framing

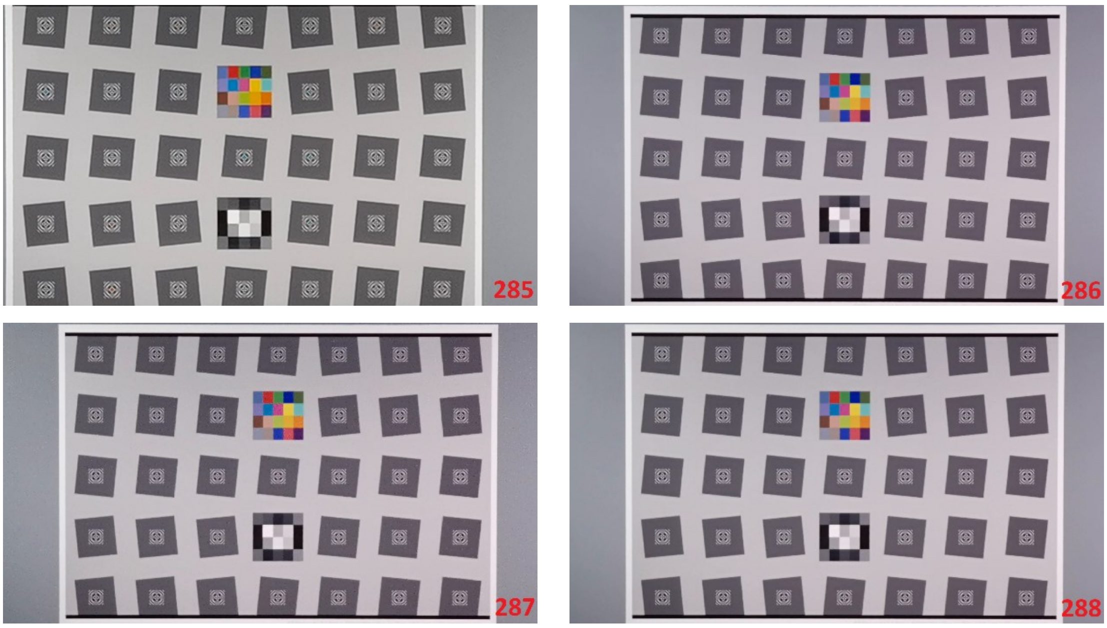
Once the optimal distances for each camera transition have been found, then each camera transition video is recorded. Then, frames are extracted from the recorded videos using a widely used image and video manipulation tool, FFmpeg. Finally, the frames are used in a further objective and subjective image quality data analysis in the following chapter.

4.3 Image quality assessment of the collected data

This chapter will subjectively analyze camera transition videos to find and verify the unpleasant behaviors that were found in the observations of chapter 4.1. In addition, it will also assess which IPP algorithms might cause unpleasant behaviors.

The white balance and brightness image quality factors are controlled by different image processing algorithms AWB and AE, correspondingly. White balance and brightness are perceptually heavily intertwined with each other. This means that it is hard to perceptually differentiate whether white balance or brightness has been adjusted. This is important to keep in mind when trying to identify which algorithm causes the unpleasant behavior.

Figure 12 illustrates the unpleasant behavior in some frames of a camera transition video from the main camera to the tele camera. The first frame is using the main camera and the other frames are using the tele camera. When observing frames 285 and 286, there is a distinct perceptual change. This is consistent with the subjective findings in chapter 4.1. It is hard to say whether the white balance or brightness has changed between the frames. From frame 286 to frame 290, there is a perceptual shift that begins to happen in each frame. Each frame begins looking perceptually more and more like frame 285. When frame 290 is reached, it begins to look quite similar to frame 285. This means that either the AWB algorithm, the AE algorithms or both have begun converging to the correct white balance or brightness values. When frames 300 and 310 are reached, they both look perceptually very similar to frame 290. This seems to subjectively indicate that either the AWB algorithm, AE algorithm or both have converged and found stable white balance value, brightness value or both, respectively.



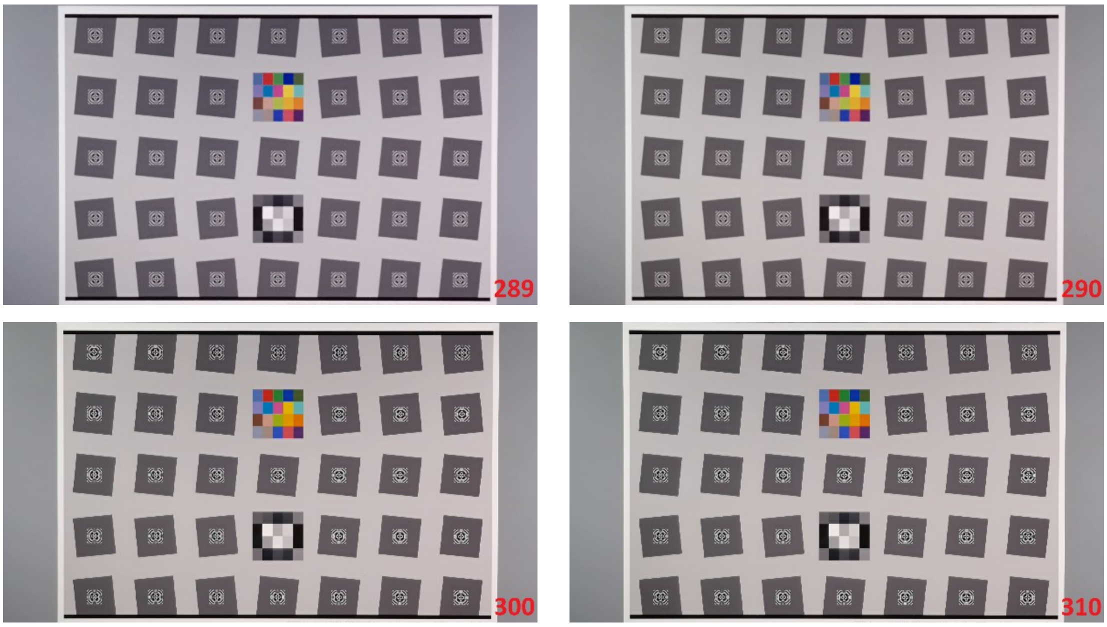


Figure 12: Some frames from a main to tele camera transition video. The numbers indicate the frame numbers. Frame 285 is using the main camera using the main and the rest of the frames are using the tele camera.

When analyzing the subjective findings it is hard to say which image quality factor brightness or white balance is the main culprit for the unpleasant behaviors. Therefore, it is important to objectively identify which algorithm might be causing more unpleasant behaviors during a camera transition. The next chapter will objectively analyze two image quality factors, white balance and brightness.

5. OBJECTIVE METRICS

Based on the subjective findings some objective metrics are proposed to objectively analyze the camera transitions. This chapter will cover what objective metrics were chosen to understand the changes in white balance and the brightness and how the objective metrics are calculated. Finally, the objective metrics will be plotted into graphs, which are then analyzed.

5.1 Objective metrics calculations

The objective metrics are calculated with MATLAB scripts. The scripts use object detection to track different objects from the SFRplus chart. Currently, the objects that are tracked are the gray areas highlighted in Figure 13. The script applies object tracking to each frame of a camera transition video. Once the tracking is done, features are extracted out of the tracked objects to calculate the objective metrics. This chapter will go through the objective metrics which are used to measure the image quality factors, brightness and white balance.

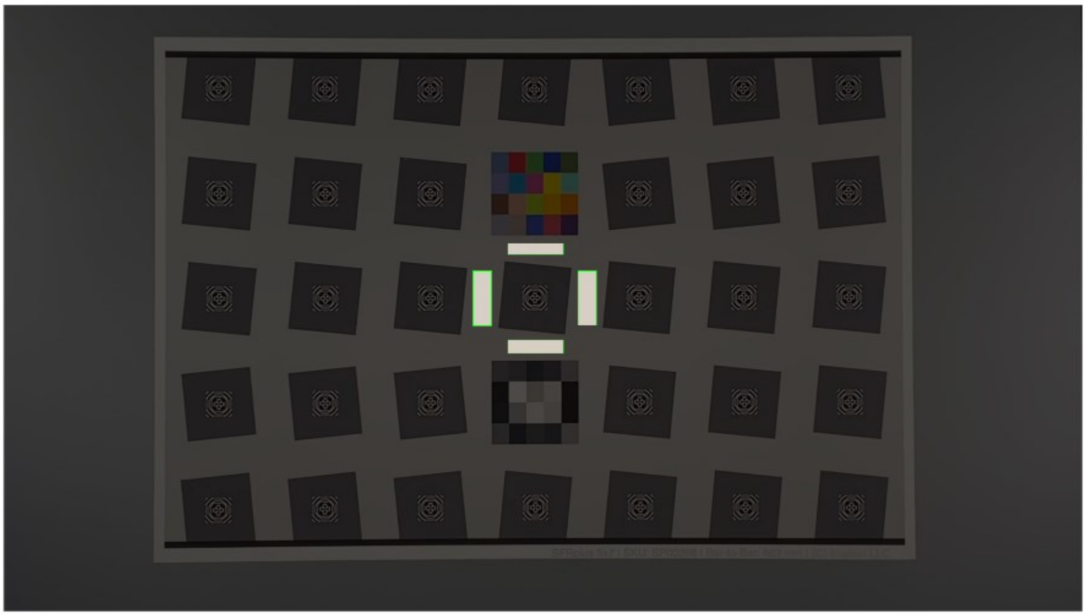


Figure 13: Gray patches detected in the SFRplus chart using object detection

5.1.1 White balance metric calculation

To measure how white balance changes through the video, the recovery angular error objective metric is calculated for each frame of the video. It is preferred to calculate white balance from a chart area which is a spectrally neutral and uniform gray area. [4,

p. 251] For this reason, the recovery angular error is calculated from the 4 gray patches in the center of the SFRplus chart which are highlighted in Figure 13.

The recovery angular error is calculated with the following formula (2)

$$err_{recovery} = arccos\left(\frac{\rho^{E} \cdot \rho^{Est}}{\|\rho^{E}\|\|\rho^{Est}\|}\right), (2)$$

where ρ^{E} is a vector of the measured RGB intensity values and ρ^{Est} is a vector of the target RGB intensity values. [17] In this case, the target RGB intensity values ρ^{Est} are all the same since they are achromatic. This means that the target RGB values can be any arbitrary values as long as they are all the same value. Using this knowledge, the formula can further be simplified into the form,

$$err_{recovery} = arccos\left(\frac{\mathrm{R \cdot G \cdot B}}{\sqrt{3 \cdot (\mathrm{R+G+B})}}\right) , (3)$$

where R, G and B are the measured red, green and blue intensity values. This shows that target intensity values of the gray patches do not need to be known if the target RGB intensity values are the same.

5.1.2 Brightness metric calculation

To measure how brightness behaves in the video, the mean lightness value of the CIE LAB color space is calculated for each frame of the video from the 4 gray patches. To obtain this lightness value, first, the linear RGB pixel values are converted with a non-linear transformation into the sRGB color space. The obtained sRGB values are then transformed into the tristimulus values of the CIE XYZ color space and finally, the tristimulus values are transformed into the CIE LAB color space [4, p. 181]. The CIE LAB color space is perceptually correlated with the human vision which was the main reason for choosing its lightness value to be the brightness metric [4, p. 181]. In this research, the brightness metric is calculated from the 4 gray patches highlighted in Figure 13, since it is a spectrally neutral and uniform gray area.

5.2 Objective metrics' results

This chapter will analyze changes in graphs of the proposed white balance and brightness metrics. The previously discussed ideas of an ideal camera change in chapter 3.1 will be used to analyze the graphs. When analyzing the graphs the main objective is not to pay attention to the absolute values of the different metrics, but to see if the subjectively unpleasant behaviors are objectively visible as changes in the graphs during camera transitions.

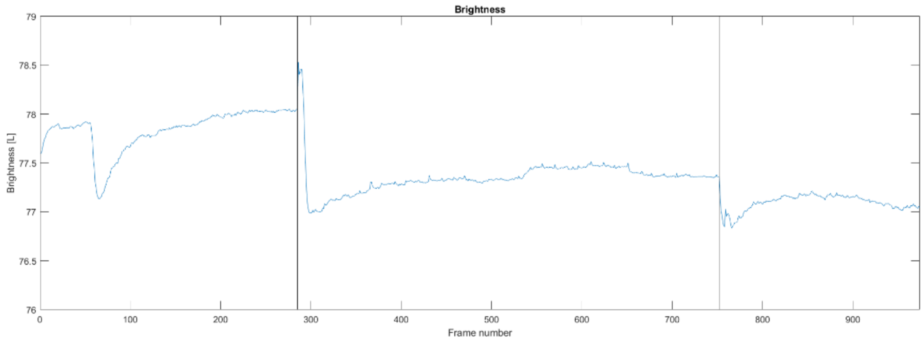


Figure 14: Brightness objective metric graph where the mobile phone first changes from the wide to the main and then back to the main

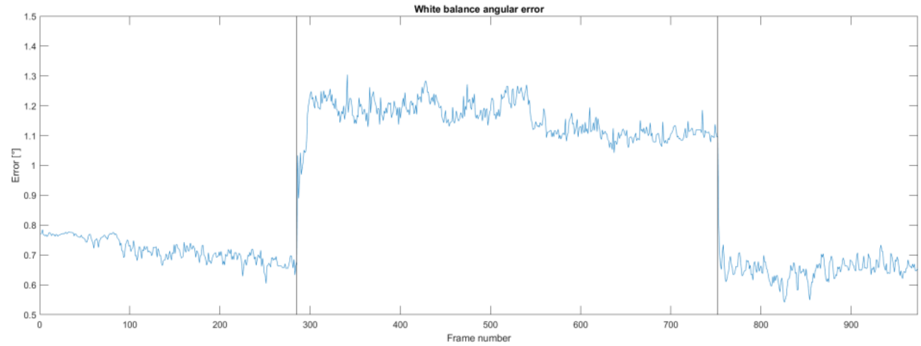


Figure 15: White balance objective metric graph where the mobile phone first changes from the wide to the main and then back to the main

The graphs are shown in Figure 14 and Figure 15 illustrate the change in brightness over time and change in white balance over time, respectively. In the graphs, the mobile phone first switches from the main camera to the tele camera and then back to the main camera. The black vertical lines in the graphs indicate the camera transition points. The second camera transition, where the mobile phone switches from the main camera to the tele camera, will not be considered in this research. Next, the first camera transition, where the mobile phone switches from the tele camera to the main camera, will be analyzed.

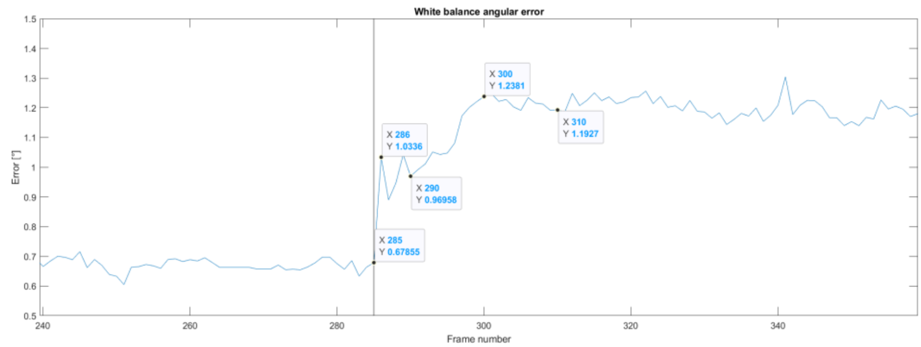


Figure 16: Brightness graph for the first camera transition point

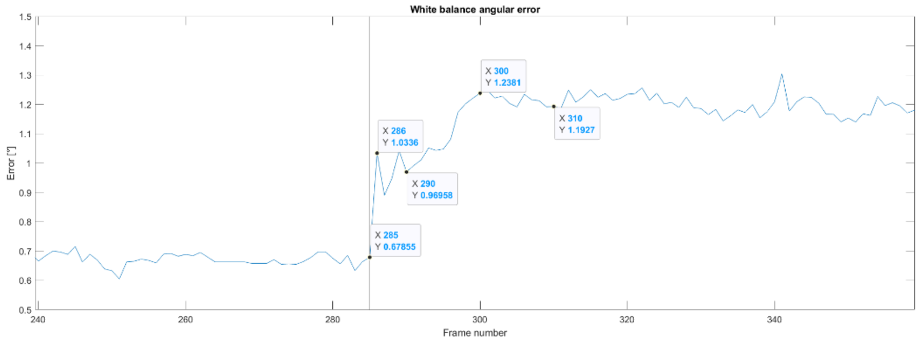


Figure 17: White balance graph for the first camera transition point

Figure 16 and Figure 17 are zoomed-in versions of Figure 14 and Figure 15, which show the first camera transition point. Both Figure 16 and Figure 17 have highlighted points that correspond to some of the frames shown in Figure 12. Next, graphs are analyzed in more detail in the frame range from 285 to 310.

From frame 285 to 286, in Figure 16 and Figure 17, there is a clear visible spike upwards in brightness and white balance. The frames from 286 to 290 have similar fluctuation patterns in white balance and brightness values. From frame 290 to 300, there is a sharp decrease in brightness and a sharp increase in white balance. The frames 300 and 310 have very similar values, which is an indication that the brightness and white balance values have converged and stabilized. These observations indicate that previously observed subjective findings of unpleasant behavior visible within objective metrics. From the graphs, it seems that both brightness and white balance are linked to unpleasant behaviors, but it is hard to pinpoint which one is the larger culprit for the unpleasant behavior during the camera transition.

6. CONCLUSION

This research sought out to figure out what caused the unpleasant behavior when zooming with a multi-camera phone. The main focus of the research became observing camera transitions since they had the most unpleasant behavior. The camera transitions were then further analyzed by identifying what image quality attributes subjectively and objectively negatively affect the zoom smoothness of the camera transitions. An automated and repeatable video capturing test environment was created to measure the image quality attributes. Videos were captured from the video capturing test environment. The videos were then subjectively and objectively analyzed. The analysis shows that the brightness and white balance changes are subjectively and objectively linked with the unpleasant camera transition behavior.

The findings in this research are useful for mobile camera manufacturers that are designing an IPP for mobile phones with multiple cameras. Overall this thesis lays a foundation for future zoom smoothness research. Future research could include measuring the center point movement during camera transitions or other image quality attributes like sharpness, noise, shading and texture and then identifying whether any of them negatively affects zoom smoothness during camera transitions. Other illumination conditions could be considered as well in the future.

REFERENCES

[1] V. T. Peltoketo, Benchmarking of mobile phones cameras, 2016.

[2] T. Kurado, Essential Principles of Image Sensors, CRC press, 2017, p. 2017.

[3] GSMArena, "Camera definition," [Online]. Available: https://m.gsmarena.com/glossary.php3?term=camera. [Accessed 31 October 2020].

[4] J. B. Phillips and H. Eliasson, Camera image quality benchmarking, John Wiley & Sons, 2018.

[5] E. Allen and S. Triantaphillidou, The Manual of Photography, 10th Oxford ed., Elsevier Science, 2011.

[6] GSMArena, "Huawei Mate 40 Pro+ technical specifications," [Online]. Available: https://www.gsmarena.com/huawei_mate_40_pro+-10550.php. [Accessed 2 November 2020].

[7] T. Lamb and J. E. Pugh, "Dark adaptation and the retinoid cycle of vision," in *Progress in retinal and eye research*, 2004, pp. 307-380.

[8] M. Delbracio, D. Kelly, M. Brown and P. Milanfar, "Mobile Computational Photography," *A Tour,* 2021 .

[9] DPMag, "Digital Photo Magazine," [Online]. Available: https://www.dpmag.com/how-to/shooting/what-is-white-balance-on-your-camera/. [Accessed 28 June 2021].

[10] Xda-developers, "Periscope telephoto camera," [Online]. Available: https://www.xda-developers.com/periscope-camera-vs-telephoto-camera/. [Accessed 27 January 2021].

[11] GadgetsNow, "Little known facts about dual cameras in smart phones," [Online]. Available: https://www.gadgetsnow.com/slideshows/dual-cameras-in-smartphones-7-little-known-facts/dual-cameras-made-debut-in-2011/photolist/56853404.cms. [Accessed 2 February 2021].

[12] DxoMark, "Multi-camera smartphones: Benefits and challenges," [Online]. Available: https://www.dxomark.com/multi-camera-smartphones-benefits-and-challenges/. [Accessed 12 January 2020].

[13] Android Debug Bridge, [Online]. Available: https://developer.android.com/studio/command-line/adb. [Accessed 15 January 2021].

[14] Imatest , "SRFplus chart," [Online]. Available: https://www.imatest.com/products/test-charts-sfrplus/. [Accessed 11 November 2020].

[15] Imatest, "SFRplus chart usage instructions," [Online]. Available: https://www.imatest.com/docs/sfrplus_instructions/. [Accessed 15 January 2020].

[16] Imatest, "Test Lab Setup," [Online]. Available: https://www.imatest.com/solutions/test-lab-setup/#Lighting. [Accessed 2 February 2021].

[17] G. D. Finlayson, R. Zakizadeh and A. Gijsenij, "The reproduction angular error for evaluating the performance of illuminant estimation algorithms.," IEEE transactions on pattern analysis and machine intelligence, 2016, pp. 1482-1488.